

---

**Cryo-EM, Biochemical, and Computational Analysis  
of *S. cerevisiae* Cell Extracts**  
*Insights into the Endogenous L-A Virus Structure*

---

Dissertation

zur Erlangung

des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität

Halle-Wittenberg,

vorgelegt

von Frau Lisa Schmidt

eingereicht am: 28.07.2023

verteidigt am: 19.12.2023

Gutachter\*innen

1. Jun.-Prof. Dr. Panagiotis L. Kastiris
2. Prof. Dr. Andrea Sinz
3. Prof. Dr. Petra Wendler

Dedicated to pressing F1  
for help

# Table of Contents

<b>List of Figures .....</b>	<b>4</b>
<b>List of Tables .....</b>	<b>5</b>
<b>List of Abbreviations .....</b>	<b>6</b>
<b>1. Introduction .....</b>	<b>8</b>
<b>1.1 Yeast as a model organism for molecular biology and biochemistry .....</b>	<b>8</b>
<b>1.2 Cell lysis to study cellular architecture – Strengths and Limitations .....</b>	<b>9</b>
1.2.1 Isolation of organelles and <i>in vitro</i> function .....	10
1.2.2 Cellular protein complexes and most abundant biomolecules .....	11
<b>1.3 Probing the structure and function of native cell extract with modern biochemical, biophysical and bioinformatic methods .....</b>	<b>12</b>
1.3.1 Cryo-Electron Microscopy and other structural methods to probe the architecture of proteins and their complexes at high resolution .....	16
1.3.2 Focus on the potential of Cryo-EM .....	18
<b>1.4 Cryo-EM and virus structure .....</b>	<b>22</b>
1.4.1 General aspects on viruses with a focus on RNA viruses .....	23
1.4.2 Viruses and protein disorder .....	24
1.4.3 Endogenous viruses .....	25
1.4.4 dsRNA viruses .....	26
<b>1.5 The helper/killer system in yeast .....</b>	<b>27</b>
1.5.1 Structural studies on yeast L-A and L-BC viruses .....	29
1.5.1.1 The crystallographic structure of the yeast L-A virus at 3.4 Å .....	30
1.5.1.2 The cryo-EM structure of the yeast L-BC virus at 2.9 Å .....	32
<b>2. Aim of this thesis .....</b>	<b>35</b>
<b>3. Materials and Methods .....</b>	<b>36</b>
<b>3.1 Materials .....</b>	<b>36</b>
<b>3.2 Methods .....</b>	<b>38</b>
3.2.1 Model Organism culture .....	38
3.2.2 Culture harvest and lysis .....	39
3.2.3 HPLC and Size exclusion chromatography (SEC) .....	40
3.2.4 Sample preparation for cryo-EM .....	41
3.2.5 Data collection of vitrified grids – microscope alignments .....	42
3.2.6 Image processing and analysis .....	43
3.2.7 Activity assay against PDHc .....	45
3.2.8 MS analysis of fractions 5 and 6 sample preparation .....	45
3.2.9 Crosslinking MS of fractions 4 and 5 sample preparation .....	46
3.2.9.1 Identification of proteins .....	47
3.2.9.2 Identification of cross-linked peptides .....	47
3.2.10 Identification of structural homologs .....	47
3.2.11 Western Blot analysis .....	47
3.2.11.1 Chemiluminescence reaction .....	49
3.2.12 High-pressure freezing .....	49
3.2.13 Mitochondria isolation and subfractionation .....	50
3.2.14 MitoTracker Orange staining .....	52
<b>4. Results .....</b>	<b>53</b>
<b>4.1 Recovery and Isolation of yeast mitochondria .....</b>	<b>53</b>

4.1.1 Mitochondrial lysate – validation of a clean preparation .....	54
4.1.2 Mitochondria lysate – observation with cryo-EM .....	56
4.1.3 Retrieving the pyruvate dehydrogenase complex from mitochondria .....	56
<b>4.2 Biochemical and structural analysis of yeast cell extract by single-particle cryo-EM.....</b>	<b>58</b>
4.2.1 The yeast L-A virus is identified in a cryo-EM-accessible eukaryotic cell extract.....	61
To identify the contents of fractions 5 and 6, LFQ-based MS quantification was employed, allowing for the detection and characterization of the L-A virus within these fractions. Figure reproduced and adapted for the Dissertation from Schmidt et al. <sup>219</sup> .....	62
4.2.2. The cryo-electron microscopy (cryo-EM) structure of the L-A virus, obtained at high resolution from a cell extract, reveals intriguing structural and functional adaptations in its native form. ....	63
4.2.3 Mediation of the mRNA decapping site trench shape by two flexible loops – proximity to a potential site for protein-protein interactions .....	70
4.2.4 Tubular Densities Inside the Capsid: Insights into RNA-dependent RNA Polymerase Localization from Asymmetric Reconstruction of the L-A Virus .....	72
4.2.5 Location and molecular modelling of the L-A virus RNA-dependent RNA polymerase .....	74
4.2.6 Higher-order interactions of the L-A virus and minimal communities: Detection and implications .	77
<b>5. Discussion and Outlook .....</b>	<b>83</b>
5.1 Challenges in cryo-EM analysis of mitochondrial extracts .....	83
5.2 On the reproducibility of the biochemical and cryo-EM analysis of yeast cell extracts .....	84
5.2.1 Proteins and pathways within the studied cellular fraction and their inter-relation to the L-A virus	85
5.2.2 Potential role of disorder in the structure, function, and life cycle of the L-A virus .....	86
<b>6. Summary .....</b>	<b>89</b>
<b>7. Zusammenfassung .....</b>	<b>91</b>
<b>8. References .....</b>	<b>93</b>
<b>9. Appendix .....</b>	<b>107</b>
<b>10. Acknowledgements .....</b>	<b>116</b>
<b>11. Resume .....</b>	<b>119</b>
<b>12. Publication list.....</b>	<b>121</b>
<b>13. Affidavit /Eidesstaatliche Erklärung.....</b>	<b>122</b>

## List of Figures

Figure 1. Model organism yeast and its cell structure. ....	9
Figure 2. Cryo-substituted mitochondria and general architecture. ....	11
Figure 3. Statistical data for number of determined structures in the Protein Data Bank ( <a href="http://www.pdb.org">www.pdb.org</a> ).....	14
Figure 4. Statistical data for percentage of determined structures in the Electron Microscopy Data Bank ( <a href="https://www.ebi.ac.uk/emdb/">https://www.ebi.ac.uk/emdb/</a> ). ....	20
Figure 5. Viruses as specimen for the progress of (cryo-)electron microscopy. ....	23
Figure 6. Cryo-EM of the L-A virus. ....	30
Figure 7. The 3.4 Å resolution of the crystallographic Gag-dimer and the potential active site for mRNA decapping. ....	32
Figure 8. Structural analysis of the yeast L-BC virus. ....	34
Figure 9. SEC triplicate for the recovery of fractionated yeast cell extracts. ....	41
Figure 10. Full workflow of native cell extracts from culture to cryo-EM micrographs.....	43
Figure 11. Experiment conducted on the native cell extract fraction to benchmark the crosslinker concentration. ....	46
Figure 12. Schematic of the Western Blot set up. ....	49
Figure 13. Mitochondria isolation via Nycodenz gradient.....	51
Figure 14. Density gradient before and after centrifugation. ....	53
Figure 15. Data supporting a clean mitochondria preparation. ....	55
Figure 16. Discovery of mitochondrial ATP synthase dimers after mitochondria isolation and lysis.....	56
Figure 17. PDHc activity assay of mitochondrial isolation steps to monitor PDHc during mitochondria isolation and lysis. ....	57
Figure 18. Size exclusion chromatography of mitochondrial lysate from the lysed with PCC including buffer shows dissociated protein content. ....	58
Figure 19. Size exclusion chromatography results of yeast cell extract. ....	59
Figure 20. Exploration of large molecular weight fraction using cryo-EM. ....	61
Figure 21. Proteomics of fractions 5 and 6 that were also analyzed by cryo-EM. ....	62
Figure 22. Fourier shell correlation (FSC) for the cryo-EM map of the L-A virus and derived model. ....	63
Figure 23. The cryo-EM map of the L-A virus and the derived model of the asymmetric capsomere.....	64
Figure 24. Discovered cation- $\pi$ interactions stabilizing the L-A virus capsid. ....	65
Figure 25. Persistent densities for the cation- $\pi$ interactions formed by Arg (a), Trp (b), Tyr (c) and Phe (d). ....	65
Figure 26. Fits and conformational variations of X-ray and cryo-EM models to resolved map. .....	67
Figure 27. Structural variation in the cryo-EM structure of the yeast L-A virus.....	68
Figure 28. Energetic calculations of the capsomere interface utilizing HADDOCK. ....	69
Figure 29. The shape of the mRNA decapping site trench is influenced by two flexible loops. .....	70
Figure 30. Identification of structural homologs for the adjacent structural region of the mRNA decapping site using the DALI server.....	71
Figure 31. View of the L-A virus by reconstructing a cryo-EM map at lower resolution by applying a larger pixel size (3.17 Å). ....	72

Figure 32. View inside the yeast L-A virus retrieved from native, endogenous yeast extracts. ....	73
Figure 33. Persistent low-resolution density in the pentameric face pore across L-A virus reconstructions in the 5-fold axis. ....	74
Figure 34. AlphaFold2-predicted Gag-Pol fusion protein. ....	75
Figure 35. Linker region of Gag-pol and the conformational space that can be explored.....	76
Figure 36. Analysis of the Alphafold model of L-A virus RNA polymerase. ....	77
Figure 37. L-A virus proximal biomolecules. ....	78
Figure 38. L-A virus particles in the native cell extract with different assembly states. ....	79
Figure 39. Statistical comparison of all states, as well as empty and full mature L-A viruses.	80
Figure 40 L-A virus communities in native cell extracts and in yeast cells. ....	81
Figure 41. Proposed scheme for the yeast L-A virus life cycle (states I-IX). ....	88

## List of Tables

Table 1. Chemicals and Enzymes used in this study .....	36
Table 2. Antibodies .....	37
Table 3. Instruments and main equipment used in this study .....	37
Table 4. Vitrification conditions .....	42
Table 5. SDS-PAGE gel ingredients.....	48

## List of Abbreviations

Ø	diameter
2D	two dimensional
3D	three dimensional
Å	Ångström
Aa	amino acid(s)
A <sub>280</sub>	Absorbance at 280 nm
AmBic	Ammonium bicarbonate
Acetyl-CoA	acetyl coenzyme A
AP-MS	Affinity purification mass spectrometry
ATP	Adenosine triphosphate
BCKDHc	Branched-chain keto-acid dehydrogenase complex
BSA	Bovine serum albumin
C-ter	carboxyl-terminus, carboxy-terminus, C-terminal tail, C-terminal end, or COOH-terminus
CAPRI	Critical assessment of protein-protein interactions
CASP	Critical assessment of protein structure prediction
CNN's	Convolutional neural networks
CoASH	Coenzyme A
Cryo-CLEM	cryo correlative light and electron microscopy
cryo-EM	cryo-electron microscopy
cryo-ET	Cryo-electron tomography, electron cryo-tomography
CTF	Contrast transfer function
<i>C.thermophilum/Ct.</i>	<i>Chaetomium thermophilum</i>
ddH <sub>2</sub> O	double-distilled water
DTT	Dithiothreitol
E1p/o/b	E1 keto acid dehydrogenase corresponding to pyruvate (p) /2-oxoglutarate (o) /branched-chain (b), with ECs 1.2.4.1, 1.2.4.2, and 1.2.4.4, respectively dihydro-lipoyl (or dihydrolipoamide) transacetylase/transsuccinylase (or acetyl/succinyltransferase)) of pyruvate (p), 2-oxoglutarate (o) and branched-chain (b), with ECs 2.3.1.12, 2.3.1.61, and 2.3.1.168, respectively
E3	Dihydrolipoyl dehydrogenase
E3BP	E3 binding protein
EBI	European Bioinformatics Institute
<i>E. coli</i>	<i>Escherichia coli</i>
EM	Electron microscopy
Exp.	Experiment
FAD	Flavin adenine dinucleotide
FAS	Fatty acid synthase
FPLC	Fast protein liquid chromatography
Fr.	Fraction
FSC	Fourier shell correlation
G	Gauge, an internationally-used scale for sizing needles Generative adversarial networks
HADDOCK	Highly ambiguous data-driven docking
IEX	Ion exchange

LC	Liquid chromatography
LD	lipoyl domain
MD	Molecular dynamics
MDa	Megadalton (1,000,000 Dalton)
MS	Mass spectrometry
microED	Microcrystal electron diffraction
min.	minimum
min	minute(s)
MRC	Multiple random coupling mechanism for oxo-acid dehydrogenase complexes
MW	Molecular weight
N-ter	Amino-terminus, NH <sub>2</sub> -terminus, N-terminal end or amine-terminus
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide hydride
<i>N.crassa</i>	<i>Neurospora crassa</i>
NMR	Nuclear magnetic resonance
NMWL	Nominal molecular weight limit or pore size
NPC	Nuclear Pore Complex
Nup107	Nucleoporin 107
OADHc	Family of $\alpha$ -keto acid/2-oxo-acid dehydrogenase complexes 2
OGDHc	oxoglutarate or $\alpha$ -ketoglutarate dehydrogenase complex
PBS	Phosphate-buffered saline
PC	Positive control
PDB	Protein data bank
PDHc	Pyruvate dehydrogenase complex
PPI	Protein-protein interactions
PSBD	Peripheral subunit binding domain
RMSD	Root-mean-square-deviation
rpm	Rounds per minute
RNN	Recurrent neural network
SAXS	Small-angle X-ray scattering
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
ScV-LA	<i>Saccharomyces cerevisiae</i> L-A helper virus
ScV-LBC	<i>Saccharomyces cerevisiae</i> L-BC helper virus
ScV-M	<i>Saccharomyces cerevisiae</i> L-M Killer virus
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
TAP	Tandem affinity purification
TCA/CAC	Tricarboxylic acid cycle or Krebs cycle or Citric acid cycle
TPP	Thiamine pyrophosphate
TEMED	Tetramethyl ethylenediamine
WB	Western blot
XL-MS	Crosslinking mass spectrometry
YPG	Yeast Extract-Peptone-Glycerol



## 1. Introduction

### 1.1 Yeast as a model organism for molecular biology and biochemistry

Biological knowledge progresses through the choice and implementation of “model organisms”, *i.e.*, suitable life forms for studying a specific trait, disease, or phenomenon, due to its short generation time, characterized genome and proteome, or similarity to humans in taxonomical terms; examples include bacteria (*Escherichia coli*), yeasts (*Saccharomyces cerevisiae*), fly (*Drosophila melanogaster*), and mouse (*Mus musculus*). Although each organism is unique and its biochemistry may deviate even within single cells, analysis of widely studied organisms has the advantage of systematizing experimental reproducibility and transferability of results across the research community.

*S. cerevisiae* in particular is an attractive model organism, and a single-celled organism, commonly used in the bread and beer-making industry (**Figure 1**). It is one of the simplest eukaryotic organisms and its genome was published in 1996, making it one of the first organisms for which the complete genome was available<sup>1</sup>. *S. cerevisiae* has rapid growth, is easy to maintain and manipulate, and resembles other systems to be compared with, comprising a very useful system to investigate its biology<sup>2</sup>. The similarity of yeast proteins and pathways to other eukaryotic organisms, for example, mammals, is also substantial<sup>3</sup>. As a result, *S. cerevisiae* has been a model system for decades to study biological pathways and phenotypes<sup>4</sup>, such as multi-drug resistance<sup>5</sup> and DNA replication<sup>6</sup>. Yeast can also use fermentation or the respiratory pathway for its carbon sources, which makes it an easily accessible target for studying respiratory diseases by isolating mitochondria<sup>7</sup>. Importantly, yeast has been used to study higher-order cellular interactions and interactomes<sup>8</sup> and comprises a promising model to advance structure-function studies.

As of June 2023, ~33,000 protein structures have been deposited in the protein data bank (PDB)<sup>9,10</sup> from yeast strains out of a total of ~200,000 protein structures (www.rcsb.org). Notably, yeast structures were pioneering for structural analysis by means of X-ray crystallography<sup>11</sup> - the resolution of yeast hexokinase at 2.1 Å allowed sequencing of the protein understanding its three-dimensional coordinates. It is also quite interesting that structural analysis of yeast proteins has also pushed cryo-electron microscopic characterization of biomolecules and their assemblies: Statistics in the PDB (as of June 2023) show that, overall, yeast biomolecules have been resolved by both X-ray crystallography (N=14,754) and cryo-EM (N=18,441). However, during the last 4 years (2020-present), 6 times more structures have been characterized by cryo-EM (N=10,094) rather than X-ray

crystallography ( $N=1,651$ ) stemming from yeast. Therefore, a model system such as yeast has driven other fields of structural biochemistry and is augmenting the “resolution revolution”<sup>12</sup> which we are presently witnessing.

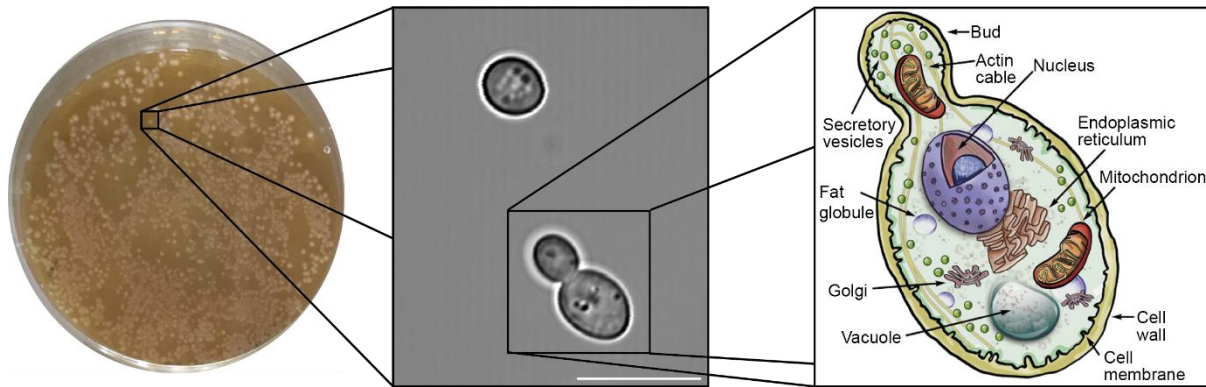


Figure 1. Model organism yeast and its cell structure.

*Saccharomyces cerevisiae* is easily propagated and maintained on agarose plates (left). Yeast cells are round or elliptical in shape with a length of 7-15  $\mu\text{m}$  and width of 2-8  $\mu\text{m}$  (middle panel). Yeast cells are budding to propagate. On the right most panel a schematic of a yeast cell is shown and reproduced with permission from<sup>13</sup>. Scale bar 10  $\mu\text{m}$ .

## 1.2 Cell lysis to study cellular architecture – Strengths and Limitations

Yeast cell extracts are used as food additives, flavorings, and as nutrients for bacterial culture media. However, they are also being used to understand cellular structure and function since decades<sup>14</sup>: Some of the earliest reports to study eukaryotic cellular function involve yeast cell extracts and focused on determining glycolytic oscillations, *i.e.*, the metabolite concentrations fluctuate around phosphofructokinase<sup>15-17</sup>. In addition, yeast cell extracts were historically used to probe the biosynthesis of saturated fatty acids<sup>18</sup>, a work in which the detailed catalytic mechanism using structure-function studies is still being deciphered<sup>19,20</sup>. Study of cell function via homogenization is very critical, as the homogenate becomes a biochemical specimen in which characterization and analysis can occur, retaining cellular function and various systematic characterization have been achieved utilizing cell extracts<sup>21</sup> such as the deciphering of the genetic code<sup>22</sup>, the visualization of chromatin threads<sup>23</sup>. The Nobel prize in Physiology or Medicine in 1974 was awarded to Claude, de Duve and Palade “for their discoveries concerning the structural and functional organization of the cell”<sup>24</sup>.

Technically, any manipulation of the cell is prone to develop artefacts, which have been recently reviewed<sup>25</sup>. Artefacts can stem from cell rupture, centrifugation, and dilution, that can alter the state of biomacromolecules. Therefore, the preparation procedure for cell extracts should be probed in a highly systematic manner. For example, experimental variables

like temperature, *pH*, contents of buffers, incubation times, and the procedure itself can alter the structure, activity, and function of extracted biomolecules and cellular material. Mitochondria is a prominent example of shape deformation; in the cell, they assume a naturally elongated shape, while they become circular after lysis<sup>26</sup>.

Additionally, diluting the cell content may lead to macromolecular dissociation, alteration of 3D structure of recovered biomolecules or loss of biomolecular interactions. Non-specific interactions are also a major concern, as compartmentalization and intrinsic concentrations are changed. Some of the cellular material will be lost in the various steps necessary for isolation of the cell extract, e.g., while filtering or ultracentrifugation steps, as these remove aggregate material and much of the membrane fraction (if soluble material is to be captured for analysis). To ensure optimal conditions and yield, every purification step must be benchmarked and controlled, for example, by immunoblotting analysis and underlying activity assays to ensure that dissociation of subunits is minimal, the native function is maintained, and optimal yield is achieved<sup>25,27,28</sup>. The resulting protocol has to be technically and biologically reproducible and as such the protocol to retrieve active keto acid dehydrogenase complexes, localized in the mitochondria, is described in a recent Ph.D. thesis by Kyrilis<sup>27</sup>.

#### 1.2.1 Isolation of organelles and *in vitro* function

One of the best-known cell organelles is the mitochondrion. It is the so-called ‘powerhouse of the cell’ involved in a variety of energy-gaining processes, such as the tricarboxylic acid cycle (TCA, Krebs cycle) which provides the necessary proton gradient for ATP generation via ADP phosphorylation. Their role in the cell is multifaceted and are virtually involved in nearly all cellular functions<sup>29-31</sup>. Its name comes from the Greek *mitos* (thread) and *chondros* (grain) as it can be found both branching and vesicle-like in the cell. Historically<sup>32</sup>, the microbiologist Richard Altmann observed mitochondria for the first time, and called them *bioblasts*, describing them as elementary organisms living inside cells to carry out vital functions<sup>33</sup>. Carl Benda introduced the name “mitochondrion” in 1898<sup>34</sup>. Structure-function correlations of mitochondria are still extremely active today, building on top of works, e.g., by: (a) Claude, who correlated morphological and biochemical properties of isolated mitochondria<sup>35</sup>; (b) Lehninger who showed that mitochondria include the enzymes of TCA<sup>36</sup> and can perform oxidative phosphorylation<sup>37</sup>; and (c) Lipmann who demonstrated respiratory control<sup>38</sup>.

Palade was the first to communicate higher-resolution structural characterization of mitochondria<sup>39,40</sup>. Mitochondria consist of an outer and inner membrane between the intermembrane space and surrounded by all, is the mitochondrial matrix at the core of the organelle (**Figure 2**)<sup>41</sup>. Although most of its genetic information was transferred into the nucleus, about 1.6 kb is still kept in the mitochondria, allowing it to have its own DNA and gene expression. That way they retain cellular control to a small degree communicating with the nucleus<sup>42</sup>, still the mitochondrion is not self-sufficient and relies on gene products of the nucleus to be imported to manage their enzymatic and signaling pathways<sup>43</sup>. However, there is significant functional interplay between the mitochondrion and the nucleus, as e.g. histone acetylation is performed by a localized TCA cycle within the nucleus<sup>44</sup>. On the other hand, mitochondrial proteins must pass the outer and inner mitochondrial membranes when transported into the mitochondrial matrix or exported to the cytoplasm. Overall, besides the numerous other cellular functions of mitochondria, they are the signature organelles for the supply of adenosine triphosphate (ATP) for the cells, driving, for example, membrane potential, muscle contraction, protein synthesis and cell division<sup>45</sup>. Due to their prominent role, mitochondria are heavily involved in stress response<sup>46</sup>, cancer development<sup>31</sup> via the Warburg effect<sup>47</sup>, and apoptosis<sup>48</sup>.

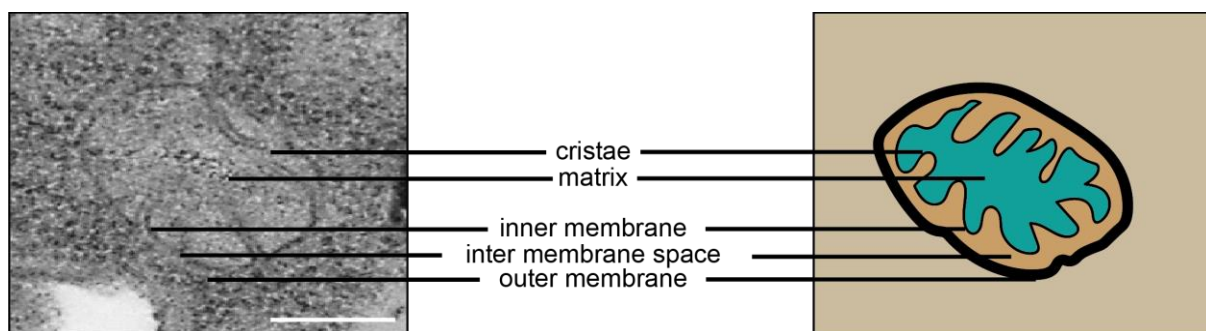


Figure 2. Cryo-substituted mitochondria and general architecture.

Shown are the outer membrane, intermembrane space, inner membrane, cristae, and matrix. Acquired on EM 900 by Dr Gerd Hause. Scale bar 2  $\mu$ m.

### 1.2.2 Cellular protein complexes and most abundant biomolecules

Instead of organelle isolation, separation of the cellular content can include retrieval of organellar and cellular extracts by further lysis which can be useful for probing cellular function; their importance was previously analyzed (§1.2). Technically, for the identification of the extract content, the right isolation and separation method should be chosen. For example, studying well-known organisms with already annotated proteomes makes it easier

to annotate the contents of cell extracts as identification by mass spectrometry is complemented by the known and deeply annotated genome. Additionally, thermophilicity can positively influence recovery of high quality cell extracts as thermophilic proteins are known for their, comparatively higher, stability<sup>49</sup>. Using the right organism and method will make the cell extract applicable to studying protein communities<sup>50</sup>, which are biomolecular complexes of a higher-order consisting of flexibly connected proteins and other molecules working in relative proximity within the same, or across different pathways. These protein communities are usually large and flexible, with prominent and therefore easy to discern in cryo-electron microscopic micrographs<sup>50</sup>. These include, but are not limited to, the synthesis of fatty acids by the fatty acid synthase molecule (FAS)<sup>50</sup>, and the mitochondrial pathways of pyruvate<sup>51,52</sup> and oxoglutarate oxidation<sup>53</sup>.

For the model organism yeast, more precise *Saccharomyces cerevisiae*, the genome as well as proteome are known. Therefore, protein abundance can be identified/verified with mass spectrometry in isolated cellular fractions<sup>54</sup>. Changes in protein-protein or other biomolecular interactions are dependent on an equilibrium between transcription, translation, and post-processing/post-translational modifications of the proteins. Changing conditions for yeast like temperature, starvation, humidity etc. will change the abundance of various proteins and this can be traced by activity assays as well as quantitative MS. The updated Uniprot database contains significant and large-scale information for over 230 million sequences, while the yeast proteome is considered a gold-standard, having high degree of annotation in terms of protein abundance, localization, interactions, and function<sup>55</sup>. Statistics from Uniprot, as of June 2023, show that yeast has 6,727 proteins for which 1974 proteins have (partially) known structure, and their existence has been proven at the protein level for 78.9% of them (N=5,307). Such statistics are among the highest across annotated proteomes in Uniprot.

### 1.3 Probing the structure and function of native cell extract with modern biochemical, biophysical and bioinformatic methods

The preceding sections provide compelling evidence for the utility of native cell extracts in research, despite the challenges associated with cell lysis. Native cell extracts offer a simplified approach to studying cellular components while retaining crucial aspects of cellular organization. They facilitate accessible biochemical and biophysical analyses, encompassing kinetic analysis, mass spectrometry (MS), cross-linking mass spectrometry (XL-MS), cryo-electron microscopy (cryo-EM), and computational studies. Although this Thesis did not explore the totality of the characterization methods of yeast cell extracts, their

potential for further investigation via transcriptomics, metabolomics, and ribosome profiling is promising. By employing diverse techniques to scrutinize native cell extracts, it becomes feasible to identify novel proteins, scaffolds, interfaces, and interactors. Consequently, investigations involving cell extracts serve as an intermediary connecting traditional structural biology techniques like X-ray crystallography and NMR spectroscopy with the latest advancements in in-cell structural biology, such as tomography of cells for comprehensive content analysis. Unlike traditional methods that often examine proteins in isolation, cryo-electron tomography of cells allows for in-situ analysis, preserving the molecular context. However, this method has limitations when visualizing highly abundant biomolecules. In contrast, cell extracts can be selectively enriched using sophisticated biochemical fractionation methods while retaining certain aspects of cellular complexity. Notably, size exclusion chromatography proves invaluable for significantly enriching rare, large macromolecular assemblies that pose challenges for direct visualization within cells. Since the seminal, Nobel-prize-winning works of Palade, de Duve, and Claude that worked on discovering and visualizing organelle structure and function<sup>24</sup>, a lot of progress has been made in modern biochemistry. I will list below those that I am aware of in the current literature and have positively affected the outcomes of this thesis, but the list is by no means exhaustive:

- 1) The first major advance is the improvements in the underlying biochemical instrumentation, in which developments have allowed highly analytical experiments to be performed, therefore, increasing dramatically the underlying reproducibility. An example of this is the automated fractionation of specimen using modern Fast protein liquid chromatography (FPLC) systems<sup>56</sup> as well as the availability of separation columns that allow high quality cell extract fractionations<sup>57</sup>.
- 2) Computing algorithms have evolved drastically, having established their presence for biophysical and physical-chemical characterization of biomolecular complexes with physics-based methods and molecular dynamics simulations<sup>58</sup>. Computer algorithms are being rapidly developed with the aid of artificial intelligence (AI) and deep learning – for example, Alphafold, a neural network architecture aiming to predict structures of proteins, has achieved unprecedented accuracy for these predictions, within experimental error<sup>59</sup>. It is astonishing, that as of June 2023, the original Alphafold publication, communicated in 2021<sup>59</sup> has amassed > 12,000 citations, showing the impact of AI in structural biology. Open availability of its source code, as

well as the creation of the associated EBI database that hosts >200 million atomic models<sup>60</sup> (<https://alphafold.ebi.ac.uk/>) has pushed structural biology forward.

- 3) The “resolution revolution” brought by cryo-EM to modern structural biology<sup>12</sup>. Currently, cryo-EM is fast-tracking to become the primary structural method for characterizing biomolecular assemblies (**Figure 3, a**). For virus structures, which is also the subject of this work, cryo-EM has recently surpassed the number of structures determined and became the major method for structural analysis of viruses (**Figure 3, b**). Cryo-EM usability was demonstrated during the COVID-19 pandemic, where structural characterization of the SARS-CoV-2 virus and its complexes with ACE2, drugs and antibodies was catalyzed by high-resolution cryo-electron microscopy characterization<sup>61</sup>.

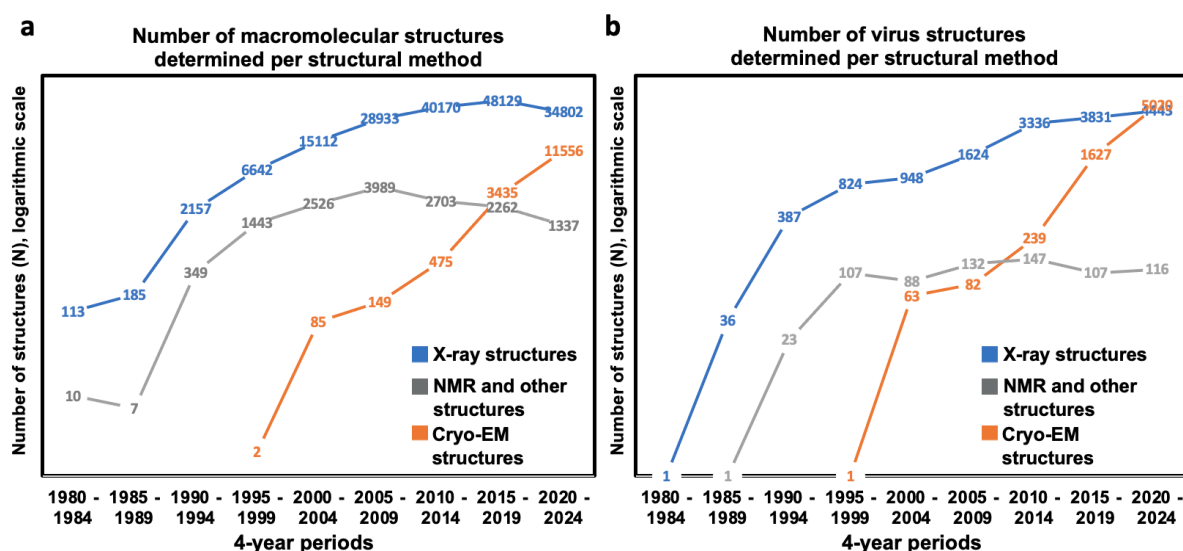


Figure 3. Statistical data for number of determined structures in the Protein Data Bank ([www.pdb.org](http://www.pdb.org)).

(a) Number of structures released in the PDB across time periods; (b) same as (a), but for the subset of molecules related to viruses, which is also what has been studied in this thesis.

- 4) Cross-linking of proteins, complexes, cell extracts and cells. Cross-linking mass spectrometry has also witnessed major advancements during the last years<sup>62</sup>. Pivotal cross-linking agents, state-of-the-art software tools, and streamlined XL-MS methodologies have recently become available to explore biomolecular interactions. XL-MS is used to illuminate the characteristics of proteins, their assemblies, and intricate interaction networks both in controlled laboratory environments and living organisms. A pivotal aspect in extracting three-dimensional structural insights from

XL-MS data lies in the indispensable role of computational modelling. Notably, the integration of XL-MS with other cutting-edge techniques in structural biology, such as cryo-EM, has yielded resounding success in addressing long-standing biological queries that were previously elusive.

- 5) Omics methods. The advent of omics methods<sup>63</sup> is transforming the investigation of native cell extract content and structure by offering a wide array of powerful and comprehensive analytical tools. Genomics, for example, have revolutionized the field by allowing researchers to sequence and analyze an organism's complete genome, providing valuable insights into genetic variations and their implications for disease<sup>64</sup>. On the other hand, transcriptomics have proven instrumental in identifying and quantifying RNA molecules within cells, shedding light on gene expression patterns and regulatory mechanisms<sup>65</sup>. The field of proteomics has enabled the identification and quantification of proteins, even, recently, for single cells<sup>66</sup>, unraveling crucial information about protein-protein interactions, post-translational modifications, and protein functions<sup>67</sup>. Additionally, metabolomics have emerged as a key technique for profiling small molecules and metabolic pathways, greatly enhancing our understanding of cellular metabolism and facilitating the identification of biomarkers<sup>68</sup>. By integrating these diverse omics approaches, it is expected that an unprecedented, detailed view on the structure and function of native cell extracts can be achieved.

Based on the above-mentioned advances, the Kastiris laboratory has investigated the structure and function of native cell extracts from *Chaetomium thermophilum*, a thermophilic filamentous fungus able to grow at high temperatures (up to 65 °C)<sup>25,50-53,69</sup>. In these works, various advances have been realized, including the merging of artificial-intelligence-based protein sequencing of cryo-EM Coulomb potential maps from native cell extracts, structure-function insights into gigantic biomolecular assemblies implicated in mitochondrial metabolism, and establishment and optimization of biochemical and structural biology methods to probe highly heterogeneous cell extract content. Despite these major advances, the methods described in these works have not been employed yet in a mesophilic eukaryote. Therefore, it is unknown if the recovered cell extracts have a conserved content across different species. It is also important to note that the methods developed only apply to high molecular weight molecules due to the limitations imposed by the column used for separating the cell extract at the FPLC. In addition, cryo-EM might have a limit on the size of the



biomolecules it is able to image, however this limit is still explored as methods to overcome this are currently being communicated<sup>70-72</sup>.

### 1.3.1 Cryo-Electron Microscopy and other structural methods to probe the architecture of proteins and their complexes at high resolution

The current surge in high-resolution structures (*i.e.*, better than 4 Å resolution) obtained through cryo-electron microscopy is commonly called the "Resolution Revolution", a term coined by Werner Kühlbrandt<sup>12</sup>. This revolution has been acknowledged through the Nobel Prize in Chemistry awarded to Dubochet, Frank, and Henderson for their contributions to cryo-electron microscopy<sup>73</sup>, enabling the high-resolution analysis of biomolecular complexes. Their contributions can be categorized into the following areas: Improvements in hardware have played a pivotal role in enhancing traditional electron microscopes. This includes the automation of data acquisition, optimization techniques to reduce specimen contamination, and the utilization of direct electron detectors<sup>74</sup>, *i.e.*, cameras that directly detect electrons. Additionally, advancements in sample preparation procedures, such as commercializing plunging machines for flash-freezing biological specimens to liquid nitrogen temperatures, have furthered progress. Computational image processing algorithms, including Bayesian methods for image classification<sup>75</sup>, precise micrograph alignment after beam-induced movement of the specimen<sup>75</sup>, and accurate calculation of image parameters<sup>76,77</sup>, have also significantly contributed to the field. It is important to know that current software for image analysis integrate most of these algorithms in user-friendly software that provide direct protocols for cryo-EM image analysis for cryo-EM map calculation, such as RELION<sup>78</sup> and cryoSPARC<sup>79</sup>.

Several noteworthy examples of structures resolved by cryo-EM include the spliceosome and its alternative splicing states<sup>80,81</sup>, a wide range of membrane proteins and complexes<sup>82</sup>, as well as the structural characterization of photosystem complexes<sup>83</sup>, which presented significant difficulties prior to the "resolution revolution." It is worth mentioning that the ribosome, which was considered exceptionally difficult to be studied by X-ray crystallography, is now a standard specimen for cryo-EM, achieving resolutions better than 1.6 Å, allowing assignment of post-translational modifications and other types of polymorphism<sup>84</sup>.

Protein X-ray crystallography stands as a distinctive scientific method utilized to investigate the precise atomic-scale three-dimensional arrangements of proteins<sup>85</sup>. The process begins with obtaining a concentrated and pure sample of the target protein. Meticulous efforts are

then made to promote the formation of crystals, where individual protein molecules align in an ordered and repetitive manner. X-ray diffraction involves directing a beam of X-rays at the crystal, which scatters as it interacts with the atoms within the protein molecules and the crystal solvent. This scattering creates a diffraction pattern captured by a specific detector. Intricate mathematical calculations are then employed to determine the distribution of electron density within the crystal, enabling the construction of a comprehensive three-dimensional model of the protein. This model reveals the precise arrangement of atoms, secondary structures, and overall architectural blueprint. X-ray crystallography studies proteins in a crystallized state, which may differ from their physiological state. Moreover, not all biomacromolecules can be crystallized due to their complex and dynamic nature, limiting access to structural data using X-ray diffraction. Nevertheless, protein X-ray crystallography provides valuable insights into the structure and function of biomolecules, enhancing our understanding of their biological roles and interactions with other molecules. This knowledge is crucial for designing and developing effective therapeutic interventions targeting specific protein structures implicated in diseases, especially since X-ray crystallography often achieves higher resolutions than cryo-EM<sup>86</sup>.

Protein Nuclear Magnetic Resonance (NMR) spectroscopy investigates protein structure and dynamics at the atomic level<sup>87</sup>. Specific atomic nuclei possess magnetic properties in a magnetic field, *i.e.*, they resonate, and protein NMR spectroscopy harnesses this principle to gain insights into protein structures<sup>88</sup>. The process involves:

- Preparing a pure and concentrated protein sample.
- Labelling the sample with heavy atoms.
- Placing the sample in an NMR tube within a strong magnetic field.

Radiofrequency pulses manipulate the atomic nuclei, inducing changes in their spin states. The NMR spectrometer, as the nuclei return to their equilibrium state, captures emitted signals that inform on the chemical environment, including positions, orientations, and interactions with neighboring atoms. Analysis and interpretation of the NMR data in the form of distance restraints provide details about the protein's three-dimensional structure, interactions, and dynamics.

Protein NMR spectroscopy offers advantages such as studying proteins in solution, which mimics physiological conditions, and sensitivity to conformational changes and dynamics; however, looking at larger proteins or multimers presents challenges due to signal overlap and low sensitivity<sup>89,90</sup>. Special labelling procedures or additional NMR methods can overcome these limitations but introduce complexity. In summary, protein NMR

spectroscopy is a powerful technique for studying protein structure, dynamics, and interactions. Although validation ensures accuracy, structure calculation remains challenging, requiring heavy labelling and complex data analysis.

### 1.3.2 Focus on the potential of Cryo-EM

Cryo-EM was initially used to study the morphology of large protein complexes that were difficult to crystallize, albeit at much lower resolutions than crystallography<sup>91,92</sup>. However, recent technological breakthroughs in sample preparation, computation, and instrumentation, discussed above, have allowed researchers to achieve near-atomic-resolution structures using cryo-EM. Cryo-EM includes various steps that would eventually lead to the derived Coulomb potential map<sup>93</sup>. The first step in a cryo-EM experiment is sample preparation. A purified protein sample is applied to a particular grid with tiny holes in a film of amorphous carbon supported by a metal frame. The sample can also be heterogeneous, as in the yeast cell extracts investigated in this Thesis. This biochemical preparation aims to distribute the protein particles evenly within the grid holes in various orientations. The grid is then rapidly frozen by plunging it into a cryogen-like liquid, ethane, creating a thin film of vitreous ice that traps the particles. This freezing process not only captures the protein structure at the moment of freezing but also protects the sample from radiation damage and prevents buffer evaporation in the high-vacuum conditions of the transmission electron microscope. Sample preparation for cryo-EM is still, mostly, a manual process and comprises the major bottleneck for structure elucidation<sup>93,94</sup>.

Improvements in sample preparation have been achieved by optimizing protocols for purifying fragile protein complexes, automating sample grid preparation, and enhancing the grids' quality<sup>94</sup>. These incremental optimizations collectively, significantly impact the success of cryo-EM experiments. In addition, various groups explore the nature of biochemical samples, which may include chemically heterogeneous specimens, membranes, organelles, or even cells utilizing cryo-electron tomography<sup>95,96</sup>.

In single-particle cryo-EM, 2D electron micrographs of individual protein particles on the sample grid are captured and are subjected to various image analysis steps<sup>97,98</sup>. Due to the need to use low electron dose to avoid damaging radiation-sensitive samples (such as biomolecules), these 2D projections are inherently noisy and do not provide atomic-level details. However, the signals can be improved by averaging many individual particle images. Since particles freeze in random orientations on the grid, sophisticated image-processing methods are employed to align and merge the data, constructing an initial 3D map. This map is then iteratively refined and validated using specialized software tools. Finally, the protein

sequence is fitted into the 3D map to generate a complete 3D model of the protein using crystallographic tools, such as COOT<sup>99</sup> and Phenix<sup>100</sup>. In the past, millions of individual particle images were required to solve high-resolution structures<sup>101</sup>. However, the development of susceptible direct-detection cameras has made it possible to achieve high-resolution structures using less particles per image<sup>102</sup>. This advancement saves time, preserves precious samples, and improves resolution. Direct-detection cameras' recent development and commercialization have been significant milestones for cryo-EM<sup>103</sup>. Unlike CCD cameras, which convert electrons into photons for image recording, direct detectors capture the incident electrons. This allows for much higher sensitivity in collecting particle images. Direct detectors are also fast, enabling images to be recorded in "movie" mode. This capability has facilitated the correction of image blurring caused by slight movements of the samples during imaging inside the electron microscopes, leading to near-atomic-resolution information. The development of direct electron detectors has revolutionized cryo-EM by capturing a series of movie frames, allowing for the correction of beam-induced motion and mitigation of radiation damage. These detectors have seen remarkable improvements in quantum efficiency and speed, and further advancements are expected in the future.

Below, I analyze factors that have already been recognized to limit resolution of cryo-EM specimen<sup>104</sup>. One challenge in cryo-EM is conformational heterogeneity<sup>105</sup>: Unlike in crystallization, where a protein is locked into its most stable orientation, cryo-EM samples can freely move until flash-frozen. While cryo-EM allows for the capture and study of conformational transitions, this heterogeneity can complicate high-resolution 3D reconstruction. Computational algorithms have progressed in classifying heterogeneous data sets into structurally homogeneous subsets, but further improvements are still needed.

The transition from photographic film to direct electron detectors marked a significant leap in cryo-EM performance. However, it is essential to acknowledge that the impressive improvement in resolution is not solely due to detector technology. Other factors, such as advancements in electron microscopic hardware, computation, and image processing software, have also played crucial roles. Automating various technological aspects has been fundamental to the success of single-particle cryo-EM and ensures its continued significance in structural biology. Despite these remarkable achievements, there are still limitations in reaching an atomic resolution for many macromolecular complexes. Statistics from the EMDB indicate that most entries fall within the 4-5 Å resolution range<sup>104</sup>, which does not allow for accurate atomic model building. This raises questions about the underlying limitations of cryo-EM that need to be addressed. Note that my re-analysis, of the statistical

data from PDB cryo-EM maps shows that cryo-EM is improving in terms of map resolution: most cryo-EM maps deposited in 2023 have better than 4 Å resolution (**Figure 4**).

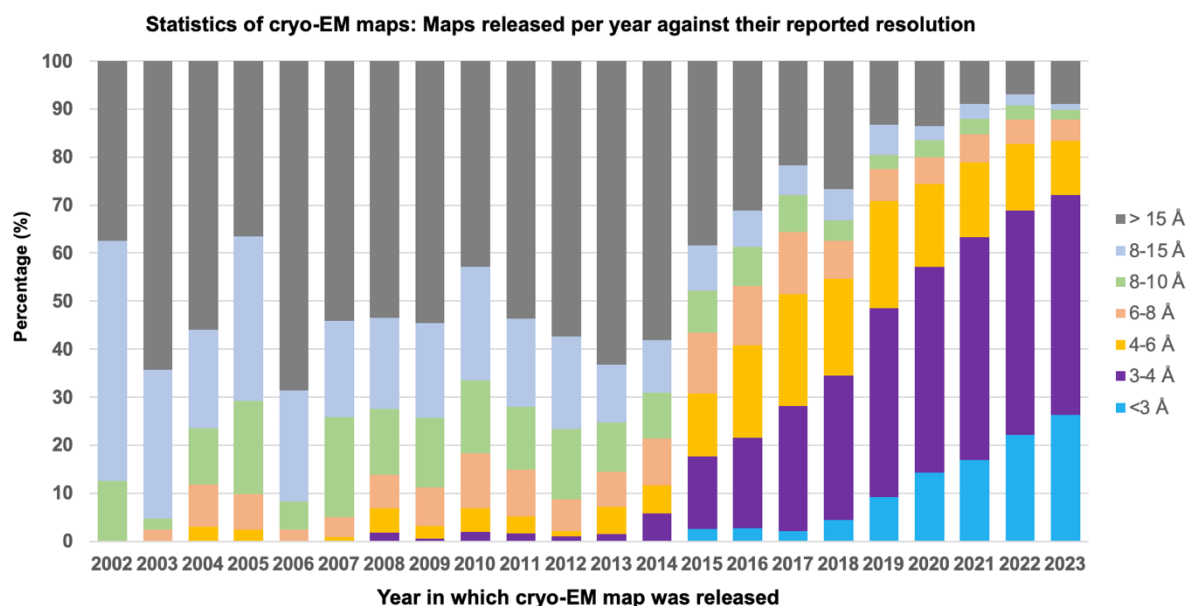


Figure 4. Statistical data for percentage of determined structures in the Electron Microscopy Data Bank (<https://www.ebi.ac.uk/emdb/>).

The resolution revolution clearly had an effect on determined cryo-EM map quality; in 2023, most cryo-EM maps are of resolution higher than 4 Å.

When it comes to cryo-EM, several factors can limit the achievable resolution. Hardware-related issues, including mechanical instabilities, vibrations, and thermal drift, play a significant role that can impact microscope performance. To mitigate these effects, microscopes are often installed in the basement of buildings. Additionally, lens aberrations arising from the quality of magnetic lenses, the electron source, the acceleration voltage, and optical system alignment can limit resolution<sup>106</sup>. Correcting these aberrations, such as defocus and twofold astigmatism, is crucial, and higher-order aberrations like axial coma become important at higher resolutions. The type of electron source used also influences resolution. Field emission electron sources offer improved spatial and temporal coherence compared to thermionic sources, contributing to the resolution revolution in cryo-EM<sup>107</sup>. Cold field emission sources with reduced energy spread perform even better in the 1-2 Å resolution range<sup>108,109</sup>. Monochromators can enhance temporal coherence, enabling sub-1 Å resolution imaging<sup>104</sup> while chromatic aberration becomes resolution-limiting at the 1 Å level and addressing it is critical<sup>104</sup>.

In conclusion, some of the limitations in cryo-EM resolution emerge from hardware-related issues, including mechanical instabilities and lens aberrations, as well as the properties of the

electron source and the performance of electron detectors. In addition, image processing, although incredibly advanced, still needs to be streamlined as different software and different operators will end up with overlapping but distinct single-particle data sets and their analysis involves various human-decision steps<sup>110</sup>. This shows that the calculated final structure represents only a minor fraction of the acquired single particles and might not be the representative structure in the vitreous ice<sup>111</sup>.

Nevertheless, recent advancements in electron microscopy (EM) have enabled the determination of atomic-resolution 3D reconstructions of proteins<sup>108,109</sup>. To achieve this level of resolution, two 300 kV electron microscopes equipped with advanced optics surpassing most high-end microscopes were utilized. These microscopes employed cold field emission electron sources, energy filters, and electron detectors that improved signal transfer for capturing high-resolution information. One of the microscopes had a spherical aberration corrector<sup>108</sup>, which minimized off-axis coma and allowed for aberration-free imaging, a capability not achievable with ordinary, non-corrected microscopes.

Determining atomic-resolution structures is crucial for various applications such as structure-based drug design and a precise understanding of chemical inhibition mechanisms – something more frequently achieved in X-ray crystallography. Even the accurate placement of a single atom can significantly impact the interpretation of these processes. However, achieving such high-resolution structures requires careful consideration of microscope capabilities and image processing methods. Fourier shell correlation (FSC) is commonly used to estimate resolution in cryo-EM<sup>112</sup>, but it does not provide quality indicators. It can be affected by systematic errors such as coma or errors in coma correction. Achieving atomic-resolution structures should follow Sheldrick's criterion<sup>113</sup>, and is only worthwhile when every atom position can be visualized and interpreted confidently, ensuring reliable chemical interpretation. The Rosenthal & Henderson plot<sup>114</sup> is a valuable tool for understanding resolution limits in cryo-EM and provides insights into microscope-detection limitations. Available microscope-detector combinations are currently suitable for structures with 2-3 Å resolutions but breaking the 1 Å resolution barrier remains a significant challenge. While most cryo-EM depositions have resolutions above 4 Å, the capabilities of electron microscopes suggest that resolutions better than 3 Å should be achievable. Factors beyond hardware capabilities, such as sample characteristics, likely influence the limitations in achieving higher-resolution structures<sup>94</sup>. Macromolecular complexes exhibit various properties, making their purification challenging. The quality of cryo-EM images may not always meet expectations due to sample instability or aggregation during purification. Factors

like pH and buffer systems used in the purification process can influence sample aggregation. Furthermore, working with dilute samples can lead to labile interactions among subunits, and purification tags can impact the composition and stoichiometry of subunits.

The air-water interface has been considered a potential denaturing environment for proteins in cryo-EM grid preparation. While direct observations of denaturation at the air-water interface are limited, it is challenging to distinguish denaturation from disintegration. Various approaches, such as support film binding and detergent addition, have successfully mitigated adverse effects at the air-water interface. However, there is still a need to understand better and address the challenges associated with the air-water interface in preparing cryo-EM grids for monodisperse and stable protein complexes<sup>115</sup>.

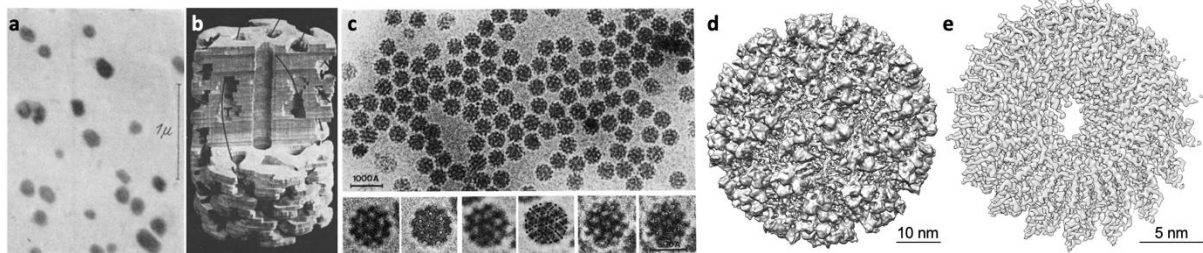
Overall, current cryo-EM equipment can determine high-resolution structures from pure samples, but technical challenges persist in achieving resolutions below 1 Å. Advances in hardware and software can support resolutions of sub-2 Å, but overcoming the challenges and limitations in cryo-EM sample preparation is essential. Improvements in grid preparation, vitrification procedures, and purification strategies are crucial for unlocking the full potential of cryo-EM in solving macromolecular structures at unprecedented resolutions.

The analysis mentioned earlier regarding the prospects and difficulties in single-particle cryo-EM can also be extended to structures obtained from natural cell extracts. Nevertheless, the presence of endogenous proteins and complexes introduces extra obstacles, including heightened chemical and structural diversity. In practical terms, the resolutions achieved in heterogeneous native cell extracts typically range from 3 to 5 Å<sup>50,52,53,69,116-118</sup>.

#### 1.4 Cryo-EM and virus structure

History of electron microscopy is tightly connected to virus specimen<sup>119</sup>: In the 1930s, Berlin witnessed the emergence of the electron microscope, driven by competition among multiple groups. AEG's Ramsauer, Brüche, Scherzer, and Mahl; independent entrepreneur Manfred von Ardenne; and the Technical High School's Knoll, Ernst Ruska, and Bodo von Borries were the key players. However, progress necessitated the involvement of industry, which came about when senior clinician Richard Siebeck convinced Siemens of the microscope's medical significance. Siemens subsequently took charge of its development and production. By 1938, Helmut Ruska had assumed leadership in sample preparation and application in the life sciences: Ruska's numerous papers were published from 1938 to 1940<sup>120-122</sup> and Siemens established a laboratory for electron microscopy under Ruska's direction. Throughout his life, Ruska left a significant impact on electron microscopy, particularly in virology and the

classification of viruses in vertebrates, plants, and bacteria. One of the first 3D reconstructions from electron microscopy was performed for the phage tail<sup>123</sup>, while cryo-electron microscopy also made a major leap through the study of viruses: The proof for structure preservation after vitrification of a biological specimen was provided by the Dubochet group in EMBL-Heidelberg utilizing different viruses as proof-of-principle<sup>124</sup>. The Semliki forest virus, which was a sample that demonstrated preservation after vitrification, was also one of the first to be determined by cryo-EM at a resolution of 9.0 Å<sup>125</sup>. Finally, one of the first high-resolution reconstructions of a virus, that of tobacco mosaic virus, at 4.4 Å was communicated before the “resolution revolution”, in 2005<sup>126</sup>. These developments are also shown in **Figure 5, a-e**. As of June 2023, more than 7,000 molecular structures of viral proteins have been deposited in the protein data bank<sup>10</sup>, and ~70% of them have reported resolutions higher than 4 Å<sup>9</sup>. Finally, during the COVID-19 pandemic, cryo-EM greatly accelerated our understanding of SARS-CoV-2 infection processes through the application of single-particle cryo-EM of viral proteins and their complexes, as well as tomography of cells and intact SARS-CoV-2 viruses.



**Figure 5. Viruses as specimen for the progress of (cryo-)electron microscopy.**

(a) Virus specimen used for showing the higher resolution of electron microscopy by Helmut Ruska; image retrieved and adapted from<sup>120</sup>. (b) First 3D model by De Rosier and Kug, that of the bacteriophage tail; image retrieved and adapted from<sup>123</sup>. (c) The Semliki forest virus, vitrified by the Dubochet group in EMBL; image retrieved and adapted from<sup>124</sup>; (d) The same virus as in (c), but reconstructed in 3D<sup>125</sup>; map was retrieved from EMDB (EMDB-ID: 1015). (e) Cryo-EM map of TMV, showing resolution close to 4 Å<sup>126</sup>. Map was retrieved from EMDB (EMDB-ID: 1316).

#### 1.4.1 General aspects on viruses with a focus on RNA viruses

Although not considered living organisms, viruses belong to life and have accompanied living organisms for a long time, even driving protein adaptation<sup>127</sup>. The overall building blocks are shared between all viruses despite being found in various sizes and shapes<sup>128,129</sup>. They consist of genetic material, RNA or DNA, double-stranded or single-stranded, enveloped by a proteolipid or protein shell. Usually, an infection with a virus will have the following six distinct steps, first attachment to the host cell, second inserting and unpacking



the viral genome, then transcription and translation while hijacking the host cells replication machinery, followed by the assembly, and last, death of the host cell and release of new virus particles into the environment where the cycle will start anew<sup>130</sup>. During this process, viral factories can be observed. These are clusters of viruses associated with the cytoskeleton, membranes, and other cell features. Most viral infections will lead to the host cells' death or damage<sup>131,132</sup>. Still, some viruses, like retroviruses or endogenous viruses, can have a neutral or positive effect on the host cell. The effects involve shaping immune responses<sup>133</sup>, adaptations, or resulting in a phenotype that helps the host-virus pair survive while other organisms or cells around them die<sup>133,134</sup>.

Overall, viruses are very diverse and can be grouped into seven classes based on how they replicate and package their genetic material (Baltimore classification)<sup>135</sup>. There are two classes of DNA viruses, three classes of RNA viruses, and two classes of reverse-transcribing viruses. Despite their different life cycles, some RNA viruses, dsRNA viruses, and reverse-transcribing viruses have similarities in how they replicate their genetic material<sup>136</sup>. They all use RNA intermediates that also act as messenger RNA (mRNA). Additionally, the structures and functions of the intracellular complexes involved in RNA replication are similar among these viruses. Some RNA viruses also share similarities with retroviruses in how they initiate genome replication using tRNA-related sequences<sup>136</sup>. They also have similarities with dsRNA reoviruses in the function and interaction of their RNA polymerase and RNA-capping enzymes. On the other hand, there are distinct features shared by other groups of viruses. For example, (+)RNA picornaviruses, dsRNA birnaviruses, and reverse-transcribing hepadnaviruses have protein-primed nucleic-acid synthesis in common. These similarities suggest that some RNA viruses, dsRNA viruses, and reverse-transcribing viruses may have evolved from a common ancestor.

#### 1.4.2 Viruses and protein disorder

Disorder in proteins is ubiquitous<sup>137</sup> and can define certain protein regions (Intrinsically disordered regions, IDRs) or entire proteins (IDPs)<sup>138</sup>. Therefore, it is of no surprise that disorder manifests in viral proteins as well, a subject extensively covered by Dunker and Uversky<sup>139-142</sup>. IDRs/IDPs have certain properties<sup>143,144</sup>, such as enrichment of disorder-promoting amino acid residue compositions, enhancing binding promiscuity, specific patterns of charge, *i.e.*, high negative or positive charge content, and ability to interact via, *e.g.*, molecular recognition features (MoRFs). Disorder in viruses, specifically, confers flexibility to adapt, survive, and evade host defenses<sup>139</sup>: Viral genomes exhibit a high mutation rate,

impacting multiple proteins due to compactness and overlapping reading frames. Throughout the virus's life cycle, interactions with host cell components occur, enabling attachment, hijacking of cellular machinery, synthesis and assembly of viral components, and release of infectious particles. Intrinsic disorder of viral proteins plays a crucial role at each stage. This is evident by the observation that viral proteins involved in different virus life cycle steps are shown to be disordered<sup>145</sup>. A very important aspect of manifested disorder in viruses is via alternative splicing or frameshifting, where long disordered regions can appear in the alternative protein products<sup>140</sup>— such a mechanism is explored in this dissertation, where the RNA polymerase of the dsRNA L-A virus, of the *Totiviridae* family, is produced via frameshifting and is connected via a disordered extension to the Gag capsid protein<sup>146</sup>.

#### 1.4.3 Endogenous viruses

Viruses simultaneously and independently developed through all parts of life including prokaryotes, eukaryotes, plants, and mammals. Then through this development endogenous viruses appeared throughout all types of eukaryotic viruses. An endogenous virus integrates into the germline genome and then is vertically inherited<sup>134</sup>. It is important to stress that viruses, traditionally considered as pathogens, can actually provide benefits to their hosts and play essential roles in their life cycles<sup>134</sup>: For instance, the polydnviruses found in endoparasitoid wasps have coevolved with their hosts, becoming indispensable. Over time, many viral genes have become integrated into the host nucleus. Additionally, endogenous retroviruses are commonly found in the genomes of higher eukaryotes and have contributed to the evolution of their hosts, notably in placental mammals. Some mammalian viruses even offer protection to their hosts by preventing infections from related viruses or diseases caused by unrelated pathogens, such as the case of bubonic plague<sup>147</sup>. In a more complex symbiotic relationship, a fungal virus confers thermal tolerance to a plant, involving both the fungal host and the colonized plant. Moreover, several acute plant viruses exhibit conditional mutualism by enhancing drought tolerance in plants. Lastly, insect viruses establish numerous mutualistic relationships with their hosts, and viruses also contribute to bacterium-insect mutualisms. As to why these viruses became endogenous, it is hypothesized that these viruses accidentally integrated in the host<sup>133</sup>. Interestingly, viruses can also safeguard their hosts by eliminating competitors, as observed with killer viruses in yeasts<sup>148</sup>.

#### 1.4.4 dsRNA viruses

The dsRNA viruses are a diverse group of pathogens that infect a wide range of organisms, including vertebrates, invertebrates, plants, fungi, and bacteria. These viruses have icosahedral capsid structures and share common replication strategies and biochemical properties<sup>149</sup>. They also have similarities in their inner capsid layers and internal enzymes. However, the outer capsid layers, which are involved in virus transmission and infection, show greater diversity in both sequence and structure. There are currently eight recognized families of dsRNA viruses, each containing viruses of medical, veterinary, or agricultural importance<sup>149</sup>. For example, rotaviruses cause a significant number of infant deaths globally<sup>150</sup>, while bluetongue virus affects ruminants, particularly sheep, causing widespread mortality<sup>151</sup>. Several plant viruses also cause diseases in major crop species like wheat and rice<sup>152</sup>. The largest family of dsRNA viruses is the Reoviridae, which includes 74 recognized species<sup>153</sup>. Despite similarities in replication strategy and some shared proteins, there is a low level of sequence similarity between different genera, making it difficult to determine protein function based on sequence alone<sup>149</sup>. To successfully infect and replicate in host cells, dsRNA viruses have evolved specific strategies. They carry their own transcription and capping enzymes to synthesize functional viral mRNAs since dsRNA molecules are not effective as templates for host cell translation<sup>154</sup>. However, host cells have antiviral defense mechanisms that can recognize and respond to naked dsRNA. To avoid detection, many dsRNA viruses keep their genomes and mRNA synthesizing enzymes within closed protein capsid shells<sup>149</sup>.

The inner capsid layers and internal enzymes of the viruses show remarkable similarities, not only within the Reoviridae but also between distant families of dsRNA viruses. The inner capsid shells are composed of large proteins and provide an assembly mechanism for the icosahedral capsid. However, some families of dsRNA viruses have different structures, indicating distinct evolutionary pathways. The outer capsid layers of icosahedral viruses are responsible for initiating infection by delivering the viral core into host cells. These outer layers face host immune defenses and are subject to selective pressure. They exhibit adaptations specific to host species, leading to a higher degree of diversity compared to the inner core proteins<sup>149</sup>.

The genomes of icosahedral dsRNA viruses are limited in size to fit within the inner capsid shell while allowing transcription. Many dsRNA viruses have segmented genomes, which provide genetic diversity and protect against lethal mutations. Each genome segment is transcribed separately, producing full-length RNA copies. The mRNAs function as templates

for protein synthesis and genome replication. Different mRNAs are produced in varying amounts, providing control over gene expression levels. Some dsRNA viruses have evolved mechanisms to generate multiple proteins from a single mRNA<sup>155</sup>. These mechanisms include sequential and unrelated ORFs, post-translational modification of polyproteins, translation "skip" sequences, multiple initiation sites, and alternative reading frames<sup>156</sup>.

#### 1.5 The helper/killer system in yeast

In 1963, Bevan and Makower made a groundbreaking discovery in yeast, known as the killer phenomenon<sup>157 and refs therein</sup>. They identified a strain of *Saccharomyces cerevisiae* in a brewery that exhibited unique characteristics. These killer yeasts secreted a toxin called killer toxin, which had the ability to kill sensitive cells of the same or related yeast genera without direct contact<sup>158</sup>. Interestingly, killer strains were immune to their own toxin but susceptible to toxins produced by other killer yeasts<sup>159</sup>. This killer phenotype was found to be widespread among yeasts and offered a competitive advantage in their natural habitats. Killer yeasts were subsequently identified in various genera, including *Candida*, *Cryptococcus*, and *Pichia* and later discovered that yeast dsRNA viruses were responsible for the killer phenomenon in *S. cerevisiae*<sup>160</sup>. These viruses, which were considered non-infectious, belong to the *Totiviridae* family, and one of the well-studied viruses is the L-A virus<sup>161</sup>.

The killer strains were categorized into three groups (K1, K2, and K28) based on their toxins and immunity components<sup>162</sup>. They relied on a helper virus for stable maintenance and replication<sup>160</sup>. L-A, an autonomously replicating mycovirus, has a distinct structure consisting of icosahedral shells with small openings for metabolite influx and protection of viral RNA<sup>161</sup>. Inside the L-A particles, there is a single linear 4.6-kb L-A dsRNA that encoded two overlapping open reading frames (ORFs) responsible for the major capsid protein Gag and the RNA-dependent RNA polymerase<sup>163</sup>. The Gag-Pol fusion protein is formed through a ribosomal frameshift event during translation<sup>164</sup>: L-A particles replicated *in vivo* by synthesizing (+) and (-) RNA strands within the viral particle, with the (+)ssRNA molecule serving as an mRNA template and being packaged into new viral particles. The M viruses, which were satellites of L-A, replicated by parasitizing the Gag and Gag-Pol proteins<sup>165</sup>.

The precise mechanism by which killer toxins functions was extensively studied<sup>158</sup>: K1 and K28 were two well-characterized virally encoded killer toxins in *S. cerevisiae*<sup>166</sup>. Although they differ in their amino acid composition and mode of action, they share similarities in their synthesis, processing, and secretion. Both toxins start as preprotoxins and undergo post-translational modifications before being secreted as mature  $\alpha/\beta$  heterodimeric protein toxins<sup>166</sup>. The processing of the toxins involves specific enzymes, such as Kex2p/Kex1p,

resulting in the secretion of active toxin proteins<sup>167</sup>. The mode of action of K1 and K28 toxins differed, with K1 forming ion channels from outside the cell<sup>168</sup> and K28 entering the cell through endocytosis<sup>169</sup>. Once inside the cytoplasm, K1 toxin disrupted membrane function, while K28 toxin transduced its toxic signal into the yeast cell nucleus, leading to cell cycle arrest.

The exact mechanism of immunity against killer toxins *in vivo* remained elusive. Several models were proposed, suggesting that immunity could be conferred by the unprocessed toxin precursor itself, acting as a competitive inhibitor of the mature toxin<sup>158</sup>. Other models suggested that immunity could be achieved through modifications of the toxin's receptors<sup>170</sup> or by diverting the toxin-receptor complex to the vacuole for degradation<sup>171</sup>. The maintenance of yeast dsRNA viruses is facilitated by the MAK genes<sup>159,164</sup>: MAK3, MAK10, and PET18 genes are required. Each of these genes contributes to different aspects of viral stability and replication. More than 30 MAK genes are involved in maintaining toxin-coding M satellites. Additionally, MAK genes encoding ribosomal proteins affect ribosomal subunit assembly. A diminished Gag to Gag-Pol ratio leads to a selective loss of the viral M dsRNA genome.

Mutations in six chromosomal SKI genes suppress MAK mutations in dsRNA-harboring killer strains. These SKI genes, known as superkiller genes, forming a complex<sup>172</sup>, are responsible for repressing translation of poly(A)- RNAs and play a crucial role in regulating mRNA degradation. The SKI gene system consists of five characterized genes (SKI2, SKI3, SKI6, SKI7, and SKI8), and SKI1 (identical to XRN1) encodes a 5'-exoribonuclease involved in general RNA degradation. Ski2p, Ski3p, and Ski8p are proposed to form a complex that functions as an RNA helicase and/or cofactor for the exosome complex through interaction with Ski7p<sup>173</sup>. Ski6p is part of the eukaryotic exosome complex responsible for the 3'-5' degradation of mRNA<sup>174</sup>. However, the exact *in vivo* function of SKI genes in yeast and higher eukaryotes is still debatable - they may serve as an antiviral defense system, down-regulating dsRNA viral copy number and repressing expression of non-poly(A) mRNAs<sup>175</sup>.

Killer yeasts have diverse industrial applications<sup>176</sup>, including combating contaminants, preserving foods, developing antimycotics, and advancing DNA technology. They excel as starter cultures in beer and wine fermentations, improving outcomes by outcompeting unwanted yeasts. Killer toxins and antifungal proteins derived from various sources show promise in treating fungal infections, especially in immunocompromised patients. Targeting yeast cell wall components like mannoproteins, glucans, and chitin, these toxins inhibit  $\beta$ -1,3-

D-glucan biosynthesis, making them potential antifungal agents. Additionally, killer toxins have been employed as models for introducing fungal resistance in transgenic plants. Their expression in crops could provide a novel approach to controlling fungal pathogens. Yeast species like *S. cerevisiae* and *S. pombe* play crucial roles in cell biology, cell cycle control, and heterologous protein production. Their post-translational modification and secretion pathways make them attractive hosts for expressing foreign proteins, facilitated by yeast or killer toxin signals and are an interesting topic for structural analysis<sup>177</sup>.

#### 1.5.1 Structural studies on yeast L-A and L-BC viruses

*Totiviridae* is a virus family with monopartite double-stranded RNA (dsRNA) genomes<sup>178</sup>. These genomes contain two open reading frames: one encoding the capsid protein and the other encoding the RNA-dependent RNA polymerase (RdRp). *Totiviridae* infects fungi (genera Totivirus and Victorivirus) and unicellular protists (genera Trichomonasvirus, Gardiavirus, and Leishmanivirus). Unclassified totiviruses have also been found in arthropods<sup>179</sup> and vertebrates<sup>180</sup>. Most totiviruses are mycoviruses that reside as complete virions in infected cell cytoplasm and spread through cell division or cytoplasmic contact. However, arthropod totiviruses and *Giardia lamblia* virus can be transmitted extracellularly. Two totiviruses infect yeasts: *S. cerevisiae* viruses L-A (L-A) and L-BCLa (L-BC). Wild yeast strains may also carry M particles, which contain a toxin-antitoxin gene and help retain the virus<sup>181</sup>. M-viruses are not equipped with capsid proteins or RdRp but are maintained by the L-A helper virus. Different types of M-viruses (K1, K2, Klus, and K28) are maintained by specific L-A viruses and compete within yeast populations, leading to co-evolution of L-A and M viruses<sup>148</sup>. The L-BC virus co-evolves with L-A in approximately 50% of yeast populations<sup>182</sup>. However, a L-BC-positive/L-A-negative yeast strain can be generated, resulting in the loss of the L-A-associated killer phenotype. This indicates that L-BC cannot act as a helper virus for the toxin-expressing M-virus. Various stable L-A/L-BC combinations can be obtained by mixing laboratory yeast strains, suggesting no interdependency between naturally co-occurring L-A and L-BC viruses<sup>182</sup>.

Capsid structures of various members of the Totiviridae family have been determined<sup>183-192</sup>. Totivirus capsids have icosahedral symmetry and a diameter of 40 nm (**Figure 6, a-c**). Each icosahedral asymmetric unit consists of two copies of a capsid protein. Similar capsid arrangements with two subunits in an asymmetric unit are found in other dsRNA viruses that infect bacteria and eukaryotes<sup>193</sup> (**Figure 6, d**). Some dsRNA viruses have a single protein layer in their capsids, while others possess additional layers with more complex structures. Structural studies have revealed the presence of a polymerase, and its incorporation into the

capsid is crucial for dsRNA virus assembly. The Gag-pol allows viral genome replication and transcription within the virion, evading the host antiviral response<sup>194</sup>. The Gag-pol is composed of a C-terminal extension of the capsid protein due to a ribosomal frameshifting sequence between the two genes: In L-A virus, the ribosomal frameshifting efficiency is 1.9%, resulting in an average of two Gag-pol per virion<sup>146</sup>.

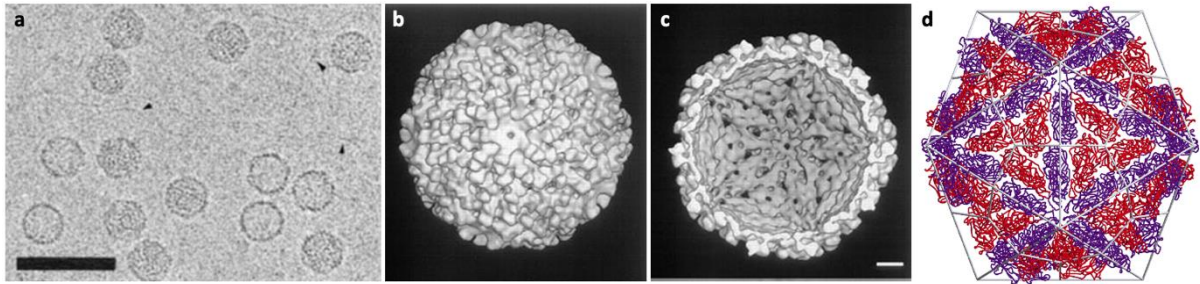


Figure 6. Cryo-EM of the L-A virus.

(a) Empty and full capsids are seen in the cryo-EM specimen; arrows indicate RNA. Scale bar is 100 nm. Figure reproduced and adapted from<sup>161</sup> (b- outer view and c- inner view of the capsid) Icosahedral reconstruction of the L-A virus at  $\sim 16$  Å resolution. Scale bar is 5 nm. Figure reproduced and adapted from<sup>161</sup>. (d) Crystallographic structure of the L-A virus capsid; the A and B subunits of the Gag are shown with red and blue colors, respectively. Figure reproduced and adapted from<sup>183</sup>.

The RNA-binding domain of Gag-pol recognizes a secondary structure at the 3'-end of the L-A virus genome's (+)RNA strand, facilitating genome encapsidation<sup>165</sup>. *Totivirus* Gag-pol generate RNA molecules lacking a cap at the 5'-ends. Consequently, host translation factors do not recognize these viral RNA molecules, making them susceptible to degradation by host cell RNA quality control systems<sup>195</sup>. L-A and L-BC particles have been found to remove the caps from cellular mRNAs, potentially overwhelming the cellular RNA degradation machinery. Moreover, both L-A and L-BC viruses employ a unique cap-snatching mechanism<sup>196</sup> that involves the transfer of the eukaryotic cap 0 (m7G) from a host mRNA to a viral transcript. Initially, m7G from a host mRNA becomes covalently bound to a histidine residue in a capsid protein and is subsequently transferred to a viral mRNA.

#### 1.5.1.1 The crystallographic structure of the yeast L-A virus at 3.4 Å

Information has been derived from the L-A virus capsid resolved by X-ray crystallography at a resolution of 3.4 Å<sup>183</sup>. As previously described, the L-A virus has a diameter of 400 Å and consists of a protein shell with 60 asymmetric dimers of the coat protein (Gag). The conformation and environment of the identical subunits in each dimer likely influence the

distinct transcription and replication process of the virus. The viral mRNA is produced within the virus particles and is released into the cytoplasm for translation by cellular ribosomes. Unlike cellular mRNAs, L-A transcripts lack a 5' cap but the Gag subunit enzymatically removes the 7-methyl-GMP cap from cellular mRNAs, facilitating viral mRNA expression. Mutations in the Gag protein at position His154 result in poor viral expression, but deleting the *ski1/xrn1* genes can restore viral mRNA levels<sup>195</sup>.

Cryo-EM showed that the L-A virus particle structure exhibits T = 1 icosahedral symmetry composed of 12 pentons<sup>161</sup> and its central core is involved in transcription, which may explain their unique virion structure. The L-A particle has a protein shell thickness of 46 Å<sup>183</sup>. It has openings with a diameter of 18 Å at the icosahedral five-fold axes, allowing entry of nucleotide triphosphates and exit of viral mRNA. The L-A particle structure confirmed its composition of 60 pseudo capsomers, each composed of two Gag molecules (A and B) with structural differences in specific surface regions (**Figure 7, a**). The active site of the decapping reaction involves His154 located at the tip of loop 1, forming a trench with three other loops (**Figure 7, b**). This trench resembles the structure of an mRNA capping enzyme called RNA guanylyltransferase<sup>197</sup>. Specific residues, including His154, play crucial roles in the decapping reaction.

The crystal structure revealed, at low resolution, an electron density map, which showed a single layer of density inside the L-A protein, that in turn identified as a portion of the L-A genome. The Gag proteins attract the genome to the inner surface of the capsid. The dsRNA of bluetongue virus (BTV) exhibits a similar layered density with three shells of dsRNA<sup>198</sup>, indicating a higher genome density compared to the L-A virus. All dsRNA viruses synthesize both (+) and (-) strands within the viral particle. The RNA polymerase is anchored to the inner wall of the particle or located at the five-fold axes, requiring movement of the template for efficient replication. The T = 1 structure of L-A virions, with 120 Gag monomers per particle, allows flexibility of the Gag protein to occupy different positions within the viral particle. However, the limited space inside the virion restricts the rate of elongation during RNA synthesis. The elongation rates for L-A and other dsRNA viruses are relatively slow compared to DNA transcriptases and some other RNA polymerases, and for the L-A virus is approximately 100 nucleotides per minute<sup>199</sup>.



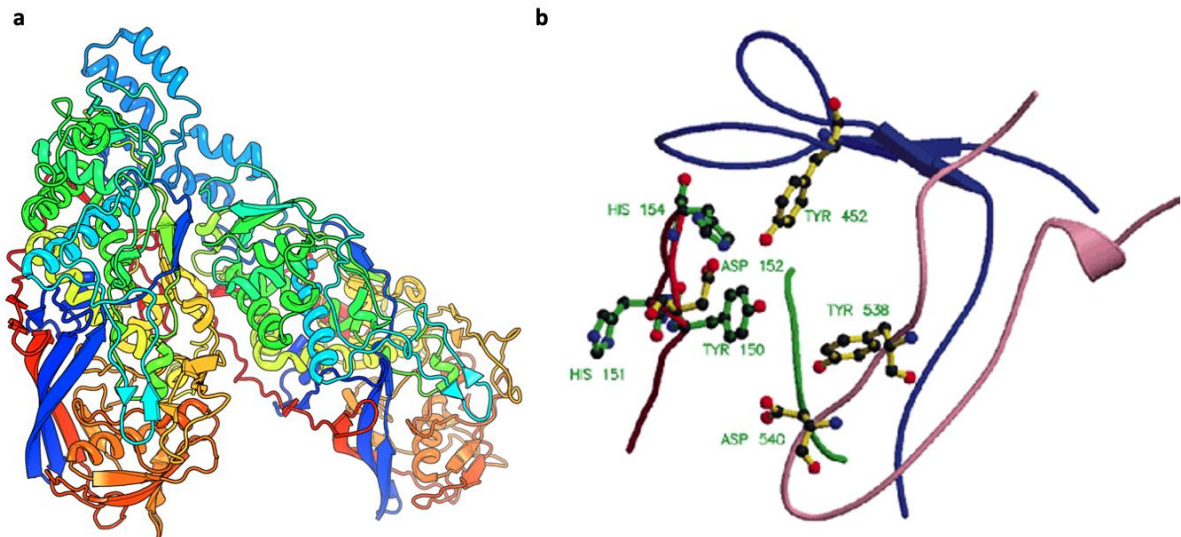


Figure 7. The 3.4 Å resolution of the crystallographic Gag-dimer and the potential active site for mRNA decapping.

(a) Cartoon representation of the two Gag proteins, showing their relative orientation within the asymmetric unit. The two Gag proteins interact via their N-terminal helix bundles. Color code is rainbow, from blue to red and denotes N- to the C- terminus. Structure was downloaded from PDB (PDB-ID: 1M1C) and visualized with Pymol ([www.pymol.org](http://www.pymol.org)). (b) The residues Tyr 150, His 151, Asp 152, Tyr 452, Tyr 538, and Asp 540 are all potential contributors to the crucial mRNA decapping reaction. Additionally, the four loops, namely loop 1, loop 2, loop 3, and loop 4, have been color-coded in red, green, blue, and pink, respectively. Figure panel reproduced and adapted from<sup>183</sup>.

#### 1.5.1.2 The cryo-EM structure of the yeast L-BC virus at 2.9 Å

Another structure confirming the biochemical and biophysical studies for the Totiviridae family is that of the L-BC capsid structure<sup>192</sup>. Here, the putative mRNA decapping site in L-BC Gag resembles that of the L-A virus. The L-BC capsid proteins form interactions that stabilize a decamer, which serves as the fundamental unit of the L-BC capsid. Additionally, they observed an open L-BC particle lacking one decamer of capsid proteins. The study also explores the integration of Gag-pol into the L-BC capsid and the role of ribosome frameshifting site in virion assembly.

When analyzing purified L-BC, the researchers classified particle images into empty particles, open particles, and genome-containing virions. They found that 87% were empty, 9% were open, and 4% contained the viral genome. Different resolutions were calculated for structures of the empty and open particles, and virions. The L-BC capsid displayed icosahedral symmetry and consisted of two subunits of capsid protein, A and B – like L-A. The A subunit connected the twofold and fivefold axes of symmetry, while the B subunits

interacted around the threefold axes. The capsid comprised twelve decamers, each composed of five dimers of the capsid protein (**Figure 8, a-c**).

The L-BC capsid exhibited similarities to the capsids of L-A virus and other Totiviruses found in protists and fungi. However, the arrangement of A and B subunits in the L-BC capsid differed from that of other icosahedral viruses. Compared to L-A, the L-BC capsid had a smoother surface, and the capsid proteins formed ridges around the threefold axes. The L-BC virion's capsid cavity contained a spherical shell representing the dsRNA genome. The interactions among L-BC Gags were stabilized by electrostatics and surface complementarity. Various amino acid residues formed salt bridges and hydrogen bonds within and between the asymmetric units, contributing to the stability of the interfaces. The interactions between decamers of capsid proteins involved the stacking of beta-sheets and hydrophobic interactions. Domain swaps between A subunits played a role in stabilizing the decamer structure.

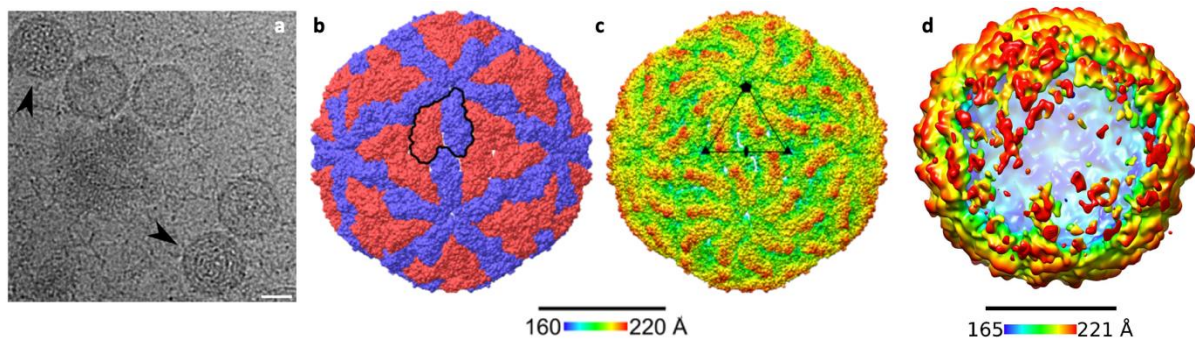
The role of specific residues in the capsid protein in binding the 7-methylguanosine (m7G) caps of host mRNAs was also discussed<sup>192</sup>: His156 of L-BC was identified as the residue involved in binding the m7G cap, and both L-BC and L-A putative decapping pockets had predominantly negative charges. The covalent attachment of the m7G cap to His156 of L-BC can well depend on Mg<sup>2+</sup> ions<sup>200</sup>.

L-BC particle image analysis revealed that ~9% of the particles had missing segments in their capsids, indicating an open structure. Asymmetric reconstructions of these open particles centered around a fivefold axis, exhibiting poorly resolved density at the borders of the capsid opening. A three-dimensional reconstruction with imposed fivefold symmetry resulted in a capsid structure with a resolution of ~10 Å, showing the absence of one decamer of Gags (**Figure 8, d**). The open particle displayed a 16 nm diameter opening. Notably, fifteen capsid proteins bordering the opening exhibited smeared densities, indicating high mobility. These border subunits were displaced by 9 to 11 Å compared to the complete capsid, moving into crevices and openings within the capsid.

The genomes of L-A and L-BC viruses share the same arrangement of open reading frames, suggesting that the Gag-pol of L-BC is also expressed as a C-terminal extension of the capsid through ribosomal frameshifting. In the L-A virus, the majority of decamers (82.5%) comprised only capsid proteins, while 16% contained one RdRp, and 1.5% contained 2 or more RdRps; such ratios are highly specific<sup>146</sup>. Purified L-BC contained a high proportion of empty particles (87%) compared to virions (4%), suggesting enrichment or induced genome release during sample preparation. Potential particle damage may have resulted from freeze-

thaw cycles or low ionic strength in the buffer. Cryo-EM studies did not reveal the polymerase inside the virus particle, likely due to variable positions of the Gag-pol domains within the RNA genome.

Existence of domain swaps between A subunits and the presence of open particles indicate that L-BC virus capsids are assembled from decamers. A proposed three-step assembly pathway may involve the formation of capsid protein dimers, subsequent decamer formation, and final capsid assembly from 12 decamers. This pathway shares similarities with the assembly of Bluetongue virus subcore<sup>201</sup>, which shares the same symmetry type as Totivirus capsids. The conserved arrangement of the capsid shell surrounding the genome among dsRNA viruses suggests a common evolutionary origin and a likely conserved assembly pathway.



**Figure 8. Structural analysis of the yeast L-BC virus.**

(a) Cryo-EM representative image showing full and empty LB-C viruses; Scale bar is 20 nm. (b-c) The icosahedral yeast LB-C virus resolved at 2.9 Å resolution. With red and blue, the Gag subunits are shown; the black line engulfs the asymmetric dimer; (c) Same as (b), but color coding indicates the distance from the geometric center of the map; the color coding is shown below the (b) and (c) panels. Scale bar is 20 nm. (d) The open conformation of the yeast LB-C virus shows the absence of a decameric face of the icosahedron. Color coding and distance from the geometric center are highlighted in a similar manner as in (b). Scale bar is 20 nm. All figure panels are reproduced and adapted from<sup>192</sup>.

## 2. Aim of this thesis

While overexpressed protein systems offer various advantages, one significant drawback is that an excessive amount of a specific protein could potentially modify the outcome concerning pathway activation compared to its natural state. In contrast, native cell extracts possess an advantage in maintaining their natural expression ratio. Consequently, this allows for direct observation of biochemical processes, higher molecular assemblies, and their complexity, as well as protein ratios. Based on these observations, the primary objective of this thesis is to identify and analyze megadalton complexes, specifically focusing on the yeast L-A virus, within native cell extracts structurally and biochemically. To accomplish this goal, the main aim is further divided into smaller, more detailed aims as outlined below:

- ❖ Adapt and refine the process of yeast cell lysis, followed by the purification of the lysate for both structural analysis and biochemical investigations. This involves evaluating and enhancing protocols for isolating both mitochondria and cell lysates. The choice of *Saccharomyces cerevisiae* as the model organism is based on its rapid growth, availability of comprehensive genomic and proteomic data, and the ability to extrapolate the findings to other eukaryotes.
- ❖ Characterize the structural signatures identified in native cell extracts by employing cryo-electron microscopy (cryo-EM) and computational pipelines. This approach aims to provide valuable insights into the complexity, location, and interacting partners of these signatures, taking the L-A virus as an illustrative example.

### 3. Materials and Methods

#### 3.1 Materials

Table 1. Chemicals and Enzymes used in this study

<b>Chemicals and enzymes</b>	<b>Source</b>	<b>Identifier</b>
<b>1,4-Dithiothreitol DTT, min. 99 % p.a.</b>	Carl Roth	6908.4
<b>Acrylamide/Bis solution, 37.5:1</b>	Serva	10688.01
<b>Agar-Agar, bacteriological highly pure</b>	Carl Roth	2266.3
<b>Ammonium acetate <math>\geq 97</math> % p.a. ACS</b>	Carl Roth	7869.2
<b>Ammonium persulfate</b>	Serva	13376.02
<b>Aprotinin from bovine lung</b>	Sigma-Aldrich	A1153-1MG
<b>Bestatin, 10 mg</b>	Sigma-Aldrich	10874515001
<b>Roti®-Quant 5x</b>	Carl Roth	K015.1
<b>Clear Western ECL substrate</b>	BIO-RAD	170-5060
<b>Calcofluor white</b>	Merck	18909-100ML-F
<b>D-Sucrose <math>\geq 99.5</math> % p.a.</b>	Carl Roth	4621.1
<b>D (+) Glucose p.a. ACS anhydrous</b>	Carl Roth	X997.2
<b>D-Sorbit <math>\geq 98</math> % C</b>	Carl Roth	6213.1
<b>Dextrin for microbiology (from potato starch)</b>	Carl Roth	3488.1
<b>di-potassium hydrogen phosphate trihydrate</b>	Carl Roth	6878.1
<b>di-potassium hydrogen phosphate trihydrate <math>\geq 99</math> % p.a. anhydrous</b>	Carl Roth	P749.1
<b>DNase I</b>	Sigma-Aldrich	10104159001
<b>E-64</b>	Sigma-Aldrich	E3132-1MG
<b>EDTA disodium salt hydrate min 99 % p.a. ACS</b>	Carl Roth	8043.2
<b>Glycine</b>	Serva	23391.02
<b>HEPES PUFFERAN®</b>	Carl Roth	9105.3
<b>HOECHST 33342</b>	Invitrogen	H21492
<b>Isopropanol</b>	Carl Roth	CP41.1
<b>Leupeptin</b>	Sigma-Aldrich	L2884-1MG
<b>Magnesium chloride hexahydrate, min 99% p.a. ACS</b>	Carl Roth	2189.1
<b>Methanol</b>	Carl Roth	4627.6
<b>Milk powder</b>	Carl Roth	T145.3
<b>Mito Track Green FM</b>	Invitrogen	M-7514
<b>Mito Track Orange FM</b>	Invitrogen	M7511
<b>Nycodenz</b>	Serva	11731508
<b>Pefabloc</b>	Sigma-Aldrich	11585916001
<b>Pepstatin A</b>	Sigma-Aldrich	77170-5MG

<b>Peptone ex casein</b>	Carl Roth	8986.1
<b>Phosphate buffered saline tablets (PB)</b>	Sigma-Aldrich	P4417
<b>Potassium chlorine min 99.5 % -1kg</b>	Carl Roth	6781.1
<b>Potassium dihydrogen phosphate <math>\geq 99</math> % p.a. ACS</b>	Carl Roth	3904.2
<b>Precision plus protein all blue standards (marker)</b>	BIO-RAD	161-0373
<b>PCC</b>	MERCK	26299-14-9
<b>Sodium chloride 99.5 % p.a. ACS, ISO</b>	Carl Roth	3957.2
<b>Sodium dodecyl sulfate (SDS)</b>	Carl Roth	0183.2
<b>Sodium hydroxide <math>\geq 98</math> % p.a. ISO in pellets</b>	Carl Roth	6771.1
<b>Sodium nitrate <math>\geq 99</math> % p.a. ACS ISO</b>	Carl Roth	A136.1
<b>TEMED</b>	Carl Roth	2367.3
<b>Tris</b>	Carl Roth	AE15.2
<b>Tris hydrochloride</b>	Carl Roth	9090.2
<b>Tween 20</b>	Carl Roth	9127.1
<b>Tryptone</b>	Sigma-Aldrich	T7293
<b>Yeast extract, micro-granulated</b>	Carl Roth	2904.3

Table 2. Antibodies

<b>Antibody</b>	<b>Source</b>	<b>Identifier</b>
<b>Rabbit polyclonal antibody a-E1<math>\alpha</math> against <i>C. thermophilum</i> E1<math>\alpha</math>-His-Tag (37-411)</b>	Custom-made by Genscript	RRID: AB_2888984
<b>Rabbit polyclonal antibody a-E2p against <i>C. thermophilum</i> E2p-His-Tag (29-459)</b>	Custom-made by Genscript	RRID: AB_2888985
<b>Rabbit Anti-GAPDH antibody loading control</b>	abcam	ab9485
<b>Mouse Anti-beta Actin antibody [8F10G10]</b>	abcam	ab170325
<b>Mouse Anti-Actin antibody [C4]</b>	abcam	ab14128
<b>Mouse Anti-VDAC1/Porin antibody [16G9E6BC4]</b>	abcam	ab110326
<b>Rabbit Anti-Capsid ScVLA</b>	Custom-made by Genscript	SC2039-PF NA-14
<b>Rabbit Anti-Polymerase ScVLA</b>	Custom-made by Genscript	SC2033-PF NA-24
<b>Goat Anti-Rabbit IgG H&amp;L (HRP)</b>	abcam	ab205718
<b>Goat Anti-Mouse IgG H&amp;L (HRP)</b>	abcam	ab205719

Table 3. Instruments and main equipment used in this study

<b>Instrument</b>	<b>Type</b>	<b>Company</b>
<b>Incubator</b>	Heracell 150i	Thermo Fisher Scientific
<b>Tabletop centrifuge</b>	Heraeus Megafuge 40R	Thermo Fisher Scientific
<b>Bead Beater</b>	FastPrep-24™ 5G	MP Biomedicals™
<b>Ultracentrifuge</b>	Optima™ MAX-XP (TLA110)	Beckman Coulter
<b>FPLC-system</b>	Äkta pure 25M	Cytiva (GE Healthcare)

<b>Plate reader</b>	Epoch 2 Microplate Spectrophotometer	Agilent (bioTek)
<b>Douncer</b>	VOS 40 Digital	VWR
<b>Gel imaging system</b>	ChemiDoc™ MP Imaging Systems	BIO-RAD
<b>Thermomixer</b>	Thermomixer C	Eppendorf
<b>Glow discharge cleaning system</b>	PELCO easiGlow	Ted Pella Inc.
<b>Vitrification instrument</b>	Vitrobot Mark IV System	Thermo Fisher Scientific
<b>Electron microscope</b>	Thermo Fisher Scientific Glacios Cryo Transmission Electron Microscope (Cryo-TEM) 200 kV	Thermo Fisher Scientific
<b>Electron microscope</b>	Zeiss Libre 900	Carl Zeiss AG
<b>Camera microscope</b>	Falcon 3EC Direct Electron Detector	Thermo Fisher Scientific
<b>Power module Electrophoresis and Blot System</b>	PowerPac™ HC High Current Power Supply	BIO-RAD
<b>Electrophoresis Chamber</b>	Mini-Protean Tetra Cell	BIO-RAD
<b>Blotting instrument</b>	Trans-Blot® Turbo™ Transfer System	BIO-RAD
<b>Ultramicrotome</b>	Ultracut S,	Leica
<b>Fluorescence microscope</b>	LSM880 Airyscan system A 40X objective	Zeiss

## 3.2 Methods

### 3.2.1 Model Organism culture

*Saccharomyces cerevisiae* Hansen (D273-10B Genotype MAT $\alpha$  mal [rho+]) was obtained from ATCC (American Type Culture Collection PO Box 1549 Manassas, VA 20108 USA; ATCC® 24657™) in a freeze-dried lyophilized ampoule. The Media and culture conditions proposed by the Company were used (Media 1245 Agar or Broth ATCC Media list). To optimize the growth conditions, different Temperatures (26 °C as submitted by ATCC as well as 30 °C as suggested by several protocols) and time points for harvesting were screened to find the optimal growing conditions. In the end, a temperature of 26 °C at 5 hours (OD<sub>600</sub> = ~5) of growth time in YPD media was considered most suitable for our experiments. For liquid YPD medium, 20 g·L<sup>-1</sup>, peptone, and 10 g·L<sup>-1</sup> Yeast extract are dissolved in 900 mL ddH<sub>2</sub>O and autoclaved. Additionally, 20 g·L<sup>-1</sup> glucose in 100 ml is autoclaved separately and added to the 900 mL YP mix after sterilization to avoid

caramelization. To make solid culture plates, 15 g agar-agar is added to the YP mix but otherwise prepared the same way and then poured into sterile Petri dishes after autoclaving. To rehydrate the freeze-dried culture, the ampoule was opened by continuously heating and cooling it by, in turn, holding the ampoule over the flame of a Bunsen burner and pipetting cold deionized water over the heated surface until it cracks. Then the glass tip is hit with forceps or a tweezer to break off. The inner vial could now be removed from the glass ampoule. Under sterile conditions, the cotton plug can then be removed. Now 0.5 mL of sterile ddH<sub>2</sub>O is added, and the pellet is resuspended. The resuspended pellet is transferred into a sterile culture tube with an additional 5 mL of sterile ddH<sub>2</sub>O and is left to rehydrate on the workbench overnight. After rehydration, 200 µL is pipetted onto 2 YPD-Agar plates and spread with a Drigalski spatula. From these plates, colonies are picked for glycerol stock creation. The subsequent protocol steps are as follows: Add one colony to 10 mL of YPD medium and let it grow to stationary phase. Stationary phase was achieved by measuring the OD<sub>600</sub> until the absorbance reached a plateau. Now 500 µL of the stationary culture and 500 µL of sterile 30 % glycerol are mixed and flash-frozen in liquid nitrogen. This stock can now be used for inoculating initial cultures and stored at -80 °C. To inoculate a culture from glycerol stock, a small inoculum is scraped off with a sterile pick or inoculation loop and then smeared onto a YPD-agar plate. Avoid thawing the glycerol stocks by working fast and keeping the stock cool. Incubate the plates at 26 °C until colonies are visible. To prepare a preculture from the plates, a colony is picked with a sterile inoculation loop and suspended in 50 mL YPD medium in a sterile Schikane flask. After incubating the flask overnight at 26 °C in a rotary shaker (195 rpm), the main culture can be inoculated. For this, 10 ml of the preculture is added to 500 ml of YPD in a new sterile Schikane flask and incubated for 5 h at 26 °C and 195 rpm.

### 3.2.2 Culture harvest and lysis

Yeast cells are harvested at 4 °C. The culture flasks are taken out of the rotary shaker and transported on ice. The liquid culture is poured into centrifuge bottles and harvested while centrifuging at 3,000 g for 5 min at 4 °C. After discarding the supernatant, the pellet is washed with ddH<sub>2</sub>O to remove the leftover media. After discarding the water after washing, the pellet is flash frozen in liquid nitrogen and stored at -80 °C until needed for experiments. Approximately 10 g can be harvested from 1 l of culture. For the lysis 10 g culture pellet is lysed in 20 mL lysis buffer (100 mM HEPES pH 7.4, 95 mM NaCl, 5 mM KCl, 1 mM MgCl<sub>2</sub>, 5% glycerol, 0.5 mM EDTA, 1 mM DTT, 10 µg·mL<sup>-1</sup> DNase, 2.5 mM pefabloc, 40



$\mu\text{M}$  E-64, 130  $\mu\text{M}$  bestatin, 0.5  $\mu\text{M}$  aprotinin, 1  $\mu\text{M}$  leupeptin, 60  $\mu\text{M}$  pepstatin A). For lysis add 10 mL glass beads and 3 mL lysis buffer into 15 mL falcon tubes for equilibration (use 4 tubes). Resuspend the pellet with the remaining 8 mL of lysis buffer and distribute evenly into the 4 tubes. Now the tubes are subjected to bead beating using a Fastprep at 4 °C with 3 rounds of shaking at 6.5 mps for 20 sec with 3 min rest on ice between each round. To remove large debris the cell lysate is centrifuged for 5 min at 4,000 g at 4 °C and right after subjected to ultracentrifugation at 100,000 g for 45 min at 4 °C. The supernatant is filtered (0.22  $\mu\text{m}$  syringe filter) and then concentrated by spin filtration (100 kDa cut-off amicon ultra centrifugal filter) at 4 °C and 4,000 g. The buffer is mixed every 10 min. Protein concentration was measured with Bradford reagent to determine the optimal protein amount for further experiments.

### 3.2.3 HPLC and Size exclusion chromatography (SEC)

For size exclusion chromatography, a Biosep 5  $\mu\text{m}$  SEC-s4000 500 Å, LC Column 600 x 7.8 mm column is injected with 500  $\mu\text{L}$  lysate at 30  $\text{mg}\cdot\text{mL}^{-1}$  protein concentration into a 500  $\mu\text{L}$  Loop. The system used was an ÄKTA Pure 25 M (FPLC) system, and the fractionation was performed at a 0.15  $\text{mL}\cdot\text{min}^{-1}$  flow rate, and fractions of 250  $\mu\text{L}$  were collected (**Figure 9**). Prior to the run, the column was pre-equilibrated with 50 mL running buffer. As running buffer, 200 mM  $\text{NH}_4\text{CH}_2\text{COOH}$  pH 7.4 was chosen. The resulting fractions were measured with Bradford reagent and then directly used for sample preparation for data collection on the cryo-electron microscope or flash frozen and stored at -80 °C until further use.

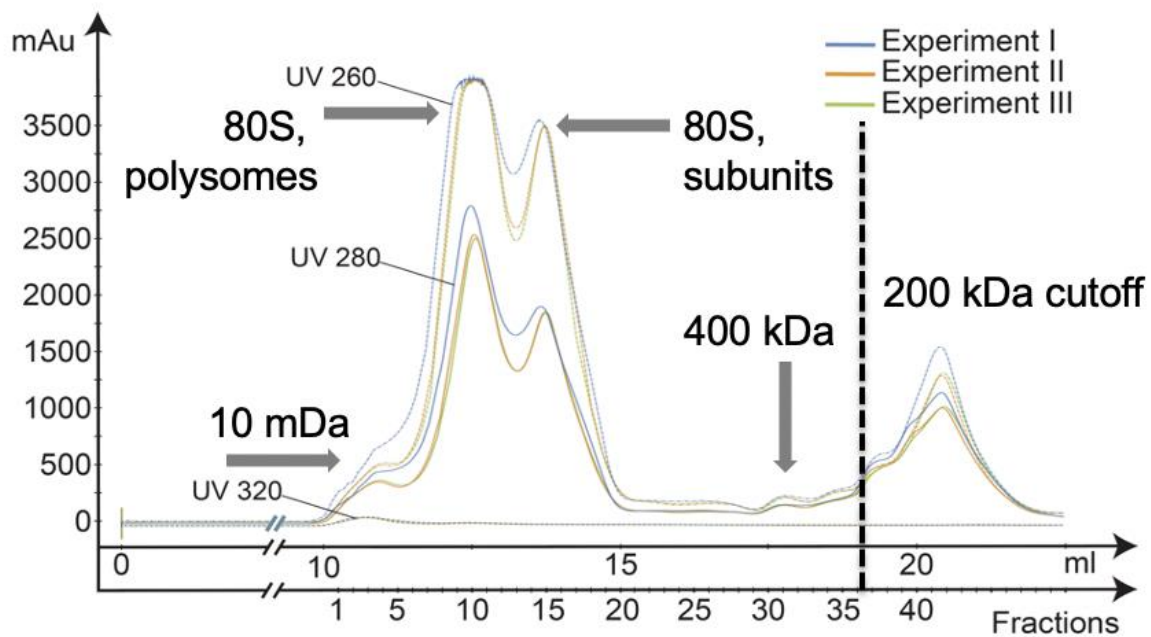


Figure 9. SEC triplicate for the recovery of fractionated yeast cell extracts.

A shoulder is seen starting at 10 mL run time of the SEC. 10 mDa complexes and macromolecules are expected at these fractions (F1-F5) as shown in publications of AG Kastiris<sup>51-53,69</sup>. At 12.5 mL the first peak includes polysomes and monosomes. Ribosome subunits are expected at the second peak (14 mL or fractions F14-F18). In later fractions proteins with lower molecular weight are found, for example proteasomes; fractions past the 200 kDa cutoff show elution of small complexes. Three different absorbances were measured (260 nm for maximal absorbance of DNA; 280 nm for maximal absorbance of proteins; and 320 nm for maximal absorbance of aggregates).

### 3.2.4 Sample preparation for cryo-EM

A regularly patterned holey carbon-coated copper grid (Quantifoil R2/1 type holey carbon film on 200 mesh support grid) was used to prepare the Cryo-EM samples. In order to make the grids hydrophilic, they were plasma treated using a PELCO easiGlow glow discharging machine. A Vitrobot Mark IV was used to plunge freeze the samples: A 3.5  $\mu\text{L}$  aliquot of the sample at a concentration of  $\sim 0.3 \text{ mg}\cdot\text{mL}^{-1}$  was applied on the carbon side of each grid. The grids were blotted using standard 595-grade filter paper and quickly plunge frozen in liquid ethane. The glow discharging and vitrification parameters are summarized in **Table 4**.

Table 4. Vitrification conditions

Vitrification Parameters	Value
Humidity	100%
Temperature	4 °C
Blotting time (sec)	6
Blot force (a.u.)	2
Sample volume (µL)	3.5
Blot offset (mm)	2
Protein concentration (µg/µL)	0.26
<b>Glow discharging Parameters</b>	
Current	15mA
Polarity	Grist Negative
Atmosphere	0.4 bar
Glowing Time	25s

The vitrified grids were clipped in the Thermo Fisher Scientific Autogrid assemblies using standard tools. They were subsequently loaded onto a Thermo Fisher Scientific Glacios cryogenic transmission electron cryo-microscope operating at 200 kV accelerating voltage, under cryogenic conditions.

### 3.2.5 Data collection of vitrified grids – microscope alignments

The screening and automated data acquisition were set up using Thermo Fisher EPU software (Version: 2.11.1.11REL). After screening the sample for various ice thicknesses and particle distributions, regions with rather thick ice were selected for capturing larger complexes. The microscope was adjusted and aligned according to the Thermo Fisher Scientific guidelines for microscope alignments (2019, Single Particle Analysis Workflow) which is further described in detail in the supplementary material of Hamdi et al.201. The images were recorded using the Falcon 3EC direct electron detector in linear (integrating) mode. The acquisition parameters are summarized in **Supplement Table 1**. The objective lens had a focal length of 3.4 mm and a spherical aberration of 2.7 mm. The beam was restricted by a 100 µm objective aperture. **Figure 10** summarizes the biochemical workflow (**Figure 10, a**) as well as the steps from the biochemical preparation to the acquisition of images at the cryo-microscope (**Figure 10, b**).

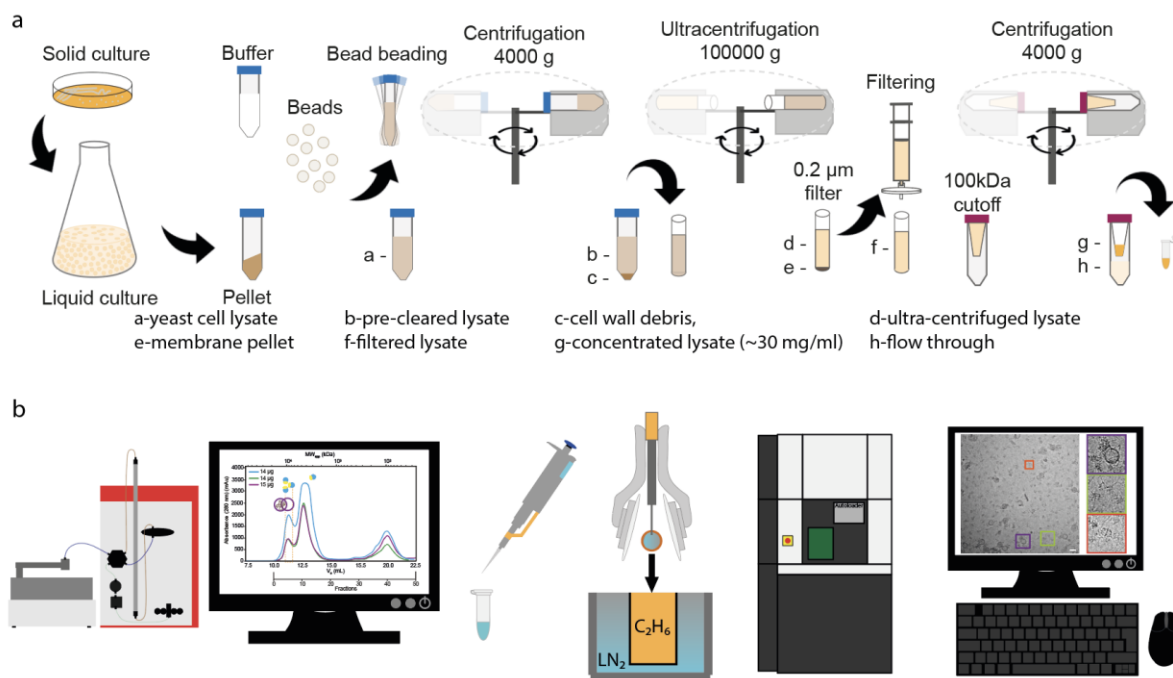


Figure 10. Full workflow of native cell extracts from culture to cryo-EM micrographs.

Yeast cells are plated on YPG media, and from there, the liquid culture for the experiment can be inoculated. Cells are washed with water after harvesting then lysed via bead beading (a = LC1) and then centrifuged to remove cell debris (b = LC2 & c = LC3). This is followed by ultracentrifugation to remove organelles and membrane parts (d = LC4 & e = LC8). The lysate was then filtered with a 0.22  $\mu\text{m}$  filter (f = LC5) and concentrated (g = LC6 & h = LC7) to approximately 30  $\text{mg}\cdot\text{mL}^{-1}$  using a 100 kDa amicon filter. The LC numbers mentioned above correspond to those given in a later figure for measured activity assay, Fig. 23. (b) The resulting concentrated lysate is then subjected to SEC column mounted on an FPLC, and the fraction of interest is vitrified with the ThermoFisher Vitrobot Mark IV and imaged using the Thermo Fisher Glacios. A data set is then acquired (b).

### 3.2.6 Image processing and analysis

Two datasets were acquired for this work. One large dataset with 10 acquisitions of fractions 5 and 6 at a pixel size of 3.17  $\text{\AA}$  and a second with 2 replicates of fraction 5 at a pixel size of 1.5678  $\text{\AA}$ . The 3.17  $\text{\AA}$  dataset was first analyzed with Scipion<sup>3202</sup>, and the refined map was then transferred to cryoSPARC<sup>79</sup> for the rest of the analysis. The dataset with the pixel size of 1.5678  $\text{\AA}$  was processed using cryoSPARC only.

The image analysis<sup>203</sup> starts with data import, motion correction<sup>204,205</sup>, and contrast transfer function (CTF) correction of the aligned movie frames<sup>76,77</sup>. Next steps, which are basically particle recognition and averaging are recently fully automated with machine learning implementations<sup>206</sup>. Particles are then picked either manually or automatically and sorted via 2D classification algorithms, and an *ab initio* reconstruction of the Coulomb potential map

can be performed. This process can be done in consecutive rounds. Then, 3D refinement, subsequent 3D classification, and final refinement are iteratively applied. The end cryo-EM map is derived when resolution cannot improve anymore; with the resulting map, a model can be built.

The analysis in Scipion3 was done as follows: After the import of the datasets, a correction of beam-induced motion drift was performed for all movies in Scipion as well as in cryoSPARC. Now, these dose-weighted, drift-corrected sets were subjected to CTF estimation. For manual picking, approximately 500 particles were picked using xmipp3<sup>207</sup> while training the algorithm for auto-picking. After auto-picking, also with xmipp3, ~230,000 picked particles were extracted using Relion<sup>208</sup>. Then cryoSPARC was used for 2D classification. The coordinates from the 2D classes were then extracted using pwem<sup>209</sup> and used for another round of auto-picking using xmipp3. Then ~65,000 particles were extracted again and subjected to 2D classification. With the help of cryoSPARC, an initial 3D model was created. This initial model was then refined using different symmetries. Symmetrized maps were superimposed to an asymmetrically refined map to derive the symmetry of the structure, which was icosahedral. The initial model was also processed in cryoSPARC. Symmetry expansion was performed<sup>79</sup>, which resulted in roughly 30,000,000 particles for 3D classification, refinement, and a 6.4 Å (FSC=0.143) icosahedral map.

The data set acquired at a pixel size of 1.5678 Å was entirely analyzed in cryoSPARC. After import, micrographs were subjected to motion correction, and after curation of the blurry or empty micrographs, also CTF correction. Since there was already a low-resolution map that could be used as reference, initial manual picking was omitted, and a template of 10 projections of the 6.4 Å (FSC=0.143) refined map was created for template picking. The picked micrographs were extracted, and 2.7 million particles were used for 2D classification. The most promising-looking classes, or to be precise classes with similar features, were selected, and another round of 2D classification was performed. From there, 17,000 particles from selected classes and as reference, the 6.4 Å (FSC=0.143) map were used for 3D reconstruction. This resulted in a 7.3 Å (FSC=0.143) asymmetric map. Then, icosahedral symmetry was applied, and symmetry expansion was performed, and the end result is a 3.78 Å (FSC=0.143) icosahedral structure of the *Saccharomyces cerevisiae* LA virus. In the end, the FSC was calculated and the local resolution was derived according to the gold standard FSC<sup>114</sup>. The icosahedrally averaged map containing two independent protomers was of sufficient quality for a C $\alpha$  trace to be built using Coot<sup>210</sup>. The X-ray structure of the L-A virus (PDB ID: 1m1c) was identified as the best hit through a subsequent DALI search<sup>211</sup>.

Thereafter, the model was rigid-body fitted into the density using Chimera<sup>212</sup> and ChimeraX<sup>213</sup> and was subjected to iterative cycles of real space refinement using PHENIX<sup>100</sup> and manual refinement using Coot to yield the final coordinates.

### 3.2.7 Activity assay against PDHc

An activity assay Kit for the PDHc activity is commercially available and was used for all steps of the lysis protocol, named LC1-LC8 (see for an explanation the **Figure 10** legend). Material derived from the isolation steps of mitochondria were also measured. The procedure was executed by the provider's instructions. This activity assay kit is extensively described in Fotis Kyrilis Dissertation<sup>27</sup>.

### 3.2.8 MS analysis of fractions 5 and 6 sample preparation

In-solution digestion was performed for sample preparation for mass spectrometric (MS) analysis. Approximately 10 µg of fractions 5 or 6 were transferred into low bind reaction tubes, and 4 equivalent volumes of ice-cold acetone were added and incubated for 60 min at -20 °C. After precipitation, the sample was centrifuged at 2,0000 g for 10 min, the supernatant was discarded, and the precipitated proteins air dried. Protein pellets were resuspended and denatured by adding 25 µL 8 M Urea / 400 mM AmBic in water into the reaction tubes and sonicated for 5 min. The samples were reduced by 5 µL 45 mM DTT (30 min at 50 °C) and alkylated with 5 µl of 100 mM 2-chloroacetamide (30 min at 37 °C) before digestion. Digesting the samples was done after adjusting the sample volume to 200 µL with water by adding trypsin (Promega Sequencing Grade Modified Trypsin) using a 1:50 (w/w) enzyme:protein ratio for 16 hours at 37 °C. Afterwards, the reaction was stopped by adding 10 µL of 10 % (v/v) TFA and cooled down to room temperature (approximately 10 min). Acidity was verified by pH paper. 20 µl of digestion mixtures were analyzed by LC/MS/MS using a U3000 nano-HPLC system coupled to a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). Peptides were separated on reversed phase C18 columns (trapping column: Acclaim PepMap 100, 300 µm × 5 mm, 5 µm, 100 Å, Thermo Fisher Scientific; separation column: µPAC 50 cm C18, Pharmafluidics). After desalting the samples on the trapping column, peptides were eluted and separated using a linear gradient ranging from 3 % to 35 % B (solvent A: 0.1 % (v/v) formic acid in water, solvent B: 0.08 % (v/v) formic acid in acetonitrile) with a constant flow rate of 300 mL·min<sup>-1</sup> over 180 min. Data were acquired in data-dependent MS/MS mode with higher-energy collision-induced dissociation (HCD), and the normalized collision energy was set to 28%. Each high-resolution full scan (m/z 375 to

1799,  $R = 140,000$  at  $m/z$  200) in the Orbitrap was followed by high-resolution fragment ion scans ( $R = 17,500$ ) of the 10 most intense signals in the full-scan mass spectrum (isolation window 2 Th); the target value of the automated gain control was set to 3,000,000 (MS) and 200,000 (MS/MS), maximum accumulation times were set to 50 ms (MS) and 120 ms (MS/MS). Precursor ions with charge states  $<2+$  and  $>6+$  were excluded from fragmentation. Dynamic exclusion was enabled (duration 60 sec, window 3 ppm).

### 3.2.9 Crosslinking MS of fractions 4 and 5 sample preparation

As an experiment prior to the crosslinking experiment itself, different crosslink concentrations were tested, and then a decision for one concentration to be used was made. Crosslinker DSBU is diluted in DMSO to 10 mM, 5 mM, 2.5 mM, 1.25 mM, 0.6 mM, and 0.3 mM. 9  $\mu$ l of fraction F5 was crosslinking using 1  $\mu$ l of the prepared crosslinker dilution. As control, 1  $\mu$ l DMSO was used. Crosslinking reaction was performed at room temperature for 30 minutes and were stopped by adding 10  $\mu$ L 2x SDS-loading dye. All reactions were then run on an 8% SDS-poly-acryl-amide-gel (**Figure 11**).

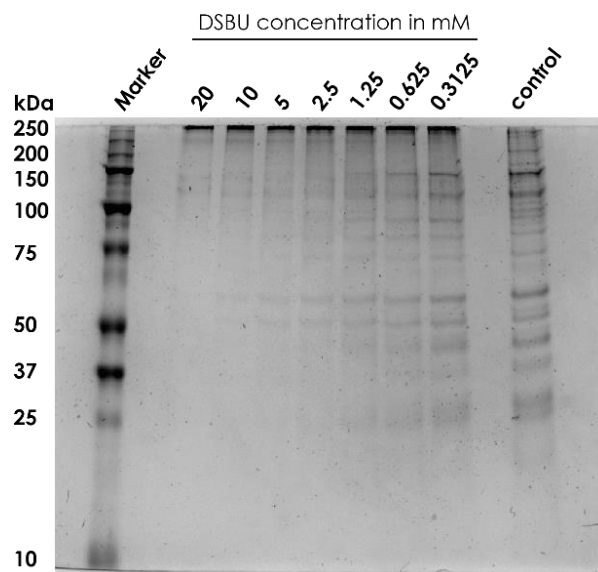


Figure 11. Experiment conducted on the native cell extract fraction to benchmark the crosslinker concentration.

Shown are SDS results for Fraction 5 of the native cell extract incubated with decreasing concentrations of DSBU crosslinker. As control the fraction was incubated with DMSO without any crosslinker. The concentration selected for the crosslinking experiment was the one with the highest number of faded bands and lowest concentration, indicating that DSBU had not yet reacted with all the proteins. For this experiment, a concentration of 10 mM DSBU fulfilled this requirement.

For the cross-linking experiment, ~0.2 µg was used, and the volume of the reaction is adjusted to 45 µL and 5 µL of 5 mM DSBU was added. After 30 min at 25 °C, the reaction was stopped by adding 5 µL 1 M Tris-HCl (pH 7.4) and mixed by vortexing. A small sample of 7.5 µL was taken out as a control for SDS-PAGE; to the rest, 4 volumes of ice-cold acetone are added, mixed, and left for 1 h at -20 °C. Then the samples were centrifuged for 10 min at 14000 g. The supernatant was removed, and the samples were left to evaporate the residual acetone. Samples were further processed as described in 3.2.8.

#### 3.2.9.1 Identification of proteins

To identify containing proteins, the MS raw files were analyzed using MaxQuant (v1.6.6), and the Yeast proteome UP000002311 as reference. Label-free quantification/LFQ and ‘Match between runs’ were additionally activated during analysis. Protein abundance was reported in the proteinGroups.txt and manually analyzed.

#### 3.2.9.2 Identification of cross-linked peptides

For identification of cross-linked peptides, the MS raw files were converted to mzXML files by Proteome Discoverer™ and the MeroX<sup>214,215</sup> software was used. As sequence database, the most abundant proteins from the proteomics analysis were used, and analysis was performed in RISE mode with default parameters. Crosslinked spectra were manually validated.

#### 3.2.10 Identification of structural homologs

The DALI webserver<sup>211</sup> was used to identify structural homologs. As input structure, the region 491-587 from the LA-virus chain A was used. To identify suitable hits, the “*Matches against full PDB*” option was used, and a scoring threshold was set to  $Z > 3.0$ , which is an optimized similarity score between two structures: The higher the Z-score, the greater the similarity.

#### 3.2.11 Western Blot analysis

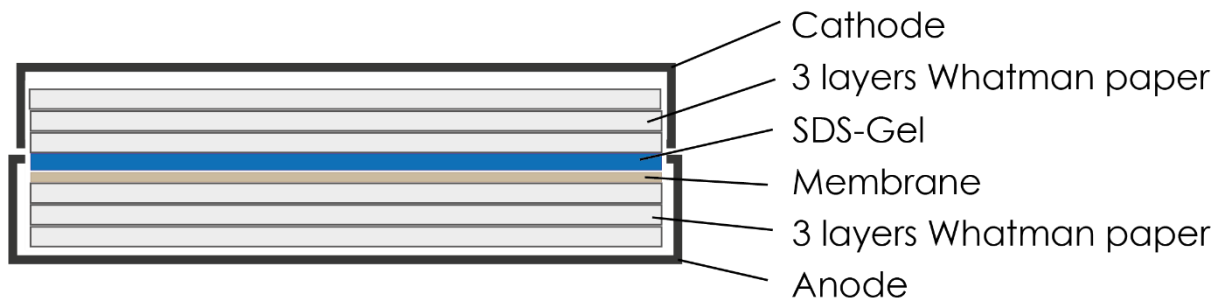
The gels were freshly casted in-house prior to the experiment using a separating gel: 10 % (w/v) acrylamide (37.5:1), 0.1 % (w/v) SDS, 0.04 % (w/v) APS, 0.002 % (w/v) TEMED in 370 mM Tris-HCl solution pH 8.8 and stacking gel: 5% (w/v) acrylamide (37.5:1), 0.1 % (w/v) sodium dodecyl sulphate (SDS), 0.04 % (w/v) APS, 0.002 % (v/v) TEMED in 125 mM Tris-HCl-solution pH 6.8 (**Table 5**). The gels cast were 1 mm thick. The samples were mixed



with a 4x loading dye (250 mM Tris-HCl (pH 6.8), 8 % w/v SDS, 0.2 % w/v bromophenol blue, 40 % v/v glycerol, 20 % v/v  $\beta$ -mercaptoethanol) and incubated for 5 min at 95 °C shaking. Roughly 3  $\mu$ g sample of the fractions and 2.5  $\mu$ g, as well as 5  $\mu$ g of the positive control, was loaded and electrophorized with a standard of 5  $\mu$ L of Precision Plus Protein™ All Blue Prestained Protein Standards (BioRad #1610373). For electrophoresis, a 1x electrophoresis buffer freshly prepared using a 10x stock solution (30.3 g Tris-base, 144 g Glycine in 1 L of deionized water) was used, and the gels run at an electrical field of 100 V for ~2 h. The gels were then transferred onto a nitrocellulose membrane using the Trans-Blot® Turbo™ Transfer System of BioRad (**Figure 12**). A pre-set protocol of 25 V (1 A) applied field for 30 min was used. After blotting, the membranes were blocked using 5 % (w/v) skimmed milk powder in TBST for 1 h at 4 °C. The primary antibody (rabbit anti-ScV-L-A peptide serum by GenScript) was diluted at 2:25,000 in 2 % (w/v) skimmed milk powder in TBST and incubated for 16 h at 4 °C. After incubation, the membrane was washed 3 times for 10 min with 2 % (w/v) skimmed milk powder in TBST. The secondary antibody (goat anti-rabbit IgG H&L (HRP) ab205718 by Abcam) was prepared the same way as the primary with a final dilution of 1:25,000. The membranes were incubated for 1 h at RT. After another 3 washing steps with 2 % (w/v) skimmed milk powder in TBST, the membrane could be imaged using the ChemiDoc MP Imaging system and a freshly prepared ECL fluorescent mixture. Antibodies were custom-made by GenScript (New Jersey, USA).

Table 5. SDS-PAGE gel ingredients

Stacking gel	Reagent	Separating gel
3.34 mL	Acrylamide/Bis solution, 37.5:1 (Stock 30 % w/v)	0.85 mL
2.46 mL	Tris-HCl pH 8.8 (Stock 1.5M)	-
-	Tris-HCl pH 6.8 (0.5 M)	1.25 mL
50 $\mu$ L	Sodium dodecyl sulfat (SDS) (Stock 20%)	25 $\mu$ L
40 $\mu$ L	APS (Stock 10%)	20 $\mu$ L
20 $\mu$ L	TEMED (Stock 6.71M)	10 $\mu$ L
4.1 mL	ddH <sub>2</sub> O	2.86 mL



**Figure 12. Schematic of the Western Blot set up.**

The cassette of the BioRad Western Blot machine is prepared by laying a stack of 3 buffer-drenched Whatman papers into it. The air is expelled from the cassette. The membrane is then placed on top of the papers, ensuring there are no air bubbles. Following the same procedure, the gel is added and covered with an additional 3-layer stack of buffer-drenched Whatman paper. The air is expelled again to ensure that neither the membrane nor the gel are displaced from the set-up, as it could result in the faulty blotting of proteins.

#### 3.2.11.1 Chemiluminescence reaction

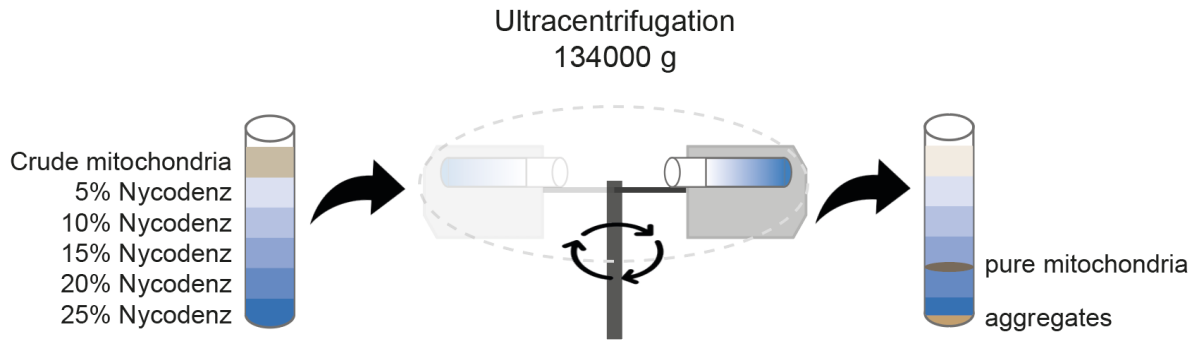
The membranes are developed in a dark environment. To prepare for each membrane, a 150  $\mu$ L ECL solution is created. This is achieved by mixing solutions A and B from the manufacturer (BIORAD) in equal proportions. The membranes are carefully placed on plastic wrap, ensuring there are no wrinkles or trapped air. The ECL solution is then sprinkled over the membranes, and they are transferred to the ChemiDoc system for evaluation. Afterward, the antibodies on the membrane are stripped for a duration of 30 minutes up to overnight, allowing them to be reused for future immuno detections.

#### 3.2.12 High-pressure freezing

Yeast cells were rapidly frozen with a high-pressure freeze fixation apparatus (HPM 010; BALTEC, Balzers, Liechtenstein). The material was cryo substituted with 0.25 % glutaraldehyde (Sigma) and 0.1 % uranyl acetate (Chemapol, Prague, Czech Republic) in acetone for two days at -80 °C using cryo substitution equipment (FSU; BAL-TEC) and embedded in HM20 (Polysciences Europe) at -20 °C. After polymerization, samples were cut with an ultramicrotome (Ultracut S, Leica, Wetzlar, Germany). The ultrathin sections (30 nm) were transferred to formvar-coated copper grids and poststained with uranyl acetate and lead citrate in an EM-Stain apparatus (Leica) and subsequently observed with a Zeiss EM 900 transmission electron microscope (Carl Zeiss Microscopy GmbH, Jena, Germany) operating at 80 kV.

### 3.2.13 Mitochondria isolation and subfractionation

Yeast is grown in liquid culture to logarithmic phase ( $OD_{600} = \sim 5$ ) and then harvested at 3,000 g at 4 °C and washed with distilled water. An aliquot of 50  $\mu$ L is taken as a control for immunoblotting (Western Blot, WB). Around 30 g yeast pellet is needed for mitochondria isolation. The pellet is resuspended in 2 mL·g<sup>-1</sup> wet weight suspension buffer (Tris-SO<sub>4</sub> 0.1 M, DTT 10 mM, pH 9.4) and incubated for 10 min at 30 °C while shaking (Sample YC). Two washing steps then follow at 3,000 g for 5 min with washing buffer (1.2 M Sorbitol, 20 mM potassium phosphate buffer pH 7.4), and after the pellet is resuspended in 6 mL·g<sup>-1</sup> wet weight washing buffer, with 3 mg·mL<sup>-1</sup>, zymolyase added. The cell wall is partially digested while incubating for 30 min whilst shaking at 70 rpm and 30 °C. The zymolyase treated sample is then washed two times with washing buffer for 5 min at 1,500 g at room temperature, and an aliquot of 50  $\mu$ L will be taken for WB (ZW). The pellet is then resuspended in 6 mL·g<sup>-1</sup> wet weight ice-cold homogenization buffer (0.6 M mannitol, 10 mM Tris-HCl pH 7.4, 10 mM DTT and 100 mM pefabloc) and homogenized with 15 strikes in ice-cold douncer (VOS 40 digital VWR). The homogenate is pelleted at 1,500 g for 5 min at 2 °C, the supernatant with the crude mitochondria saved, and the resulting pellet is re-homogenized with 1 volume homogenization buffer. After washing the re-homogenized sample, the two supernatants are combined, and an aliquot of 50  $\mu$ L is taken for WB as well as an aliquot from the discarded pellet (WB samples SNZ & PZ). The crude mitochondria are centrifuged at 10,000 g in a Beckman tabletop ultracentrifuge Max4000 for 10 min at 2 °C (save aliquot of supernatant for WB; SNU). The resulting pellet (WB sample PU) is resuspended in 4 mL resuspension buffer (0.6 M Mannitol, 10 mM Tris-HCl pH 7.4), and a gradient is prepared. For the gradient, Nycodenz is diluted in gradient buffer (10 mM MOPS-KOH pH 7.2, 1 mM EDTA, 1.2 M Sorbit) to reach concentrations of 5 %, 10 %, 15 %, 20 %, and 25 % for the layers of the gradient. The gradient is stacked as shown in **Figure 13** and then centrifuged for 1 h in a swinging bucket rotor at 134,000 g and 4 °C.



**Figure 13. Mitochondria isolation via Nycodenz gradient.**

Nycodenz is diluted to 5-25% in steps of 5% in gradient buffer and stacked from highest concentration first to lowest concentration last into the centrifuge tube. Mitochondria will travel until the interface formed between 15% and 20% Nycodenz gradient buffer, while aggregates will sediment.

The intact mitochondria are removed and washed at 10,000 g for 15 min at 2 °C with wash buffer twice. 50 µL aliquots are taken from the mitochondria band (ML), top layer opaque debris (GU) and aggregates band (GS) for WB. The mitochondria pellet is resuspended in one volume resuspension buffer (10 mM Tris-HCl pH 7.4, 0.6 M Mannitol) and stored at -80 °C.

The mitochondria pellets (800 mg wet weight) are diluted in 500 µL hypotonic buffer (20 mM MOPS, 0.5 mM EDTA, protease inhibitors: 2.5 mM pefabloc, 40 µM E-64, 130 µM bestatin, 0.5 µM aprotinin, 1 µM leupeptin, 60 µM pepstatin A, 1 mM DTT, 10 µg·mL<sup>-1</sup> DNase) and mixed extensively by pipetting. Then 3 volumes of hypotonic buffer are added and a 50 µL sample is taken for PDHc activity assays (mito). The sample is kept on ice for 1 h to swell the mitochondria and then dounced with 50 strokes. The homogenate is centrifuged at 21,000 g and the pellet is dounced again. After douncing the homogenate was centrifuged at 90,000 g for 30 minutes at 4 °C. A sample each of 50 µL pellet (pellet/PCC<sub>pellet</sub>) and supernatant (mito lysate/PCC<sub>lysate</sub>) was saved for activity assays. The supernatant was then concentrated by spin filtration (100 kDa cut-off amicon ultra centrifugal filter) at 4 °C and 4,000 g. Also 50 µL flow through (flow through/PCC<sub>flow-through</sub>) and concentrated sample (lysate conc./PCC<sub>conc. lysate</sub>) were saved for PDHc activity assays. This protocol was performed with and without 0.1% PCC added to the homogenization buffer. The resulting concentrated lysate was subjected to size exclusion chromatography in the same manner as the whole yeast lysate and every 3 resulting fractions pooled and treated for cryo-EM.

### 3.2.14 MitoTracker Orange staining

MitoTracker Orange (Invitrogen M-7511; molecular weight: 392,9273 g·mol<sup>-1</sup>) stains mitochondria in live cells and is accumulated upon membrane potential. The protocol for staining is as follows: Dissolve the lyophilized solid in DMSO to get a stock solution of 1 mM (0.05 mg in 127 µl). From this stock, which can be stored frozen, a working solution (1 mM) is prepared while mixing 1 µL stock solution with 99 µL DMSO. The working solution is then used at a concentration of 50 nM (1 µL dye in 50 µL mitochondria sample) in the isolated mitochondria sample and 200 nM (5 µL dye in 100 µL cell sample) in the yeast cell controls. The samples were then incubated for 45-50 min in the dark and 30 µL from each is transferred to a glass slide with a cover slip and imaged at the LSM (excitation 554 nm / emission 576 nm). Images were acquired with a Zeiss LSM880 Airyscan system A with the 40X objective without immersion. Images were captured with the ZEN Black image analysis software. Analysis of images was performed using the Fiji software<sup>216</sup>.

## 4. Results

### 4.1 Recovery and Isolation of yeast mitochondria

The method for mitochondria isolation via Nycodenz density gradient centrifugation was adapted from Gregg *et. al*<sup>217</sup> for the yeast strain of the laboratory. As part of the results, the careful stacking of gradients without disturbing the layers was crucial. Two methods were employed for this purpose. The first method involved starting with the least dense gradient buffer concentration and gently underlaying the higher concentrated buffers using a syringe inserted through the previous layers. The second method consisted of slowly overlaying the less dense buffer solutions on top of the higher concentrated one using a syringe, transfer pipette, or micro pipette. Each method has its own advantages and disadvantages. For the purposes of this thesis, the second method was utilized based on optimization. To ensure optimal results, the gradients were stacked cleanly, as depicted in **Figure 14**. While it is possible to freeze and store these gradients at -20°C, it should be noted that fresh gradients were prepared for all experiments conducted in this thesis. Both sucrose density gradients and Nycodenz gradients were employed in the experiments, and clear phases were observed before and after centrifugation using both types of gradients. However, considering that Nycodenz is not osmotically active like sucrose and is therefore less stressful for the mitochondria, it was determined to be the preferred gradient for retrieving mitochondria.

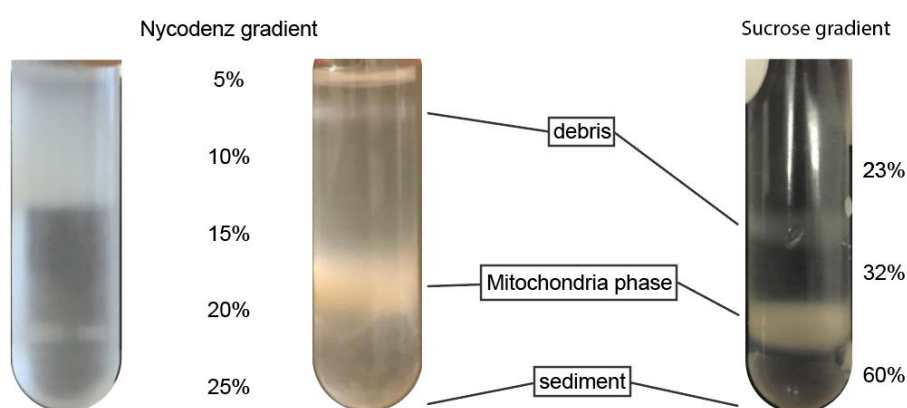


Figure 14. Density gradient before and after centrifugation.

Before and after images of Nycodenz and sucrose density gradients are shown in this image. The clean mitochondria can be retrieved after ultracentrifugation at approximately 16% Nycodenz concentration from the gradient. An aliquot each for WB analysis was taken from the debris (GU), mitochondria layer (ML) and sediment (GS).

#### 4.1.1 Mitochondrial lysate – validation of a clean preparation

To assess the effectiveness of mitochondrial purification and the vitality of the isolated organelles, several probes and analyses were performed. Firstly, the samples were probed using a mitochondrial stain called MitoTracker™. Positive controls consisted of yeast cells with active mitochondria (**Figure 15, a and d**), while negative controls included heat shock-killed yeast cells (**Figure 15, b and e**). These controls were used to establish the staining patterns before examining the isolated mitochondria (**Figure 15, c and f**).

In the positive control images, the alive cells (**a and d**) exhibited active mitochondria as small, distinct dots throughout the cells. In contrast, the dead cells (**b and e**) showed a diffuse signal, lacking any specific cellular pattern, regardless of whether they were in focus or not.

The isolated mitochondria were then stained, revealing small dots evenly distributed without any cellular structures surrounding them. Additionally, a quality control step was conducted using Western Blot analysis (**Figure 15, g**). Antibodies were employed not only on the resulting mitochondrial fraction but also throughout all isolation steps to identify potential losses of mitochondria during isolation and removal of unwanted organelles.

The VDAC1 antibody was used to detect mitochondria. This antibody specifically identifies the Voltage-Dependent Anion-Selective Channel 1 present in the outer mitochondrial membrane and is commercially available. The antibody was applied to both supernatants and pellets at various stages, such as during washing (**Figure 15, h ZW**) and centrifugation to remove cell wall debris (**Figure 15, h ZW and PZ**). If mitochondria were lost at any point, it would result in decreased yield in subsequent steps, indicating a potentially harsh protocol.

The applied protocol exhibited a clear band for mitochondria in the isolation steps where further processing occurred, as well as a strong signal for mitochondria in the gradient band that was isolated and stored for subsequent experiments. The antibody also revealed loss of mitochondria in the upper gradient layer and gradient sediment, which was expected due to the presence of broken mitochondrial membrane parts and aggregates in these layers.

The interaction between the endoplasmic reticulum (ER) and actin with mitochondria was of interest, necessitating their probing. Both proteins showed signals in the supernatant after ultracentrifugation, but lower signals were observed in the crude mitochondrial pellet. However, in the upper gradient layer, a strong signal was detected for both actin and the ER, while no signal was detected in the mitochondrial layer, confirming the absence of these proteins.

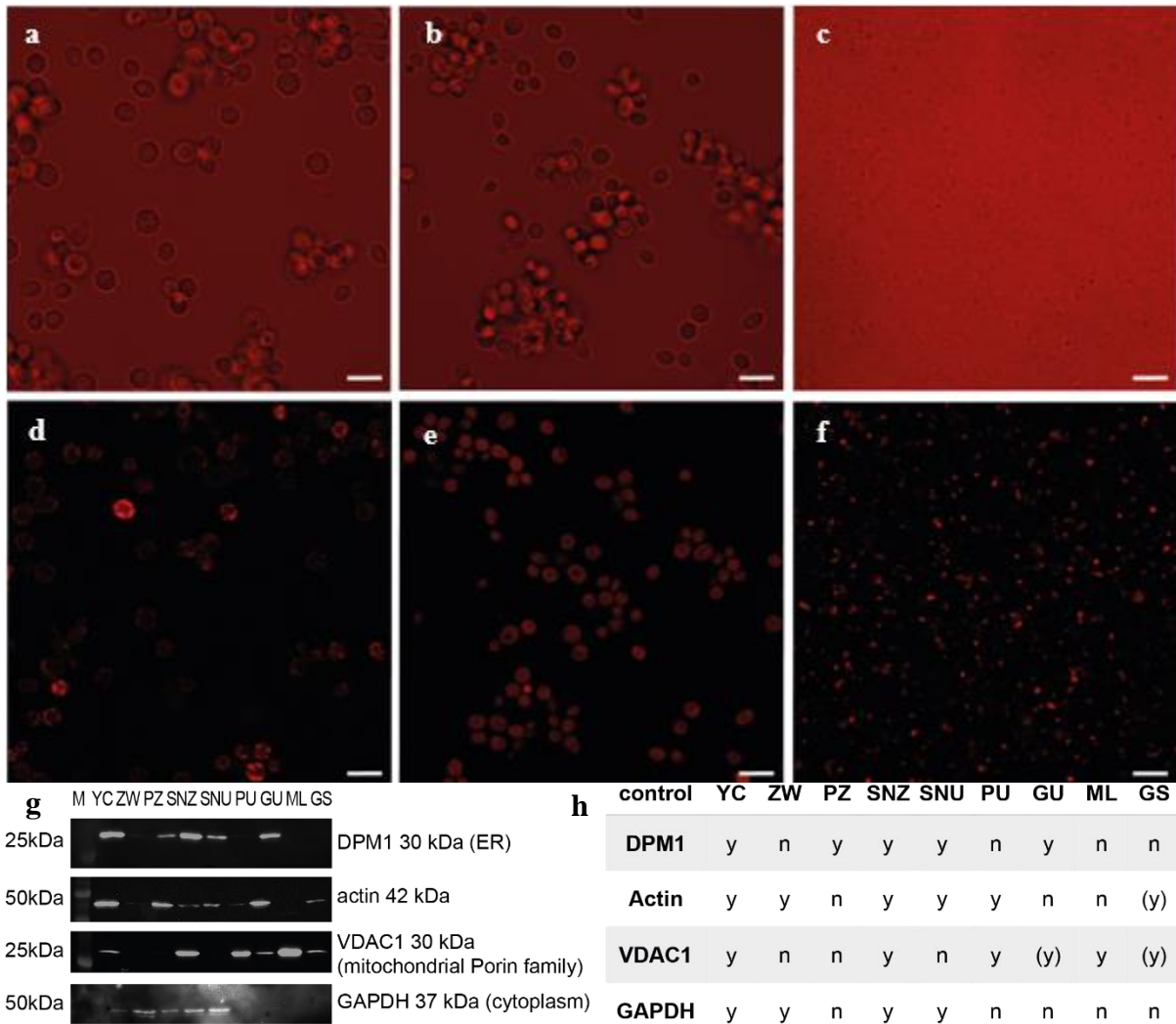


Figure 15. Data supporting a clean mitochondria preparation.

The yeast cells (a-b and d-e) and isolated mitochondria (c and f) were tested for their vitality with the MitoTracker stain. As positive control live yeast (a light microscopic image, d fluorescent image) was probed, as negative heat shocked yeast (b light microscopy, e fluorescence) was also stained after ensuring that yeast cells were still visible but wouldn't grow on solid media. The isolated mitochondria (c light microscopy and f fluorescence) were stained and checked for vitality. Mitochondria isolation steps were probed via Western blot analysis. Shown in this figure are the results for anti-ER antibody DPM1, anti-actin antibody, anti-cytoplasm antibody GAPH and antibody against the mitochondrial porin family VDAC1. The abbreviations are as follows: M Marker, YC yeast cells, ZW washing step after Zymolyase, treatment SNZ supernatant after light centrifugation step, PZ pellet after light centrifugation step SNU supernatant after strong centrifugation step, PU pellet after strong centrifugation step GU gradient upper layer ML mitochondrial layer GS gradient sediment. h) Expected signal for western blot control. In the matrix y stands for yes, a signal for this antibody is expected and n for no, a signal is not expected, whereas (y) can have faint signal due to broken mitochondria. Scale bars 5  $\mu$ m.

To assess the presence of cytoplasmic content, the GAPDH antibody was employed. This antibody detects the enzyme Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH)



involved in cytoplasmic glycolysis. Immunoblotting analysis demonstrated a strong cytoplasmic signal until the supernatant, but no signal was detected for the crude mitochondria and gradient centrifugation steps, indicating the successful removal of cytoplasmic content during the process. Overall, the combination of staining, Western Blot analysis, and antibody probing provided comprehensive insights into the purification efficacy, vitality, and absence of unwanted proteins during the isolation of mitochondria.

#### 4.1.2 Mitochondria lysate – observation with cryo-EM

ATP synthase dimers of the inner mitochondrial membranes, the protein complex that generates most of the ATP, were retrieved by Blum *et al.* from *Yarrowia lipolytica* and visualized by cryo-electron tomography<sup>218</sup> (**Figure 16, a**). In the mitochondrial fractions stemming from yeast vesicles with highly similar decoration of membrane proteins could be identified with cryo-EM (**Figure 16, b-c**). These ATP synthase vesicles differ in size and shapes, mostly resembling vesicles or flattened discs with perturbing small spikes. As these structures were not structurally analyzed further, the observation was an additional validation for the successful mitochondrial preparation, amenable for cryo-EM investigations.

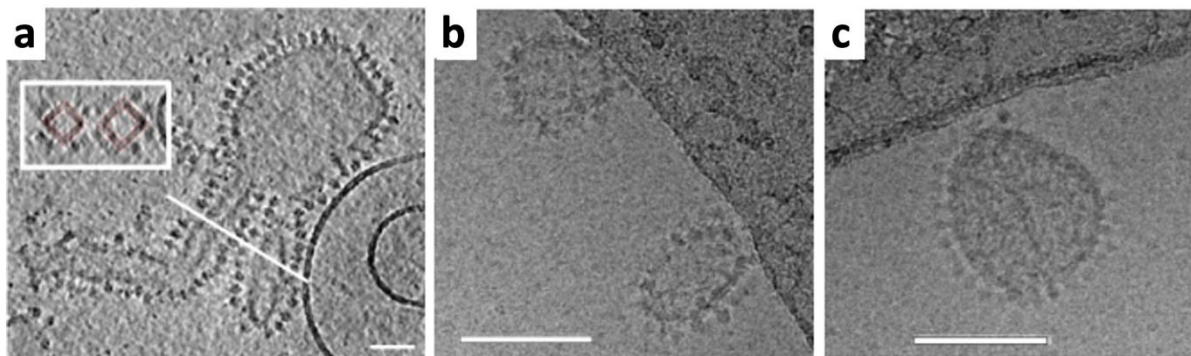


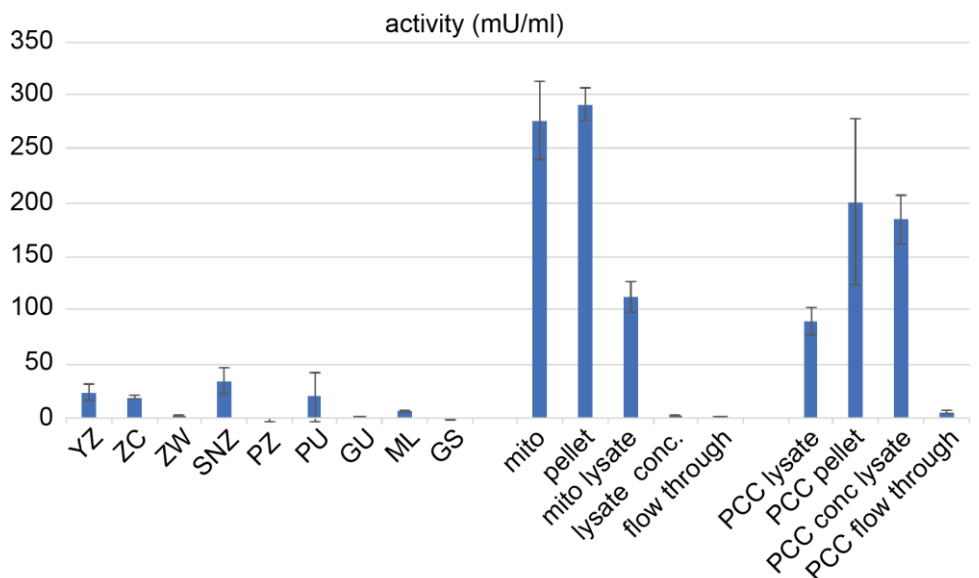
Figure 16. Discovery of mitochondrial ATP synthase dimers after mitochondria isolation and lysis.

Proteoliposomes of reconstituted *Y. lipolytica* ATP synthase dimers from Blum *et al.*<sup>218</sup>(a; scale bar 100 nm). Similar shaped features in mitochondrial fractions after mitochondria isolation and lysis of *S. cerevisiae* mitochondria (b-c; scale bars 100 nm).

#### 4.1.3 Retrieving the pyruvate dehydrogenase complex from mitochondria

Mitochondria were isolated and ensured to be clean and ready for use through the application of a vitality dye, western blot, and cryo-EM analysis (see above). Then, the ability of the mitochondria to oxidize pyruvate, which is carried out by the mitochondrial pyruvate dehydrogenase complex (PDHc), was examined. PDHc activity was expected to be observed in both yeast cells before and after zymolyase treatment. During the washing steps (ZW), it was discovered that there was minimal activity, indicating that the washing process was

gentle and did not result in significant loss or breakage of mitochondria (**Figure 17**). Subsequently, the washed and concentrated mitochondria were subjected to activity testing before being lysed. They exhibited high activity. Following lysis via osmotic shock and douncing, both the lysate and the remaining solid portion (pellet) continued to display PDHc activity, although the activity in the pellet was substantially higher. This mirrored the observations made in the lysate experiment, where PDHc was also found in the pellet. No PDHc activity could be detected whatsoever after concentrating the mitochondrial lysate. However, when a detergent called PCC (pyridinium chlorochromate) was utilized, a distinction in the activities of the pellet and the concentrated lysate samples became apparent. The concentrated lysate showed slightly less activity compared to the pellet from the experiment with PCC. This indicated that a significant amount of PDHc was bound to the membranes. Nonetheless, a substantial quantity of PDHc was still recoverable for subsequent analysis. Overall, these findings indicate that the washing and isolation process effectively obtained clean mitochondria with functional PDHc activity and provided evidence that yeast PDHc has affinity for the mitochondrial membranes.



**Figure 17.** PDHc activity assay of mitochondrial isolation steps to monitor PDHc during mitochondria isolation and lysis.

ZW washing step after zymolyase treatment SNZ supernatant after light centrifugation step, PZ pellet after light centrifugation step SNU supernatant after strong centrifugation step, PU pellet after strong centrifugation step GU gradient upper layer ML mitochondrial layer GS gradient sediment. Mitochondria lysed in osmotic active buffer without and with detergent PCC. Lysis steps involve intact pooled mitochondria (mito), pellet after osmotic shock and homogenizing (pellet and PCC pellet) and lysate (mito lysate and PCC lysate), concentrated lysate (lysate conc. and PCC lysate conc.) and flow through (flow through and PCC flow through). For visualization of the purification steps, please refer to **Figure 10, a** and its legend.

The mitochondrial lysate obtained from PCC including buffer lysis was concentrated to  $\sim 30 \text{ mg}\cdot\text{mL}^{-1}$  and  $100\mu\text{L}$  was subjected to size exclusion chromatography as done with the native cell extracts of *C. thermophilum*<sup>50-53,69</sup> and yeast (see next chapter). The chromatogram (**Figure 18**) has no peak at the higher molecular weight fractions but shows a peak at the fractions with molecular weight  $<100\text{kDa}$ . The sample was concentrated prior to SEC in an amicon with a cut-off of  $100 \text{ kDa}$  and peaks for proteins above that were expected. The chromatogram shows one large peak at low molecular weight fractions suggesting that the proteins were dissociated in any step between injection and elution.

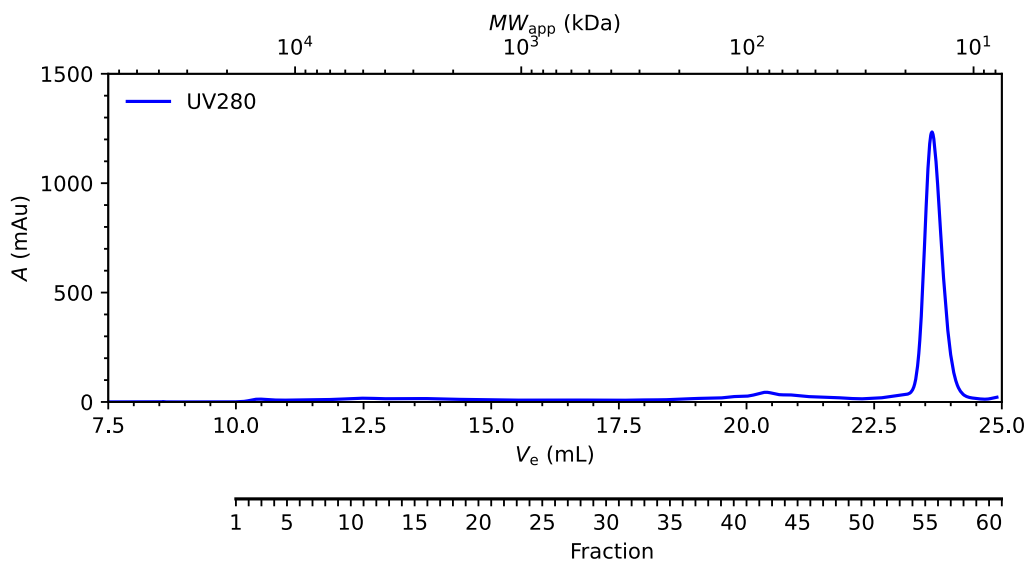


Figure 18. Size exclusion chromatography of mitochondrial lysate from the lysed with PCC including buffer shows dissociated protein content.

$100 \mu\text{l}$  ( $30 \text{ mg}\cdot\text{mL}^{-1}$ ) mitochondrial lysate obtained from lysis with 0.1% PCC on Sepharose Q column.

4.2 Biochemical and structural analysis of yeast cell extract by single-particle cryo-EM  
 Size exclusion chromatography (SEC) was employed to separate complexes of similar molecular weight from the native cell extract, with the resulting fractions being collected within an efficient timeframe of less than 8 hours after yeast cell harvesting (**Figure 19, a-b, Methods**). The approach utilized in this study was based on fractionation experiments performed on the thermophilic mold *Chaetomium thermophilum*, with minor modifications made to adjust the initial cell volume and the process of resuspending the starting material prior to lysis (Methods). This allowed the retrieval of fractions containing significantly concentrated proteins suitable for direct cryo-EM analysis (**Figure 19, c**). Following the SEC step, fractions 5 and 6, which encompassed megadalton (MDa) assemblies, were vitrified (**Figure 19, a**, indicated by orange dotted line).

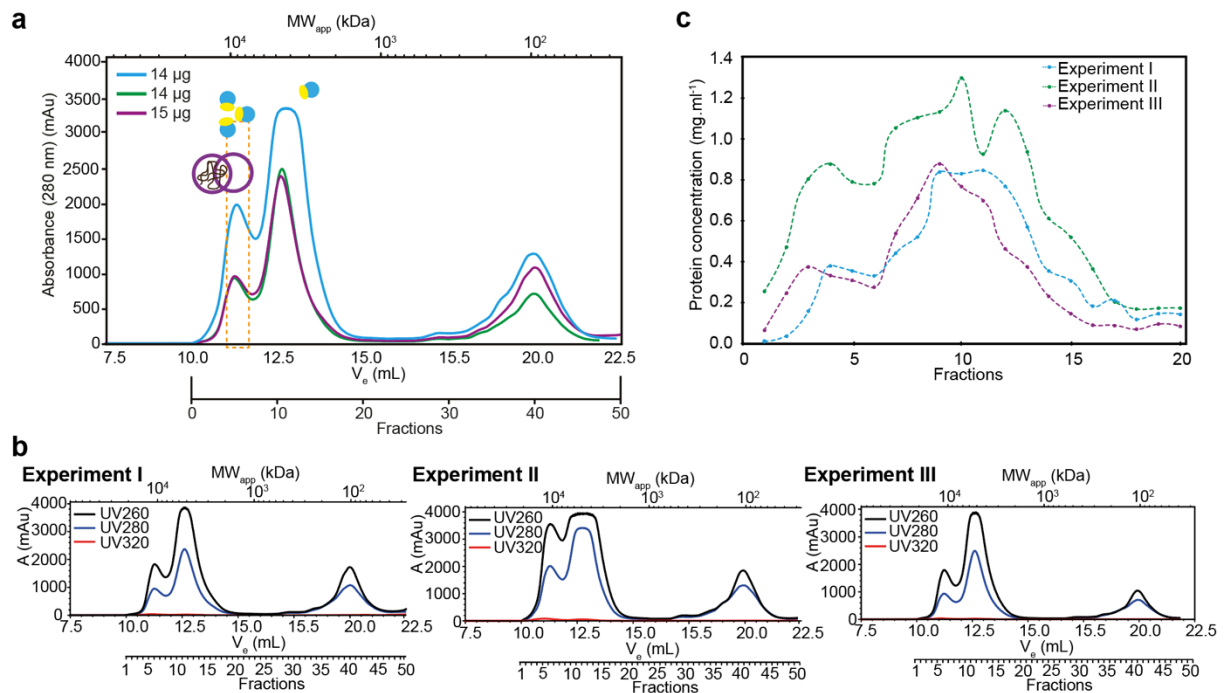


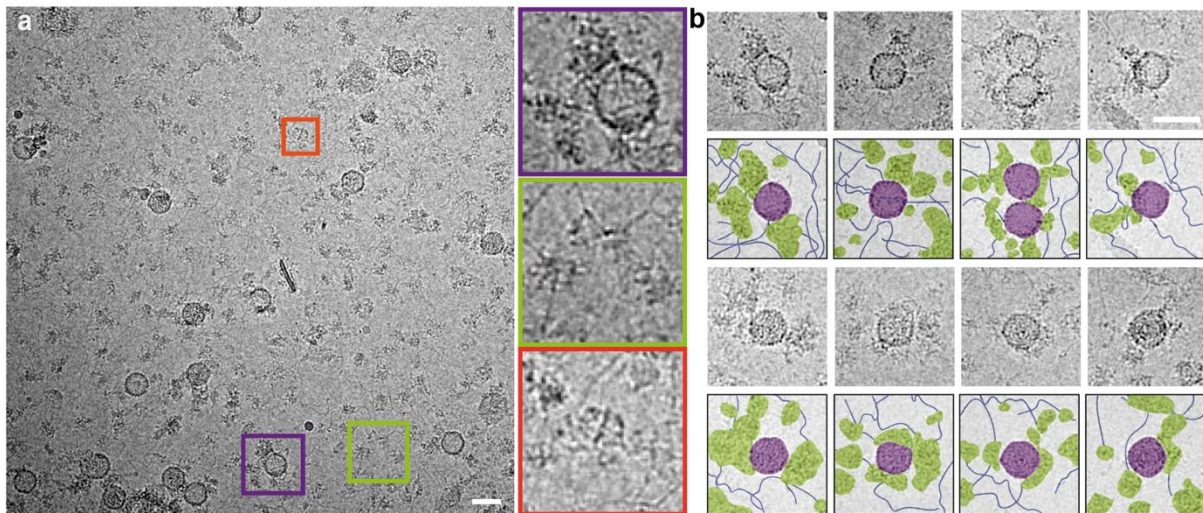
Figure 19. Size exclusion chromatography results of yeast cell extract.

(a) In this study, the *S. cerevisiae* native cell extract was subjected to SEC profiling. Fractions 5 and 6, which corresponded to  $\sim 10$  MDa complexes, were further investigated (orange box). The locations where full and empty L-A virus are expected based on relative retention are indicated by purple circles. Consequently, the first peak is anticipated to contain polysomes, and the second peak is expected to contain monosomes. Triplicate SEC measurements were conducted at 260 nm, 280 nm, and 320 nm absorbance (b). Additionally, the resulting fractions after SEC were analyzed using the Bradford reagent, and high in-fraction protein concentration was observed in the triplicate experiment (c). Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

The cryo-EM micrographs obtained revealed the presence of large intact complexes. Notably, large biomacromolecules, including fatty acid synthase (FAS)<sup>50</sup>, were observed (**Figure 20, a**). The existence of such intact complexes indicates the preservation of their structural integrity during the sample preparation and imaging process, enabling a detailed examination of their architecture. Furthermore, the micrographs also exhibited the presence of polysomes (similar to the ones we previously observed in a cleaner sample<sup>220</sup>), which are clusters of ribosomes actively engaged in protein synthesis (**Figure 20, b**). These polysomes signify the cellular machinery's active involvement in protein production. The identification of polysomes within the cryo-EM micrographs provides crucial evidence of the ongoing translation process within the cellular context under investigation.

In addition to the larger complexes and polysomes, distinct spherical structures were also prominently observed within the cryo-EM micrographs (**Figure 20, a-b**). These spherical structures, with their well-defined boundaries, represent discrete entities within the cellular

landscape. Their presence suggests the existence of unique cellular components or compartments, the nature of which is yet to be fully understood. Importantly, during the meticulous analysis of the micrographs, a consistent observation emerged regarding the close proximity of these spherical structures to filamentous formations and potential polysomes (**Figure 20, a-b**). This intriguing spatial relationship implies potential functional associations or interactions between the spherical structures and the surrounding cellular components. The proximity of the spherical structures to filamentous formations and polysomes may indicate their involvement in important cellular processes, such as organelle organization, molecular transport, or signal transduction. Taken together, the cryo-EM micrographs not only unveiled the presence of large intact complexes like FAS<sup>140</sup> and polysomes but also shed light on the existence of distinct spherical structures within the cellular milieu. The spatial relationship observed between these structures and other cellular components hints at their potential functional significance, stimulating further investigation into their roles and interactions within the intricate cellular machinery.

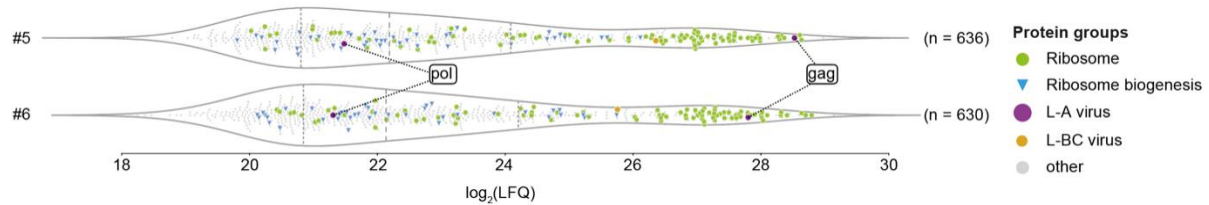


**Figure 20. Exploration of large molecular weight fraction using cryo-EM.**

(a) A micrograph representing fraction 5 displays a diverse array of biomolecules, showcasing their heterogeneity within the sample. Notably, the presence of key components such as fatty acid synthase (highlighted in the red box), ribosomes (highlighted in the green box), and the L-A virus (highlighted in the purple box) can be observed. Additionally, in the micrograph crops presented in (b), the L-A virus is depicted near ribosomes, fibrillar structures, and other proteins. These representative crops provide a glimpse of the spatial relationship between the L-A virus and various cellular components, emphasizing its potential interactions within the complex cellular environment. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup> Scale bars 50 nm.

4.2.1 The yeast L-A virus is identified in a cryo-EM-accessible eukaryotic cell extract. A total of 7011 movies were collected, each captured at a pixel size of 1.57 Å (**Methods**). These movies were then subjected to image analysis, specifically focusing on the spherical structures observed (**Supplement Figure 1**). The analysis yielded an averaged cryo-electron microscopy (cryo-EM) map at a resolution of 3.77 Å, achieved through icosahedral averaging (FSC=0.143, **Supplement Table 1**). Furthermore, based on this map, an initial backbone trace was feasible to be constructed due to the achieved resolution. The resolution, indeed, could accommodate the majority of the backbone to be modelled. In order to find structurally related molecules, a search was conducted using DALI search<sup>211</sup> against structures in the Protein Data Bank (PDB), a software that uses distance-based matrix alignments<sup>221</sup>. The search identified a potential match with the 3.4 Å crystal structure of the *S. cerevisiae* L-A virus<sup>183</sup>, which happened to be the only available structure. To verify the presence of L-A virus specific proteins, mass spectrometry (MS)-based protein identification was performed on the fractionated cell extract. The analysis confirmed the abundance of the L-A virus proteins (Gag, Gag/Pol), as well as the presence of the L-BC virus, albeit in much lower abundance. Notably, structures of L-A and L-BC viruses are very different<sup>192</sup>, especially in the high-resolution regime that was achieved. To enhance the accuracy of the identification,

the endogenous protein sequences corresponding to these viruses were incorporated into the existing yeast proteome (UP000002311, <https://www.uniprot.org/proteomes/UP000002311>) (**Figure 21**).



**Figure 21.** Proteomics of fractions 5 and 6 that were also analyzed by cryo-EM.

To identify the contents of fractions 5 and 6, LFQ-based MS quantification was employed, allowing for the detection and characterization of the L-A virus within these fractions. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

An analysis was conducted to annotate the proteomic content of the studied extract, specifically focusing on the two fractions containing the L-A virus, and employing biological triplicates for reliable measurements. By utilizing the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database<sup>222</sup>, the identification and classification of proteins were streamlined. As a result, ribosomal proteins were identified with exceptionally high abundance (**Figure 21**). Furthermore, within these fractions, the capsid (gag) and polymerase (pol) of the L-A virus were also identified in significant quantities (**Figure 21**). Cross-linking verified also the presence of these molecules (**Supplement Figure 2**).

To gain further insights, the capsid:pol ratio per fraction was calculated based on the mass spectrometry (MS) data, similar to the approach used for studying the endogenous pyruvate oxidation metabolon<sup>51</sup>. The analysis revealed an average of  $1.8 \pm 0.49$  pol proteins per capsid, which aligns with previous predictions derived from frameshifting efficiency of the gag open reading frame<sup>146</sup>. In comparison to previous studies, the reported results showed the following:

- The L-A virus was enriched within rapidly retrieved yeast extracts, which were directly amenable for cryo-EM and mass spectrometry analysis.
- The "environment" and components of the L-A virus were investigated through quantitative mass spectrometry.
- The L-A virus was identified by employing the high-resolution features of the calculated cryo-EM map and utilizing C $\alpha$ -trace model building.

4.2.2. The cryo-electron microscopy (cryo-EM) structure of the L-A virus, obtained at high resolution from a cell extract, reveals intriguing structural and functional adaptations in its native form.

The derived cryo-EM map of the L-A virus, obtained at a resolution of 3.77 Å (FSC=0.143) (**Figure 22, a**), allowed for the construction of a high-quality de novo model. Through image processing (**Supplement Figure 1**), the model accurately accounted for the observed densities (**Figure 22, b**).

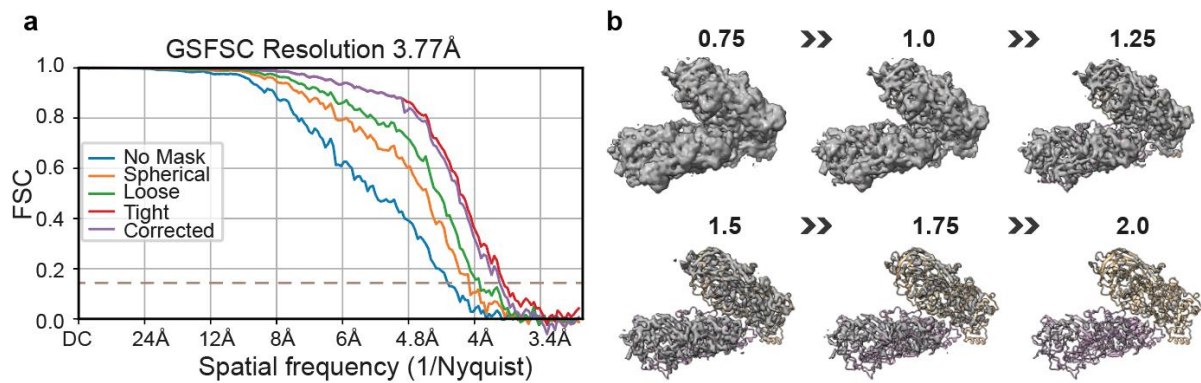


Figure 22. Fourier shell correlation (FSC) for the cryo-EM map of the L-A virus and derived model.

(a) FSC of the cryo-EM derived map for the L-A virus utilizing icosahedral symmetry. (b) Asymmetric dimer of the yeast L-A virus model superimposed on the capsid densities at different thresholds (increase in steps of 0.25 sigma) set in ChimeraX. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Fitting and refinement of the model in the cryo-EM map showed a reasonable fit, both in terms of main chain and in terms of side chain placement. Although the cryo-EM map was resolved at 3.8 Å, zooming into the densities, clear resolution of both main-chain and often side-chain details were apparent (**Figure 23, a-b**).



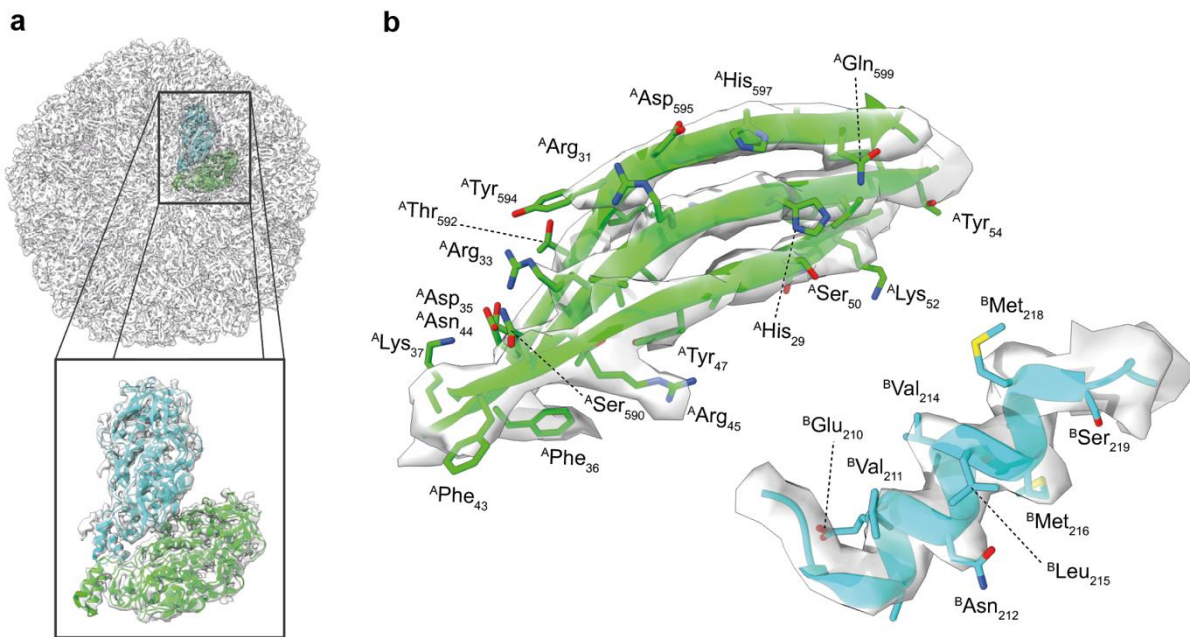


Figure 23. The cryo-EM map of the L-A virus and the derived model of the asymmetric capsomere.

(a) The cryo-EM map and model of the asymmetric subunit were derived, with the magnified fitting shown in the insert. (b) The atomic model fitted in the cryo-EM map exhibits representative resolution features. A close-up of a  $\beta$ -sheet is shown in green, where the separation of the single strands is clearly visible in the map's density. Additionally, an  $\alpha$ -helix is displayed in blue, revealing the helical pitch at a resolution of 3.78 Å, with reliable side chain density often observed. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

The icosahedral protein shell of the L-A virus exhibited a triangulation number (T) of 2, consisting of 120 protomers. The overall dimensions of the virus, such as its diameter (400 Å) and thickness (46 Å), were comparable to its crystallographic counterpart. Additionally, the opening diameter at the icosahedral fivefold axes, measuring 18 Å, served as a gateway for the viral mRNA to exit and for nucleotide triphosphates to enter. This consistent architecture of the L-A virus, maintained in both cell extracts and the crystal structure, indicates a high level of rigidity in the capsid and its openings, highlighting their critical structural stiffness.

During the L-A virus analysis, previously unreported stabilizing interactions, cation- $\pi$  interactions, were identified. These interactions play a crucial role in facilitating connections both within and between L-A protomers (**Figure 24**). Within the asymmetric unit, a total of 19 side chains participate in forming 8 distinct cation- $\pi$  interaction networks, which are formed at both monomeric and dimeric interfaces (examples shown in **Figure 24**). Notably, the residues involved in these interactions exhibit higher resolution compared to side chains of the same type that are not involved in cation- $\pi$  interactions (**Figure 25**).

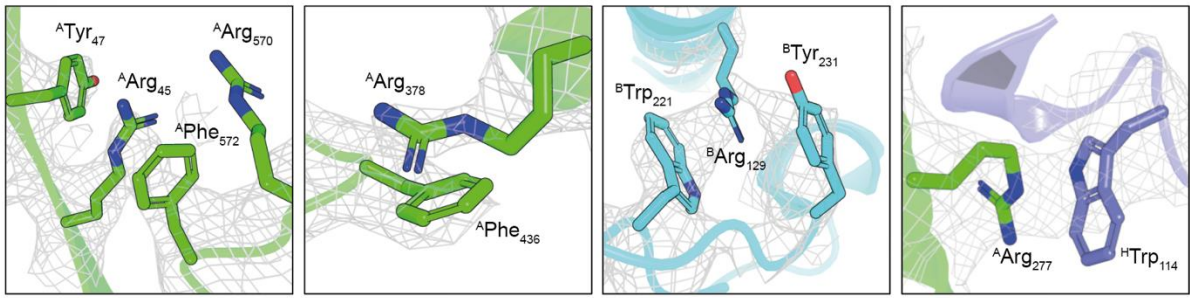


Figure 24. Discovered cation- $\pi$  interactions stabilizing the L-A virus capsid.

Representative cation- $\pi$  interactions are shown between Arg and Tyr, Trp or Phe within chain A (green), chain B (blue), or between chain A and an adjacent chain B (slate). Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

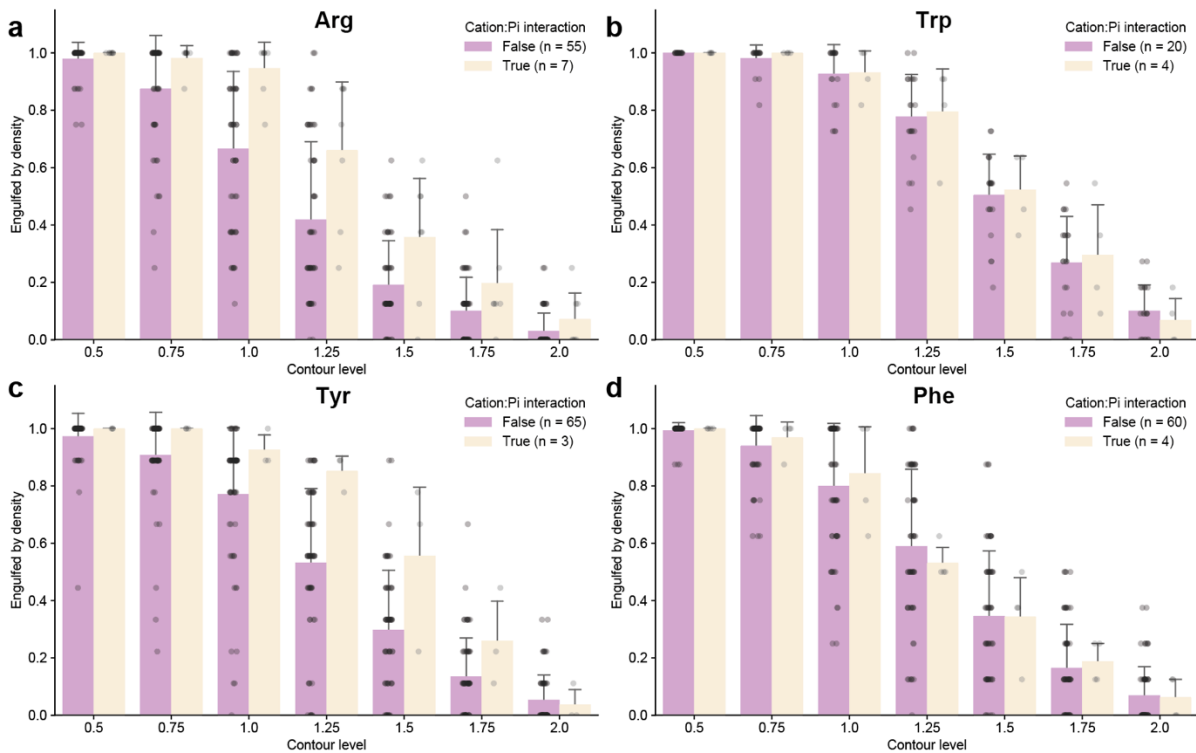


Figure 25. Persistent densities for the cation- $\pi$  interactions formed by Arg (a), Trp (b), Tyr (c) and Phe (d).

Bar plots of the resolvability of residues involved in cation- $\pi$  interactions (papaya whip) against residues that are not involved in such interactions (purple). Different contour levels were calculated in ChimeraX, and these results show the persistence of these interactions in the cryo-EM map as compared to other residues that are not involved in such interactions. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Cation- $\pi$  interactions are robust non-covalent bonds<sup>223</sup>, exemplified by their structural presence in bacteriophage T5, where a single cation- $\pi$  interaction persists independently of capsid conformational changes as revealed by cryo-EM<sup>224</sup>. Another example is the prediction of a cation- $\pi$  interaction in a dsRNA virus (bluetongue virus) following mutagenesis<sup>225</sup>. The observed extended network of cation- $\pi$  interactions is considered highly specific, as it differs from the non-specific hydrogen-bonding patterns typically observed. The formation of these interactions requires the proper alignment of a delocalized  $\pi$ -electron system with a cationic moiety in the correct orientation. In the context of the L-A virus, cation- $\pi$  interactions likely contribute to both the folding of the individual building blocks and the assembly of the virus into higher-order structures. Substantial conformational differences are observed in the crystallographically determined regions of the capsomere between the two monomers present in the resolved asymmetric unit<sup>183</sup>. These differences encompass residues Asn8–Lys12, Gly82, Thr96–Ile99, Ile111–Thr112, and Gly387–Asp396, which are also evident in the cryo-EM model. Furthermore, the cryo-EM model reveals specific adaptations that occur within the cell extract, distinct from the effects of dense crystal packing on the capsid conformation (**Figure 26, a**). The asymmetric homodimer obtained from cryo-EM exhibits superior fits in both backbone ( $CC_{\text{cryo-EM}}=0.85$  vs.  $CC_{\text{x-ray}}=0.79$ ) and side-chain ( $CC_{\text{cryo-EM}}=0.84$  vs.  $CC_{\text{x-ray}}=0.79$ ) conformations compared to the crystal structure (**Figure 26, a**). Structural changes compared to the crystal structure are detected in both capsomere subunits (**Figure 26, b**), particularly in regions Glu92–Val104, Gly387–Ser393, and Phe525–Thr535, as well as regions Ala494–Asp505 of protomers A and B, respectively. These regions play direct roles in capsid assembly (**Figure 27, a-d**).

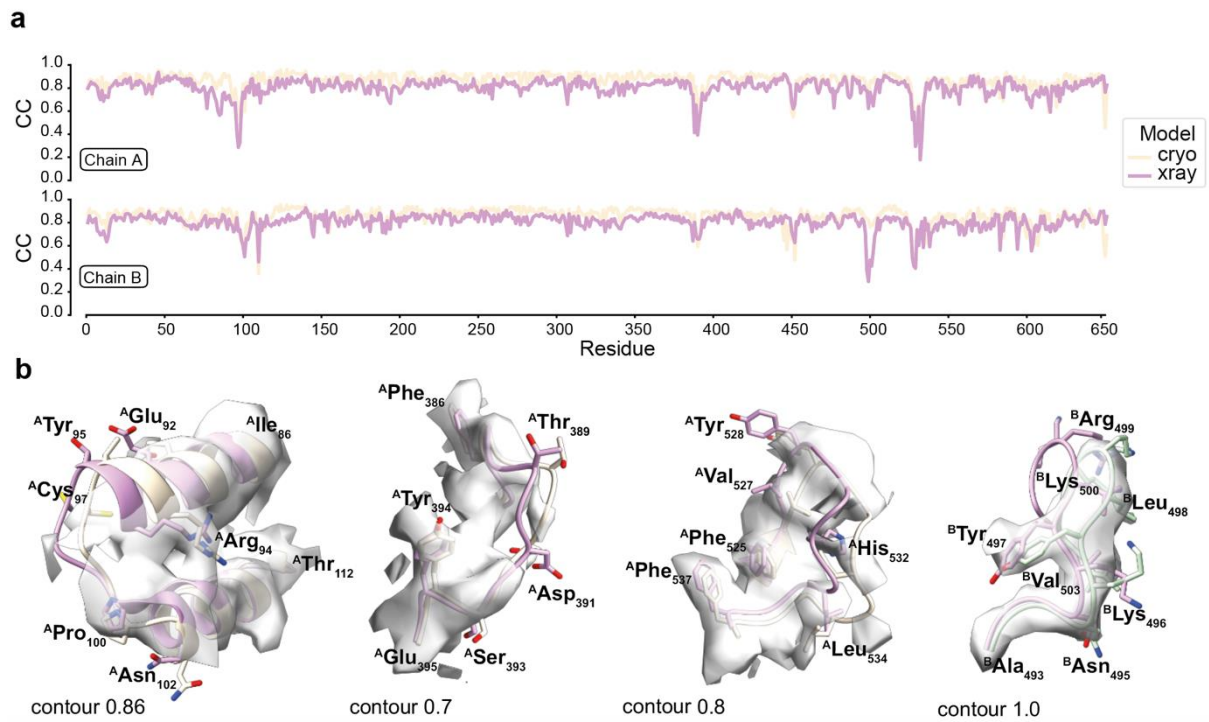


Figure 26. Fits and conformational variations of X-ray and cryo-EM models to resolved map. (a) Cross-correlation analysis of chain A and chain B models, represented in peach and purple colors respectively, reveals local conformational variability as indicated by dips in cross-correlation values. (b) The asymmetric homodimer fits derived from cryo-EM (peach) and X-ray (purple) models are shown, demonstrating backbone ( $CC_{\text{cryo-EM}}=0.85$  vs.  $CC_{\text{X-ray}}=0.79$ ) and side-chain ( $CC_{\text{cryo-EM}}=0.84$  vs.  $CC_{\text{X-ray}}=0.79$ ) conformations. Additionally, a comparison is made between cryo-EM (beige) and X-ray (purple) models after fitting into the reconstructed capsid density (grey), highlighting significant local conformational differences. Overall, these analyses provide insights into the fits and variations of X-ray and cryo-EM models with the resolved cryo-EM map, shedding light on the conformational flexibility of the studied system. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

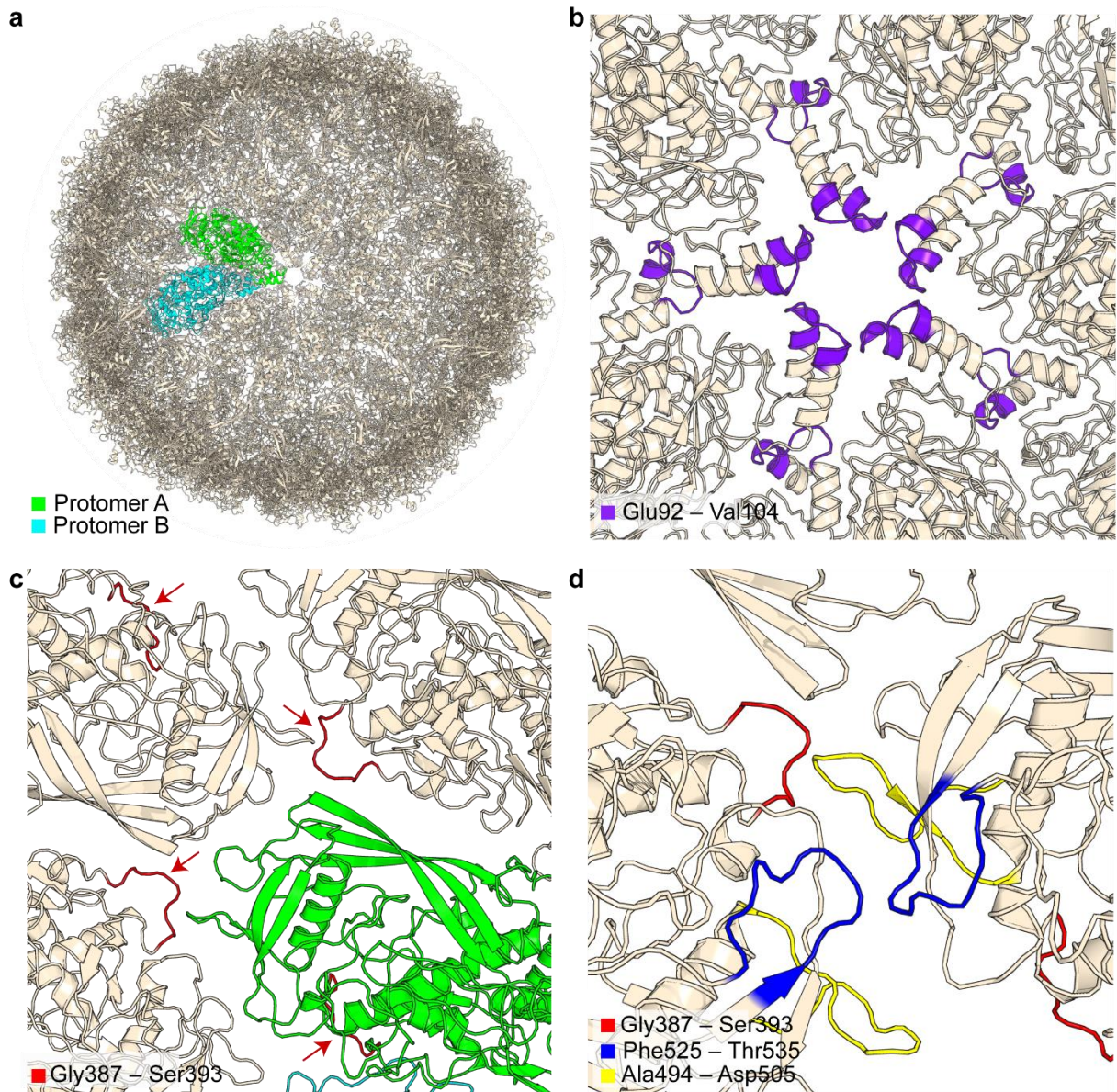


Figure 27. Structural variation in the cryo-EM structure of the yeast L-A virus.

(a) Overview of full capsid structure and the asymmetric unit (capsomere) composed of protomer A and B (b) protomer's helix-turn-helix fold at region Glu92–Val104 localizing at the five-fold axis, structuring and supporting the mRNA exit tunnel/nucleotide entry tunnel (c) loop Gly387–Ser393 is buried in the capsomeres dimeric interface and shown with red arrows (d) loop Phe525–Thr535 is involved in higher-order binding of the capsomeres communicating with other regions of identified flexibility, involving loops Ala494–Asp505 and Gly387–Ser393. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

The functional pores at the icosahedral five-fold axes are formed by the helix-turn-helix fold region E92–V104 in both protomers A and B (**Figure 27, b**, purple-blue). The cryo-EM structure exhibits flexibility in this region, which can be attributed to its dual role in supporting the capsid and forming the pores. Loop Gly387–Ser393 is buried in the dimeric interface of the capsomere (**Figure 27, c**, red), while loop Phe525–Thr535 participates in

higher-order binding through loop-loop interactions (**Figure 27, d**, blue). The flexible loop Ala494–Asp505 in capsomere protomer B interacts with loop Gly387–Ser393 and further stabilizes the capsid (**Figure 27, d**, yellow). The distinct flexibility differences described above can be quantitatively assessed by calculating energetic contributions using the macromolecular modelling software HADDOCK<sup>226</sup>. The embedded capsomere interface (**Figure 28, a**) buries a substantial surface area of 3400 Å<sup>2</sup> (**Figure 28, b**), with electrostatic forces playing a significant role (**Figure 28, b**). Extending energetic calculations to all proximal interfaces of the capsomere reveals a diverse range of energetic contributions, where van der Waals forces, electrostatics, and desolvation energies uniquely define each of the nine surrounding interfaces (**Supplement Figure 3**). Based on these unique interface energetic calculations, we propose a model of capsomere stability, involving 86 distinct steps of monomeric association to complete the capsid (**Supplement Figure 3; Supplement movie 1**).

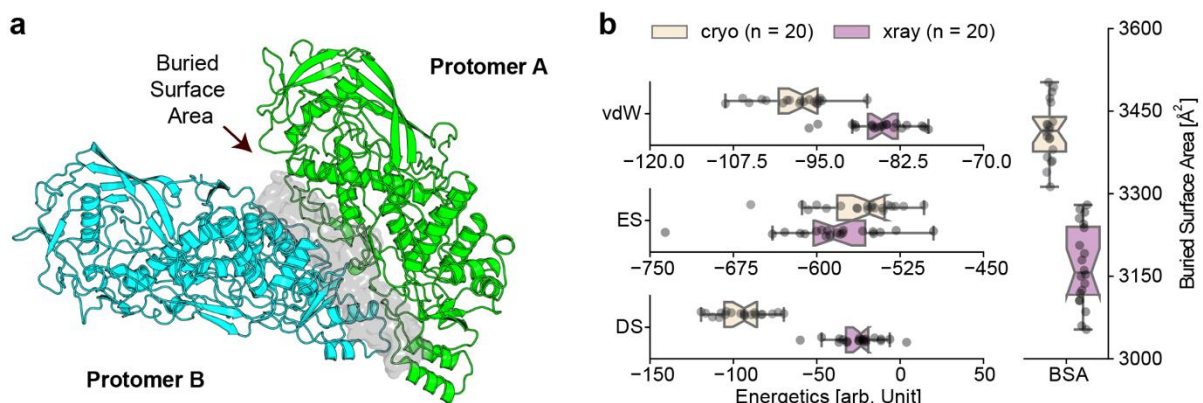


Figure 28. Energetic calculations of the capsomere interface utilizing HADDOCK.

View of the capsomere composed of A and B chains, highlighting the surface area buried, forming the capsomere interface (3400 Å<sup>2</sup>). Chain A is colored in peach, chain B in mint green, and the interface in grey. (f) Scoring components for the interface after refinement of the capsomere shown in (e) calculated with the molecular modelling software HADDOCK<sup>226</sup> (link used is: <https://alcazar.science.uu.nl/services/HADDOCK2.2/haddockserver-refinement.html>). Non-covalent interactions calculated include van der Waals (VdW), electrostatics (ES), and desolvation energy scores (DS), all in arbitrary units (a.u.), and buried surface area (BSA), in Å<sup>2</sup>. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

#### 4.2.3 Mediation of the mRNA decapping site trench shape by two flexible loops – proximity to a potential site for protein-protein interactions

The location of mRNA decapping activity, although not conserved across protomers, is widely observed among dsRNA virus families<sup>154</sup>. In both the X-ray and cryo-EM L-A virus structures, the region encompassing Gln139-Ser182 (including the active site His154) (**Figure 29, a-d**) contributes to the outer surface of the capsid, facilitates cellular mRNA decapping, and facilitates the transfer of the 7-methyl-GMP (m<sup>7</sup> GMP) cap from the cellular mRNA 5'-end to the viral RNA 5'-end, counteracting the host exoribonuclease's targeting of uncapped RNAs. This site, therefore, governs competition for the utilization of the translation machinery.

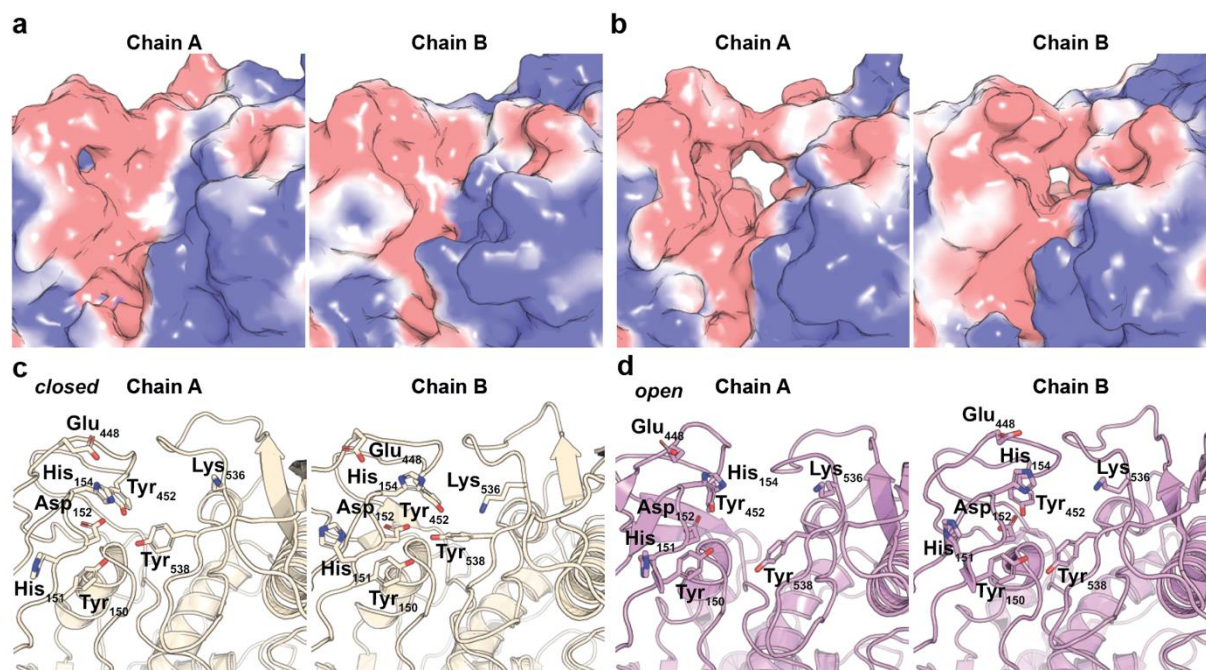
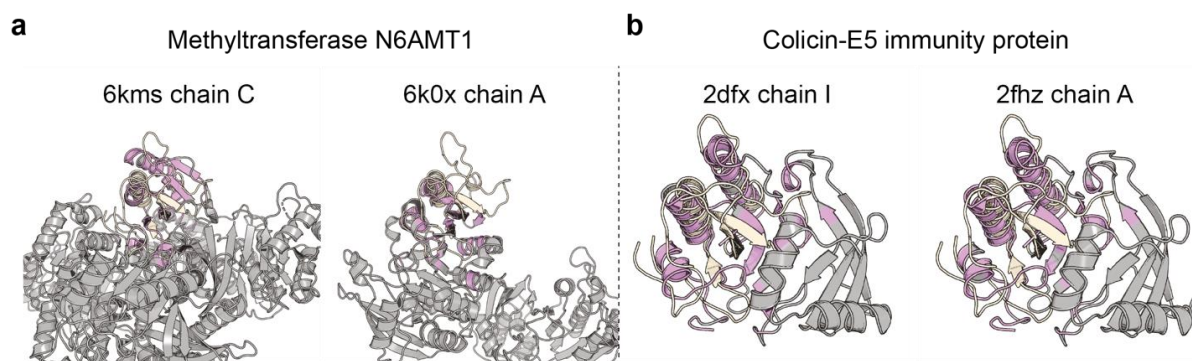


Figure 29. The shape of the mRNA decapping site trench is influenced by two flexible loops. The decapping site is presented in surface (a, b) and cartoon (c, d) representations. The conformations observed in the cryo-EM structure of the L-A virus capsid are depicted in panels (a) and (c), while panels (b) and (d) illustrate the conformations observed in the crystal structure. The capsomere subunits A and B are represented by Chain A and Chain B, respectively. The decapping center of the L-A virus capsid, as observed in the cryo-EM data, is identified to be in an occluded state. Conversely, the x-ray data shows the decapping center in an open state. The outer capsid surface, which forms a trench, is assembled by flexible loops contributed by regions Gln139-Ser182. The active residue for RNA decapping, His154, is shown in panels (c) and (d). Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

A trench, formed by loop regions and containing residues involved in the active site, was previously described<sup>26</sup>. These loops maintain similar conformations in the cryo-EM structure, resulting in a comparable local Coulomb surface potential (**Figure 29, a, b**).

However, an additional pair of loops, P449-Y452 and T531-T535, located distally from each other with the latter also contributing to the flexible region F525-T535 (**Figure 28, d**), influences the diameter of the trench pore. In the cryo-EM structure, their proximity is 3 Å closer, allowing H532 to close the trench (**Figure 29, c, d**). While the presence of the open state cannot be disregarded in the native state, the more occluded trench is prominent in the derived average. This indicates that a selectively open conformation of the active site is not universally present in all capsomeres, but logically must exist near the pore at the 5-fold axis. However, icosahedral averaging does not permit the quantification of open and closed active site states.

Structural alignment of the structured region adjacent to the active site (residues 491-587) was conducted using the DALI webserver, revealing domains with structural homology despite having very low pairwise sequence identity (<10%) (**Supplement Table 2, Figure 30, a, b**). This structural domain comprises a minimal fold characterized by a surface-exposed antiparallel β-sheet formed by 3 β-strands (2 long and 1 short), along with inner α-helical and loop elements. It is buried within the protomer and involves two critical residues, Tyr538 and Asp540, which play a role in the mRNA decapping trench of the active site (**Figure 29, b, d**). Notably, this conserved minimal fold shares similarities with the methyltransferase N6AMT and the cytotoxic nuclease domain of immunity protein 5 (im5)<sup>227</sup>, exhibiting a main-chain root-mean-square deviation (rmsd) of 3.2 Å and 2.8 Å, respectively, as reported by DALI (**Supplement Table 2**). Overall, the reported findings reveal an alternative conformation of the mRNA decapping trench, as well as distant homologs that may have evolutionary implications.



**Figure 30.** Identification of structural homologs for the adjacent structural region of the mRNA decapping site using the DALI server.

Identified homologous domains of (a) the methyltransferase enzyme family and (b) the Colicin immunity protein family, both aligned to the mRNA decapping active site (residues I491-V587). Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>



#### 4.2.4 Tubular Densities Inside the Capsid: Insights into RNA-dependent RNA Polymerase Localization from Asymmetric Reconstruction of the L-A Virus

In order to gain a deeper understanding of the L-A virus within cell extracts, a collection of 10,067 movies was obtained from the fractionated extract at a lower pixel size (3.17 Å). This was done to increase the number of imaged single particles available for analysis. The resulting lower-resolution cryo-EM map (Figure 31, a), reconstructed using 103,214 particles in total, exhibited 2D class averages covering various views across the 2-fold, 3-fold, and 5-fold axes (Figure 31, b). The map reached Nyquist at 6.4 Å with a high signal-to-noise ratio (Figure 31, c), as evidenced by the map x,y slices (Figure 31, a-d).

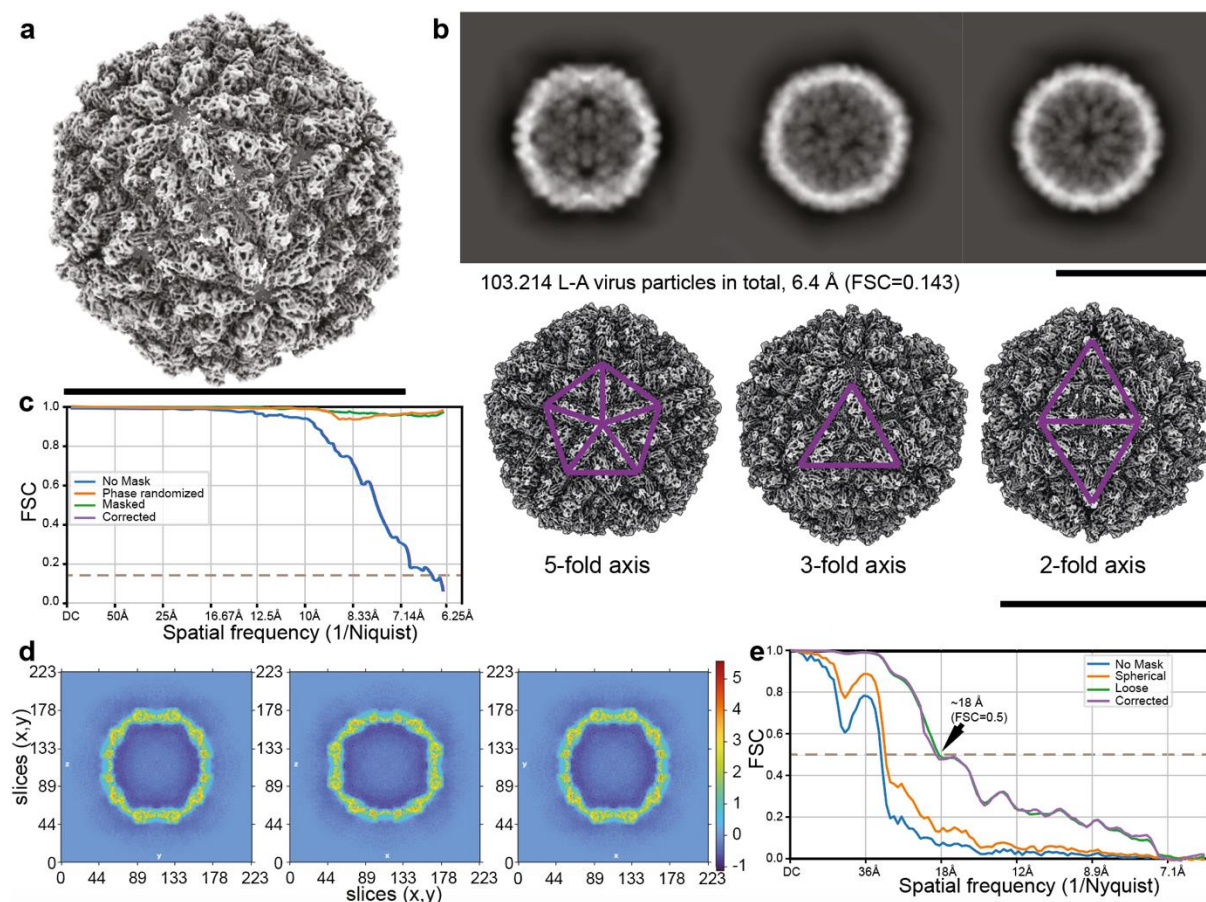
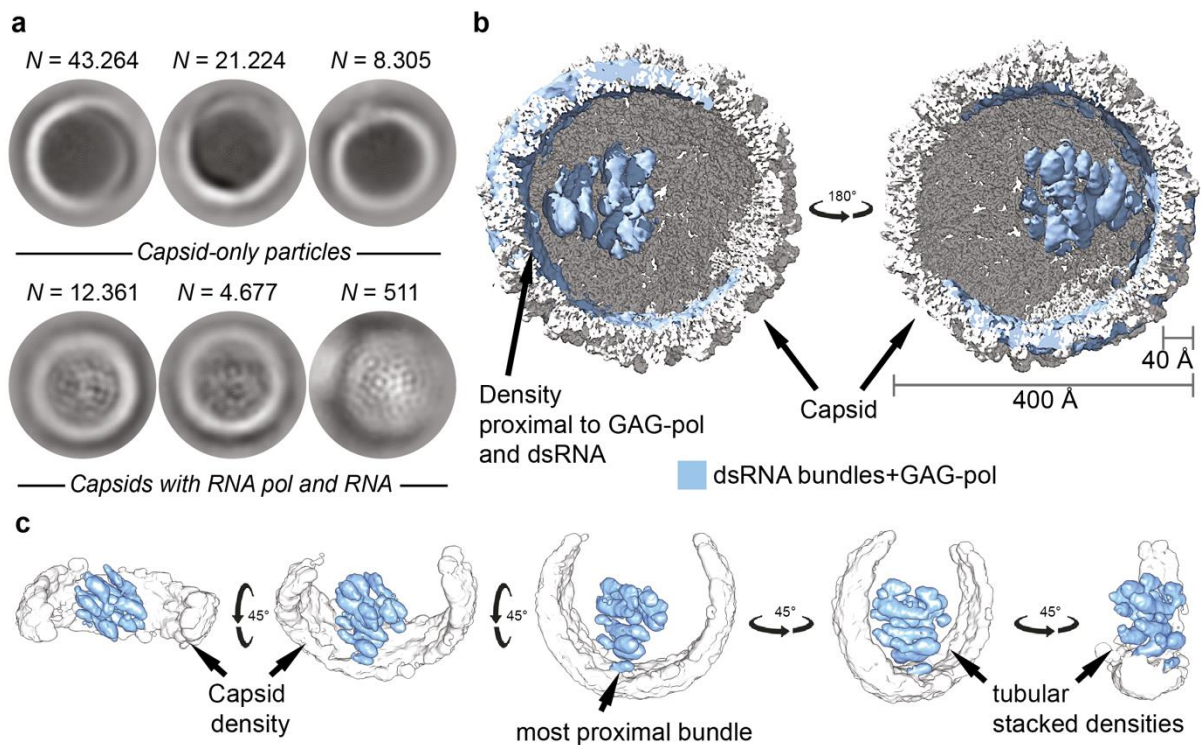


Figure 31. View of the L-A virus by reconstructing a cryo-EM map at lower resolution by applying a larger pixel size (3.17 Å).

(a) 3D reconstructed map of the L-A virus at 6.4 Å. (b) 2D class averages covering views of the 2-, 3- and 5- fold axes. (c) FSC of the 6.4 Å reaching Nyquist. (d) A high signal-to-noise ratio is achieved as shown by the map x,y slices. (e) FSC at 0.5 of the asymmetrically reconstructed inner densities after particle subtraction. Scale bars at 45 nm. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

From the same data, an asymmetric reconstruction was performed after particle subtraction reaching 18 Å (FSC=0.5, Figure 31, e). In detail, following the application of particle subtraction to eliminate the dominant capsid density, the subsequent 2D classification

analysis revealed a diverse population of viral particles (**Figure 32, a**). The majority of the particles appeared to be empty or contained inner signals that were averaged, suggesting the absence of distinct structural features. However, a subset of capsids exhibited pronounced inner densities, indicating the presence of unique components or arrangements within these viral structures. Among the extensive dataset comprising 17,549 particles that exhibited prominent inner density features, an advanced step was taken to reconstruct the asymmetric inner densities. This reconstruction, achieved at an approximate resolution of 18 Å (FSC=0.5, **Figure 31, e**), provided valuable insights into the internal organization of the viral capsid (**Figure 32b-c**). Notably, the reconstructed inner density displayed a remarkable asymmetrical position within the virus capsid, originating from below the 18 Å diameter opening located precisely at the 5-fold axis.



**Figure 32.** View inside the yeast L-A virus retrieved from native, endogenous yeast extracts. (a) 2D classification of inner densities of “full” L-A virus particles after subtracting the capsid density with CRYOSPARC. 2D class averages with capsid-only particles were discarded, and 2D class averages containing particles with internal density were used for 3D reconstruction. (b) 3D reconstruction of inner densities, colored light blue, superimposed on the capsid map (light grey). Arrowheads point at the capsid density and the most proximal inner density bundle. (c) Different views of the reconstructed inner density. Appearing capsid densities colored light grey and inner densities colored light blue shown in different views rotated 45° around the x and y axes. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Within this intriguing inner density, the presence of distinct tubular structures became apparent upon closer examination (**Figure 32, c**). These tubular structures exhibited a striking resemblance to canonically packed double-stranded RNA (dsRNA), suggesting a potential role of this viral entity in encapsulating and protecting genetic material. The tubular arrangement within the inner density offered valuable clues about the potential mechanisms involved in genome packaging, viral assembly, and the structural integrity of the viral particles.

4.2.5 Location and molecular modelling of the L-A virus RNA-dependent RNA polymerase  
 While the structure of the RNA-dependent RNA polymerase of the L-A virus remains unknown and cannot be reliably placed at this resolution, low-resolution densities were observed in proximity to the capsid and within the pentameric face pore (**Figure 33**).

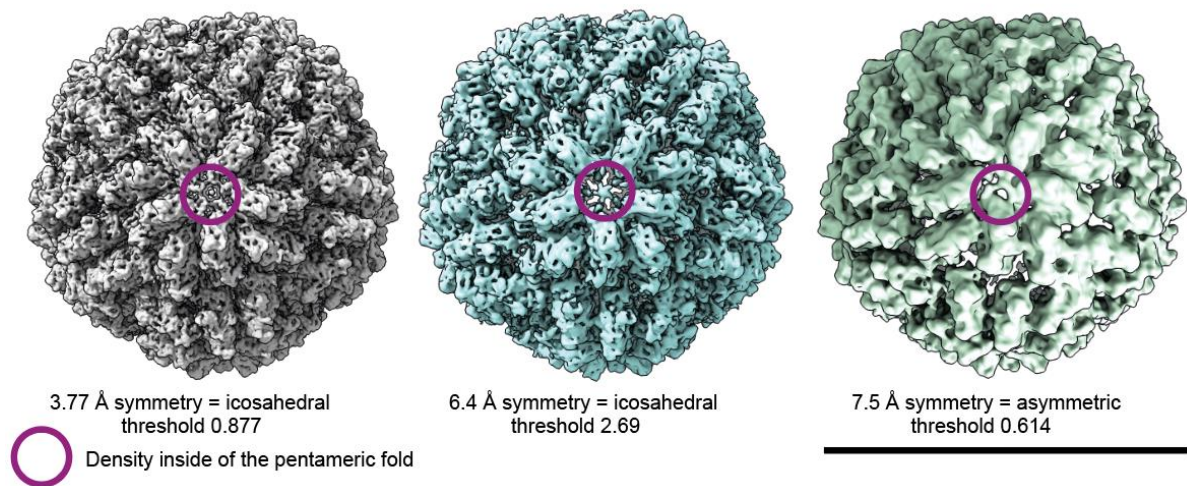
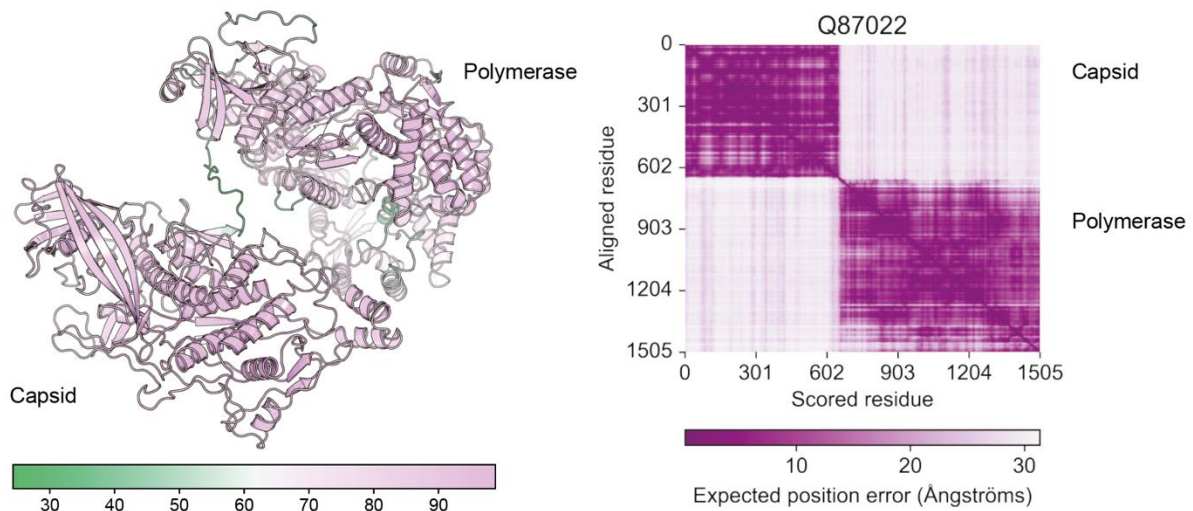


Figure 33. Persistent low-resolution density in the pentameric face pore across L-A virus reconstructions in the 5-fold axis. Scale bar 45 nm.



**Figure 34. AlphaFold2-predicted Gag-Pol fusion protein.**

The predicted structure of the polymerase is connected via a 62-residue flexible linker to the capsid structure. The heat bar below shows the confidence of AlphaFold2 for the model, the heatmap in purple shows the expected position error of each residue, while dark purple represents a low error and lighter color higher errors. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

To determine its relative positioning within the capsid interior, artificial intelligence (AI)-based structural modelling was employed for the 1505 amino acid gag-pol fusion product synthesized upon ribosomal frameshifting. The AI-based model provided a high-quality atomic model for the individual gag and pol domains, with experimental error accounted for in each modeled fold (**Figure 34**). Particularly, the AI-based model of the pol domain demonstrated high quality due to the availability of experimentally determined structural templates in the Protein Data Bank (PDB) (**Figure 34**).

Interestingly, regions with ambiguous modelling included loops, specifically the unstructured linker connecting the C-terminus of gag and the N-terminus of pol (**Figure 35**). The linker, which is highly charged, may play a role in positioning the pol domain relative to the gag domain, potentially folding around the ordered pol molecule in a compacted state.

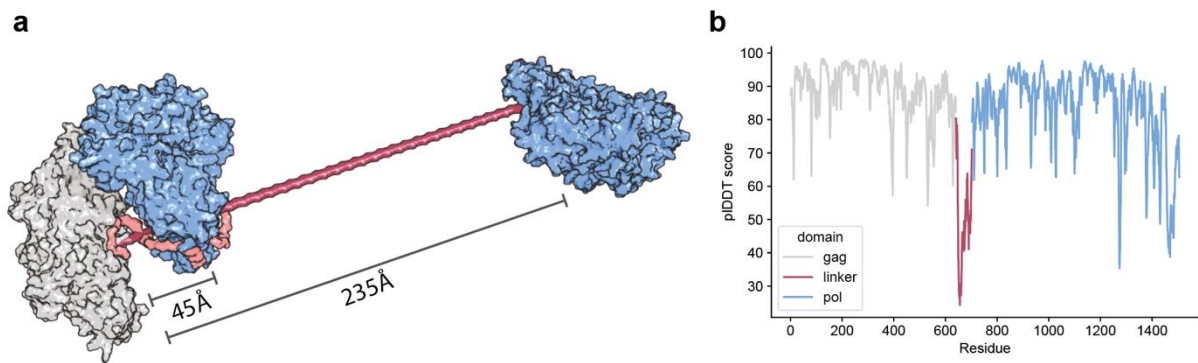


Figure 35. Linker region of Gag-pol and the conformational space that can be explored.

The linker region (purple) shown in (c) is highly positively charged and supposedly positions the Pol (blue) relative to the Gag (grey). AlphaFold2-predicted and maximum distances are drawn at 45 Å and 235 Å, respectively. On the right, the pLDDT score shows the low confidence prediction for the flexible linker region. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Physical-chemical calculations considering the molecular weight and length of the disordered linker suggested various distances separating gag C-terminus and pol N-terminus. In scenarios where the linker is completely folded, the gag C-terminus/pol N-terminus distance was estimated to be approximately 30 Å. In other possible structural states, such as molten globule, unfolded, and/or disordered conformations, distances ranging from 35 to 45 Å were calculated. These calculations indicated that the pol domain primarily explores distances proximal to the pore leading to the five-fold axis, aligning with the proximity of multiple resolved tubular densities. Looking specifically the AlphaFold homologues of the pol domain, results show that the fold is predicted within experimental accuracy, except the N-terminus which is connected to the Gag and must also be flexible in order to move within its restricted nanocompartment (**Figure 36**).

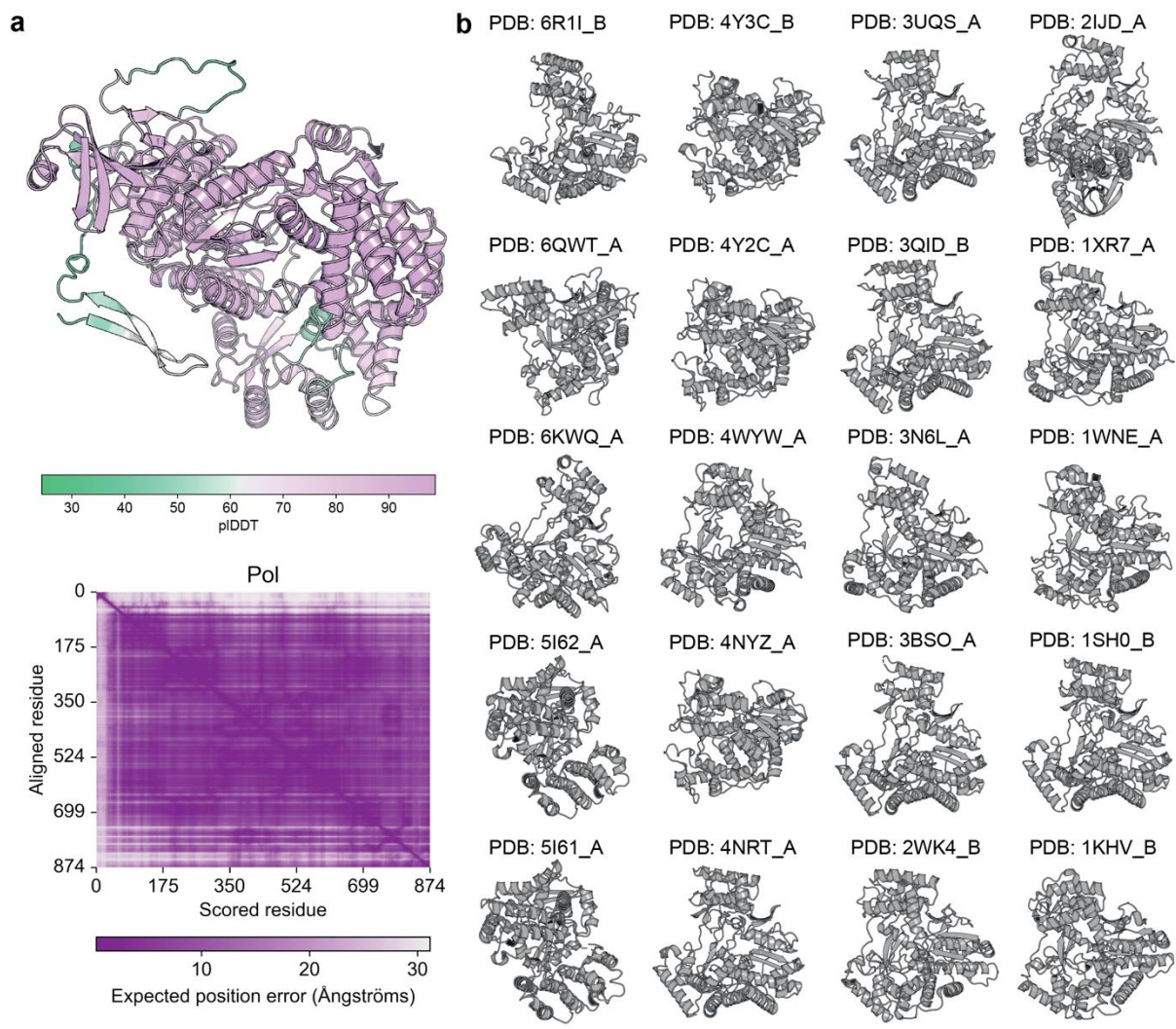


Figure 36. Analysis of the AlphaFold model of L-A virus RNA polymerase.

(a) AlphaFold predicted polymerase. The heat bar below shows the confidence of AlphaFold for the model it created, the heatmap in purple shows the expected position error of each residue while dark purple represents a low error and lighter color high errors. (b) Shown are the templates AlphaFold used for predicting the yeast L-A virus polymerase model. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Taken together, these findings suggest that RNA replication within the L-A virus occurs within an asymmetric nano compartment with a potential size of less than 50 Å per gag-pol protomer. The complementary charge of the linker to dsRNA further supports this confinement of RNA replication within the virus.

#### 4.2.6 Higher-order interactions of the L-A virus and minimal communities: Detection and implications

The cryo-EM acquisition with a lower pixel size (3.17 Å) was analyzed to enhance statistical significance, resulting in a dataset of 10067 movies. Electron-dense material in close proximity to pleomorphic L-A viruses was systematically detected (**Figure 37, a,b**). By

conducting 2D averaging of selected L-A virus particles, we obtained 9711 single-particles exhibiting structural signatures in close proximity. However, the direct proximity must be considered highly flexible due to the diffused densities observed (**Figure 37, b**).

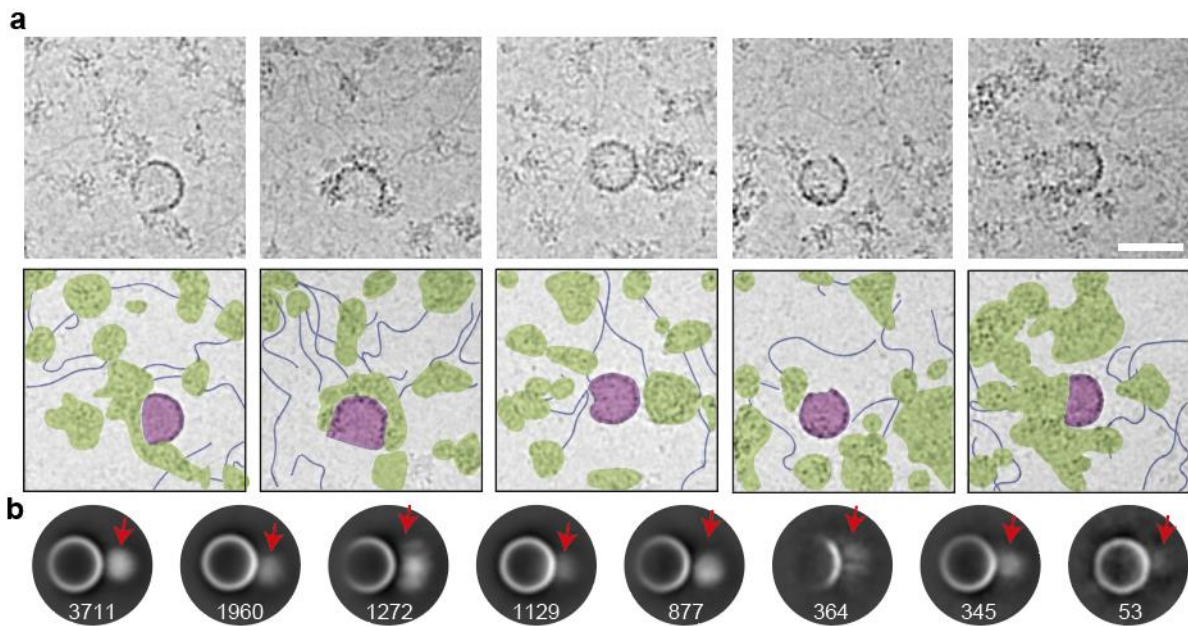


Figure 37. L-A virus proximal biomolecules.

(a) A representation of assembly states corresponding to the viral lifecycle is shown in crops and impressions in distinct steps from I to IX, like Fig. 1a. The scale bar is 50 nm. (b) 2D classes of the LA virus capsid with diffused densities in proximity; ribosomes are often bound as shown in, e.g., panel (a). Figure reproduced and adapted for the dissertation from Schmidt et al.<sup>219</sup>

Additionally, a 3D reconstruction of a ribosome at a resolution of 8.1 Å (unmasked, FSC=0.143) was achieved using 627748 particles (FSC=0.143, **Supplement Figure 4, a Supplement Table 1**), further confirming the high abundance of ribosomes in the same fractions as the L-A virus.

The unambiguous 2D classification of the selected single particles clearly demonstrated the presence of polysomes in the cell extract fraction (**Supplement Figure 4, b**), exhibiting a 3D conformation resembling previously published polysome conformations<sup>220</sup> and stalled disome states<sup>228</sup> (**Supplement Figure 4, b**). From the 3D reconstruction of the ribosome (**Supplement Figure 4, a, c, Supplement Table 1**), we performed 3D variability analysis<sup>229</sup> to characterize different ribosomal classes. Subsequent 3D reconstructions aimed to resolve discrete heterogeneity across individual ribosomes, revealing the presence of all major translation states with tRNAs bound at specific sites within the decoding center, achieving resolutions ranging from 10 to 13 Å (FSC=0.143, **Supplement Figure 4, c and Supplement Figure 5, a-f**). These states were similar to those derived from human ribosomes in multiple

samples<sup>230</sup>. Notably, we determined the occupancies of tRNAs at the exit, peptidyl, and aminoacyl sites, as well as the movements of the L1 stalk (**Supplement Figure 4c**). These findings establish the presence of the L-A virus in the cell extract, in close proximity to ribosomes engaged in active translational processes.

During the analysis of cryo-EM micrographs, we identified the L-A viruses, which exhibited pleomorphic characteristics in addition to their proximity to ribosomes and flexible macromolecules (**Figure 37, Figure 38**).

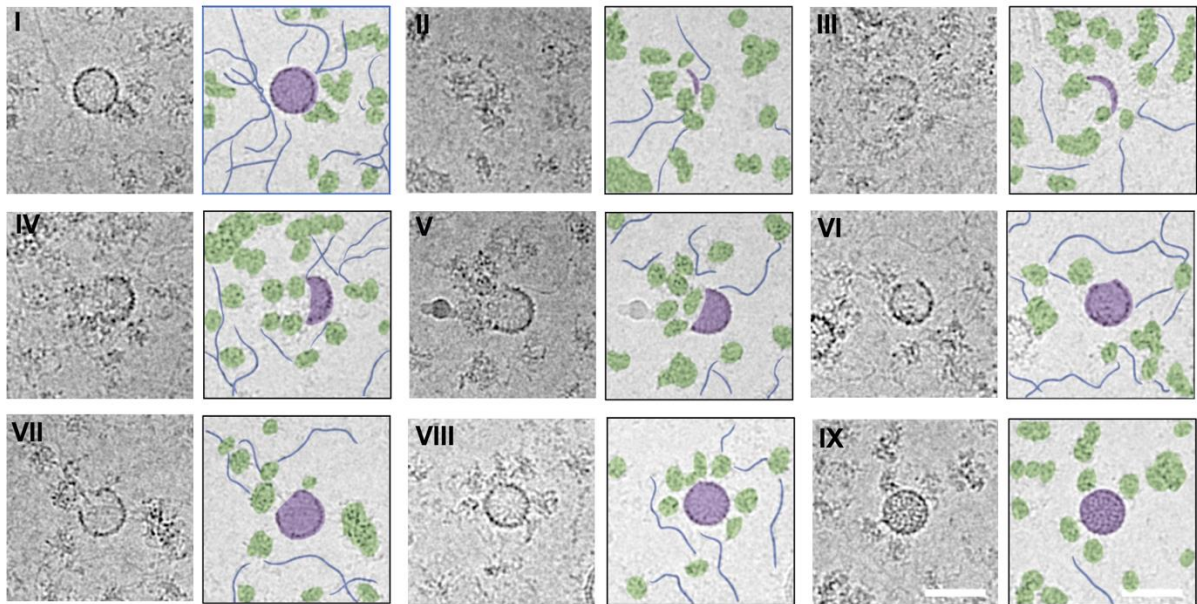


Figure 38. L-A virus particles in the native cell extract with different assembly states.

States were statistically assessed and grouped based on the hypothesized assembly pathway for the L-A virus during its life cycle<sup>231</sup>. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup> Scale bar 50 nm.

Statistical analysis of the native cell extract revealed that the majority of the identified particles were in their mature state (**Figure 38 (I, VIII, IX), Figure 39, a**), while rare pleomorphic states were also observed (**Figure 38 (II-VII)**). Additionally, it is worth noting that indented particles might have been included in the count due to their similarity in shape to the identified states. All particles observed were found to be in close proximity to ribosomes and irregularly shaped macromolecules, which could potentially be mRNA and may be implicated in mRNA decapping, considering the function of the L-A virus capsid.

To assess whether these particles were disrupted during cell lysis, the fractions were subjected to western blotting analysis targeting the L-A virus gag protein (Methods). The results demonstrated a high abundance of the virus in the high molecular weight SEC fractions (**Figure 39, c, Supplement Figure 6**). Conversely, immunodetection of the viral capsid components in the lower-molecular weight fractions did not yield any detectable



signals fractions (**Figure 39, b, Supplement Figure 6**), suggesting that the observed states could correspond to particles minimally damaged by the biochemical treatment. Additionally, the distribution of pleomorphic states and their non-random distribution, indicating a higher prevalence of empty and full viruses, provided further evidence of relatively undisturbed particles (**Figure 39b**). The packaging efficiency of the virus, defined as the ratio between randomly counted full (N=400) or empty (N=661) mature capsids, was estimated to be approximately 40% (**Figure 39b**). However, due to the resolution limitations, no conclusions could be drawn regarding the assembly mechanism at this stage<sup>232</sup>.

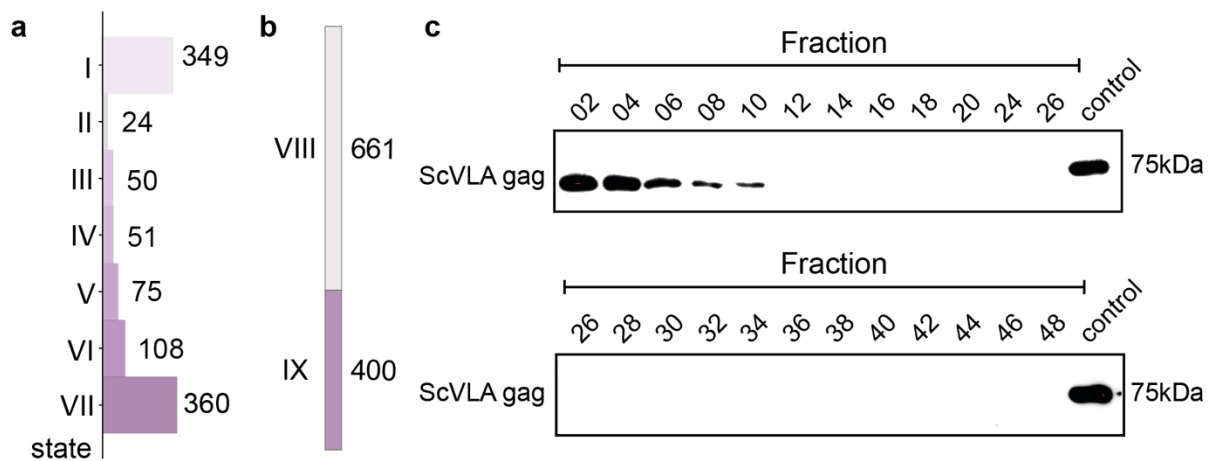


Figure 39. Statistical comparison of all states, as well as empty and full mature L-A viruses. (a) A statistical representation of the assembly states I to VII in absolute numbers. (b) Statistical comparison of full and empty viruses. This shows the ratio between randomly counted full (N=400) or empty (N=661) mature capsids corresponding to states VIII and IX of Figure 38. Western blot analysis against gag shows signal only in high molecular weight fractions indicating that the virus capsids are not likely to be severely damaged. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Another frequent observation in the cryo-EM micrographs of L-A virus-containing yeast native cell extracts was the formation of small groups of the L-A virus (**Figure 40 and Supplement Figure 7**). These groups, composed of an average of  $12 \pm 4$  viruses, predominantly consisted of mature L-A viruses and were found to be in close proximity to other proteins and RNA, while remaining separate from individual particles in the images (**Figure 40, a**). To further investigate this phenomenon, transmission electron microscopy was employed to image high-pressure frozen and vitrified 30 nm thick sections of whole yeast cells. These sections displayed similar groupings, with an average of  $16 \pm 10$  viruses per group (**Figure 40, b**). Statistical analysis demonstrated comparable distances between the nearest viruses and their measured diameters (**Figure 40, c-d**), supporting their correspondence between the in-extract and in-cell environments.

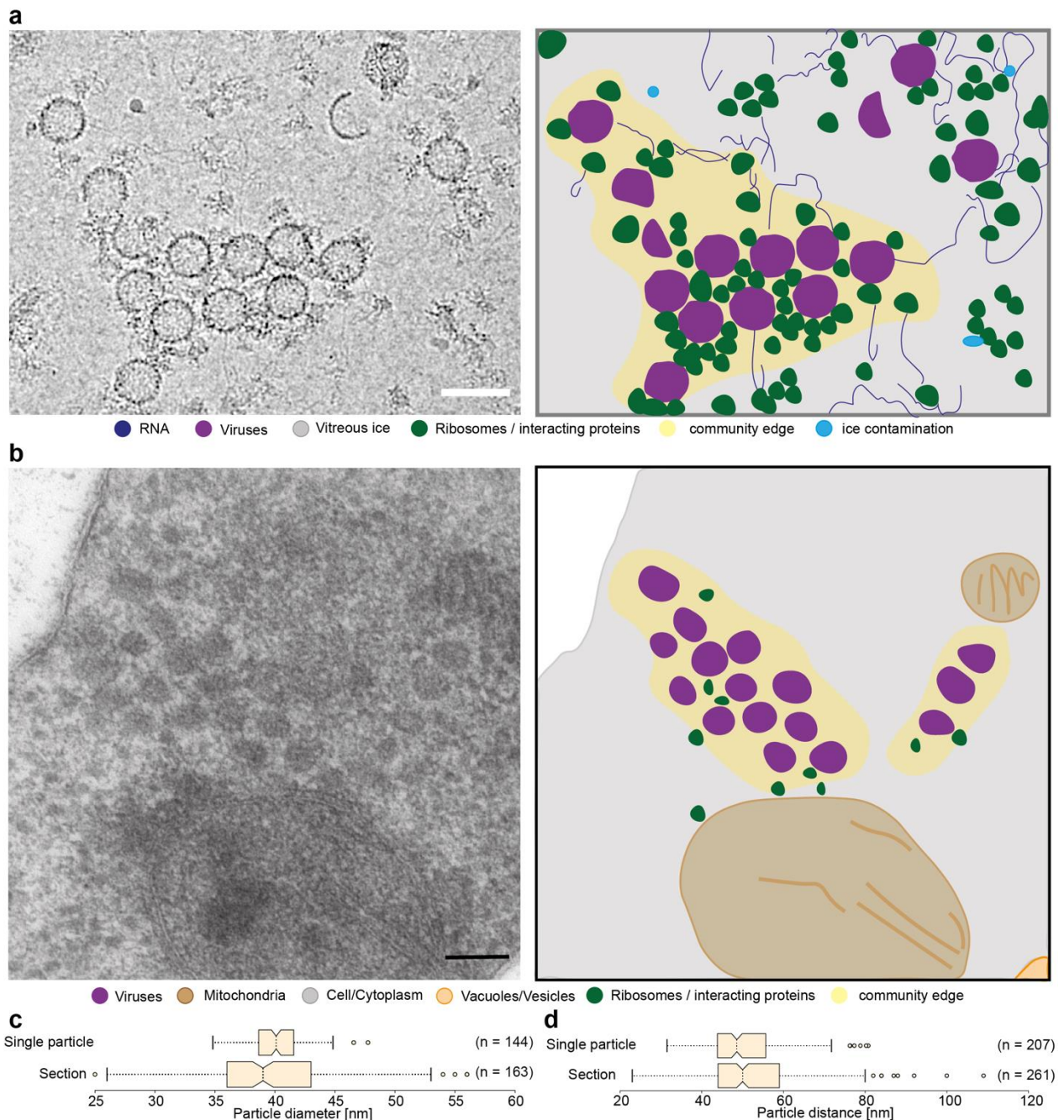


Figure 40 L-A virus communities in native cell extracts and in yeast cells.

(a) Proposed viral communities are shown in a crop of a micrograph. Molecules are colored, such as RNA in dark blue, ribosomes in dark green, ice contamination in light blue and L-A virus in purple. The space around the group of viruses and surrounding proteins is colored yellow. (b) Cryo substituted, in resin embedded *Saccharomyces cerevisiae*, cut into sections of 30 nm in width and imaged on the EM 900. Shown is a crop through yeast on the left and an impression on the right. Color coding stays the same as in (a). The measured particle diameter (c), as well as the distances between L-A viruses (d), are represented by box plots with the standard deviation shown in the whiskers. All scale bars represent 50 nm. Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>

Such assemblies have been widely documented for exogenous viruses and are often associated with membranes or cytoskeletal features, such as connections to the Golgi apparatus or interactions with actin filaments. In the cell section images, the L-A virus communities did not exhibit evident interactions with the cytoskeleton or membranes. However, they were found to colocalize with the translation machinery, predominantly ribosomes, and irregular, electron-dense macromolecules that could represent mRNA (single-particle micrographs). These viral communities appear to function as nano compartments that are relatively isolated from other cellular materials.

## 5. Discussion and Outlook

Insights into the organization and structure of large biomolecular assemblies can be obtained through the application of cryo-electron microscopy (cryo-EM) on native cell extracts. In the past, traditional methods such as protein overexpression, purification, crystallization, heavy labelling of proteins for NMR, or embedding of samples in resin followed by thin sectioning and staining with contrasting agents were employed to probe cell contents. These approaches provided insights into cellular organization and high-resolution structures of proteins. However, their utility in studying interactions or higher order complexes like metabolons at higher resolution was limited.

The resolution revolution in cryo-EM has indeed transformed the field, enabling direct imaging and analysis of cellular content by simply lysing cells and simplifying them using size exclusion chromatography or sucrose gradient centrifugation. Although sucrose gradient centrifugation is less suitable for cryo-EM due to its cryoprotectant properties, native cell extracts, while not as complex as whole cell lysates, still preserve fundamental principles of organization and remain amenable to various other techniques.

Consequently, the identification of actively translating ribosomes within the fractions and the observation of unique features elucidating the organization of virus particles in host cells have become possible and shown in this Dissertation. The key strength of cryo-EM of native cell extracts lies in its ability to combine multiple methods, such as quantitative mass spectrometry (MS) for protein identification, cross-linking MS (XL-MS) for exploring interactions, and the application of cryo-EM and computational methods for structural reconstruction and analysis. Additionally, traditional embedding and sectioning of whole cells can be employed to gain insights into in-cell organization and provide validation for findings obtained through cryo-EM of native cell extracts that, due to cell lysis artifacts, could provide insights that are not necessarily native.

### 5.1 Challenges in cryo-EM analysis of mitochondrial extracts

The choice of organelle isolation method can both resolve and introduce challenges and issues. Employing a protocol with fewer but distinct centrifugation steps and utilizing a non-osmotically active chemical for gradient centrifugation can enhance the vitality of mitochondria and result in higher protein concentrations. In the first method, careful placement of the buffer without disturbing the layers above is possible during pipetting. However, when using a syringe tip to traverse through all layers, there is a potential for

mixing the layers. Conversely, in the second method, the need to traverse layers is eliminated as they are stacked atop each other. Nevertheless, excessive pipetting force can disrupt the interfaces, leading to less distinct boundaries.

While achieving clean mitochondria preparations with sufficient enzymatic activity for PDHc was accomplished by lysing mitochondria using a buffer containing PCC, retrieving the PDHc intact from the fractions proved infeasible due to the likely dissociation of large protein complexes during size exclusion chromatography. PCC might be the issue here, because a detergent can dissociate soluble proteins, even if used in minimal concentrations. In yeast, PDHc appears to exhibit strong associations with membranes or membrane proteins, posing challenges in obtaining the complete complex while maintaining its activity. The complex has been purified with many more steps previously<sup>233</sup> but waiting times could inactivate the complex even if it is amenable to low-resolution cryo-EM<sup>234</sup>. In comparison, the thermophilic fungus *Chaetomium thermophilum*<sup>27,51</sup> exhibited more success in retrieving the full, active PDHc complex. Notably, mesophilic proteins derived from native cell extracts exhibit reduced stability and prove more challenging to detect and reconstruct compared to their thermophilic counterparts.

## 5.2 On the reproducibility of the biochemical and cryo-EM analysis of yeast cell extracts

In this thesis it was possible to adapt the protocol to retrieve native cell extracts from the thermophilic *Ct.* for cryo-EM shown in the work of Kyriulis *et. al*<sup>27,51</sup> to the mesophilic fungus *S. cerevisiae*. The method posed challenges for yeast as proteins seemingly dissociated upon SEC. Still the resulting lysate could be used in sufficient concentration reproducibly for proteomics and structural biology. Through the method of cryo-EM of native cell extracts interesting features could be observed and identified through quantitative MS and modelling. As expected, an enrichment for different higher molecular structures such as ribosomes, also in the form of polysomes, as well as the yeast L-A virus could be achieved. Although the lysis protocol worked efficiently, PDHc could not be isolated. High activity was detected in the pellet after the 100000 g centrifugation step, indicating an association of PDHc with membranes or membrane proteins. During this thesis, the enrichment of PDHc with the method of native cell extracts from yeast was being performed: This marks a major difference between applying the method to a thermophilic fungus or a mesophilic fungus and has implications on the unique biology of each fungus.

### 5.2.1 Proteins and pathways within the studied cellular fraction and their inter-relation to the L-A virus

The fast and reliable screening and analysis of viruses within the native assembly context of eukaryotic cells is crucial for the rapid development of anti-viral treatments, especially in light of the increasing number of viral infection outbreaks witnessed in the 21st century, such as SARS-Cov1 and -Cov2, multiple Ebola outbreaks, and zoonotic-influenced strains like swine flu, avian flu, and the Zika virus<sup>235</sup>. In this study, we present a straightforward and mild protocol that preserves the native state of viral structures, enabling their analysis even for endogenous viruses that are present in low cellular abundance. While viruses employ various mechanisms to survive and replicate within host cells, including the manipulation of cellular pathways, nuclear import (e.g., HIV)<sup>236</sup>, and integration into the host genome, the formation of the viral capsid always represents the final step during maturation. Using the L-A virus as a model<sup>194</sup>, we demonstrate that this process occurs not randomly throughout the cytoplasm but within communities that may involve multiple steps in capsid maturation, closely associated with the translation apparatus.

Furthermore, our research provides valuable insights into the structure of viruses in their endogenous context. The cryo-EM reconstruction of the L-A viral capsid from native cell extract achieves a resolution comparable to the previously published highly purified structure obtained through X-ray crystallography<sup>183</sup>. Notably, this resolution is achieved despite the sample being captured within a highly heterogeneous environment, interacting with other electron-dense cellular material. As expected, the core regions of the capsid exhibit no significant differences, while the flexible loop regions, accessible to solvent, display distinct deviations, which may indicate functional variations within the native environment. Intriguingly, we identified multiple cation- $\pi$  interactions within and across the capsid subunits. These interactions not only contribute significantly to the stability of the capsid but may also facilitate folding and higher-order assembly. Cation- $\pi$  interactions rely on precise positioning of both partners—the delocalized  $\pi$ -electron system and the cationic moiety. Once formed, these interactions are rigid and exhibit clearly resolved side chains.

As previously described, the L-A capsid possesses decapping and 'cap-snatching' activity<sup>196,200</sup> located on the capsid surface. These activities allow the virus to cap its own mRNA, thereby increasing its half-life, and generate mRNA decoys for degradation. In our analysis, we identified a structured region at the active site, sharing structural homology with methyltransferases and the colicin-E5 immunity protein. Since the L-A virus decapping activity specifically targets methylated cap structures, it is plausible that a methyltransferase

activity (or the recruitment of endogenous methyltransferases) assists the endogenous decapping process by methylating endogenous mRNA prior to decapping. Additionally, the E5 immunity protein, which can neutralize colicin-E5—a tRNAse—by mimicking RNA, may serve as an interaction point for RNA binding proteins in the context of the L-A virus. These proteins could potentially degrade viral mRNA or recruit RNA-specific enzymes, such as methyltransferases, to the decapping site, thereby optimizing intrinsic decapping activity.

Overall, through our analysis of the L-A virus within its near-native environment, we unveiled an additional layer of viral architecture. By forming communities, newly transcribed viral mRNA is promptly translated by associated ribosomes, resulting in the formation of distinct polysomal structures at the site of capsid formation. This tight association of ribosomes with viral mRNA prevents mRNA degradation. In the context of the 'cap-snatching' mechanism, capping the viral mRNA may not solely serve as a protective measure but also enhance the translation of viral proteins by decorating the mRNA with a dense layer of ribosomes. In the future, we aim to resolve the L-A virus at higher resolution, as well as resolve the intermediate states at higher resolution as well. A dataset has been acquired with an even smaller pixel size (0.97 Å), which provided a structure of the L-A virus at 3.2 Å (results not shown).

#### 5.2.2 Potential role of disorder in the structure, function, and life cycle of the L-A virus

The L-A virus, a dsRNA virus, exhibits intriguing structural and functional adaptations in its native form, as revealed by high-resolution cryo-EM studies. The cryo-EM structure of the L-A virus obtained from a cell extract provides valuable insights into the importance of molecular disorder in the virus.

One of the key findings is the identification of stabilizing interactions called cation- $\pi$  interactions within the capsid of the L-A virus. These interactions, formed at both monomeric and dimeric interfaces, play a crucial role in the folding of individual building blocks and the assembly of the virus into higher-order structures. The presence of cation- $\pi$  interactions highlights the significance of molecular disorder in facilitating specific and robust non-covalent bonds within the virus. The cryo-EM structure also reveals conformational differences in the capsomere between different monomers and the presence of specific adaptations within the cell extract. These variations demonstrate the flexibility of the virus and its ability to adapt to different environments. Flexible loops and regions within the capsomere, such as the helix-turn-helix fold region and loop regions involved in mRNA decapping activity, contribute to the structural dynamics and functional versatility of the

virus. Furthermore, the cryo-EM analysis uncovers tubular densities inside the capsid, resembling canonically packed dsRNA. These observations suggest that RNA replication within the L-A virus occurs within an asymmetric nano compartment, confining the replication process within the virus. The proximity of the RNA-dependent RNA polymerase to the capsid and the presence of tubular densities provide insights into the localization and function of the polymerase in the virus. The cryo-EM studies also detect electron-dense material near pleomorphic L-A viruses, which are found in the cell extract along with ribosomes engaged in active translational processes. These findings indicate higher-order interactions between the L-A virus and other biomolecules, such as ribosomes, within the cellular environment. The presence of polysomes and ribosomes in the vicinity of the virus suggests potential interactions and crosstalk between translation and viral replication processes. Overall, these results can be put in context and propose a life cycle for the L-A virus based on the data collectively shown in this thesis (**Figure 41**).



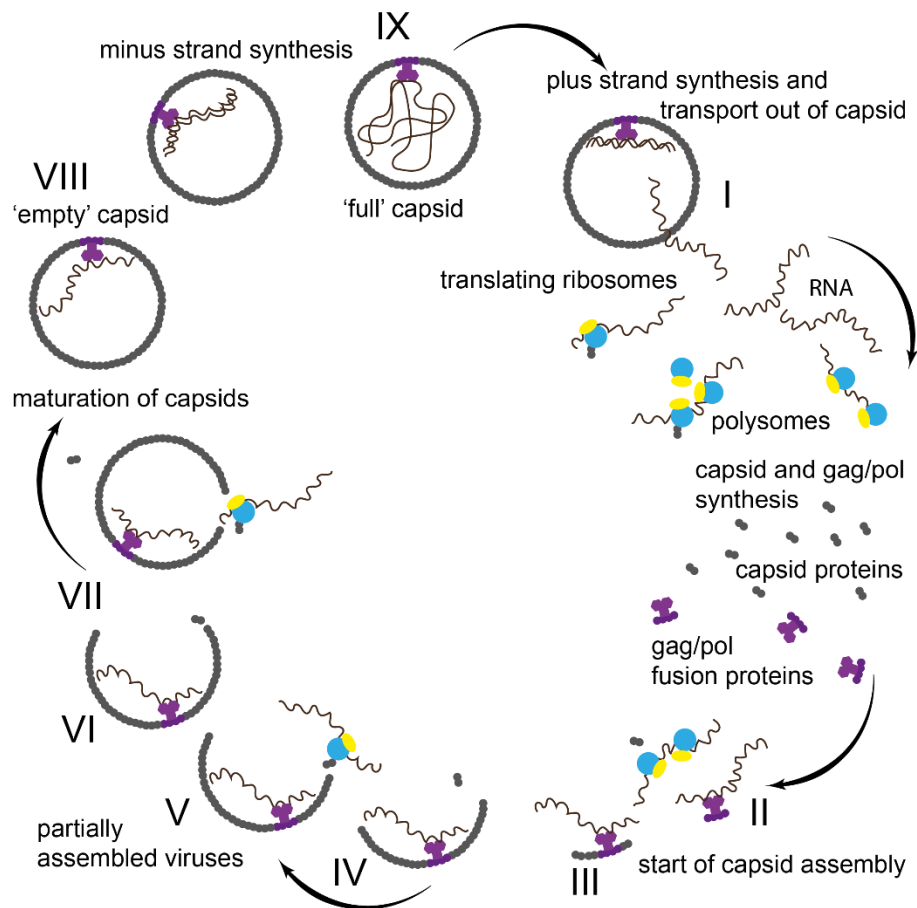


Figure 41. Proposed scheme for the yeast L-A virus life cycle (states I-IX).

Replication starts with plus-strand synthesis and the transport out of the capsid (state I). Yeast ribosomes then translate the ssRNA and new viral particles (capsids) form (states II-VII). Last, the minus-strand is synthesized (states VIII-IX). Genome replication occurs in assembled viruses (state IX). Figure reproduced and adapted for the Dissertation from Schmidt et al.<sup>219</sup>.

In conclusion, the cryo-EM analysis of the L-A virus emphasizes the importance of molecular disorder in its structural and functional adaptations. The presence of stabilizing interactions, conformational flexibility, and higher-order interactions with cellular components highlight the dynamic nature of the virus and its ability to interact with its environment. Understanding the significance of molecular disorder in the L-A virus can provide valuable insights into the mechanisms of viral assembly, replication, and interaction with host cellular machinery.

## 6. Summary

This dissertation aimed to investigate the feasibility of using organellar and cellular extracts from *Saccharomyces cerevisiae* for high-resolution structural analysis. The use of cell extracts offers significant advantages over purified or overexpressed equivalents as it allows for the study of specimens that closely resemble their natural state, thus retaining cellular interactions.

The first part of the dissertation focused on the feasibility of using mitochondrial cell extracts. The choice of overlaying less dense buffer solutions on top of higher concentrated ones was based on optimization, with the Nycodenz gradient preferred over the sucrose gradient due to its minimal impact on the mitochondrial structure. The effectiveness of the mitochondrial purification could be validated via western blot analysis and vitality staining via MitoTracker, while cryo-electron microscopy (cryo-EM) provided additional evidence by visualizing ATP synthase on the mitochondrial membranes. The study also focused on retrieving the pyruvate dehydrogenase complex (PDHc) from the mitochondria, demonstrating functional PDHc activity in clean mitochondria obtained through the isolation process. PDHc was found to be bound to the mitochondrial membranes. Overall, this study successfully recovered active yeast mitochondria, validated the purity and vitality of the organelles, and provided insights into the interaction of specific proteins with mitochondrial membranes.

In the second part of the dissertation, cryo-EM combined with modelling techniques was utilized to analyze a yeast cell extract, with a specific focus on the L-A virus capsid and its interactions. This cryo-EM map of the L-A helper virus consisted of 120 protomers arranged in an icosahedral protein shell with consistent architecture and the resulting model showed a reasonable fit into it. The study identified previously unreported cation- $\pi$  interactions within the capsid, which were crucial for folding and assembly. Flexibility and conformational variations were observed in different regions of the capsid, particularly in areas involved in assembly and stability, deviating from the crystal structure. The study proposes a model for capsomere stability based on energetic calculations, suggesting a multistep process for capsomere association. Furthermore, the study explored the location and characteristics of the mRNA decapping site within the capsid, revealing the outer surface as the site for decapping, protecting against host exoribonuclease targeting. Flexible loops in this region influenced the shape of the decapping site trench, with the cryo-EM structure revealing a more occluded state compared to the crystal structure. The localization of RNA-dependent RNA polymerase

within the capsid was visualized using cryo-EM and AI-based modelling, which revealed the arrangement of dsRNA within the capsid through tubular densities, also revealing a flexible linker connecting the polymerase to the capsid structure. The charged linker was suggested to position the Pol domain relative to the Gag domain, with physical-chemical calculations indicating varying distances between their termini. Further cryo-EM analysis indicated the presence of ribosomes in close proximity to the L-A virus, suggesting their involvement in virus-associated translation processes. The study also identified pleomorphic L-A viruses and the formation of small groups of viruses functioning as nano compartments. Overall, this study provided an enhanced understanding of the L-A virus capsid's structure, interactions, flexibility, and functional sites, shedding light on its assembly, stability, and molecular mechanisms.

The studies conducted in this dissertation highlighted the strengths of cryo-EM in combination with various techniques for analyzing mitochondrial and cellular extracts from *S. cerevisiae*. The challenges and reproducibility of biochemical and cryo-EM analysis of yeast cell extracts were also discussed. Overall, cryo-EM of native cell extracts provided valuable insights into the cellular organization and the structure of biomolecular assemblies, with the potential to uncover unprecedented information in the future.

## 7. Zusammenfassung

Ziel dieser Dissertation war es die Möglichkeit zu untersuchen, organellare und zelluläre Extrakte aus *Saccharomyces cerevisiae* für die hochauflösende Strukturanalyse zu nutzen. Die Verwendung von Zellextrakten bietet erhebliche Vorteile gegenüber gereinigten oder überexprimierten Äquivalenten, da sie die Untersuchung von Proben ermöglicht, die ihrem natürlichen Zustand ähneln und somit die zellulären Interaktionen erhalten bleiben. Der erste Teil der Dissertation konzentrierte sich auf die Möglichkeit der Verwendung von mitochondrialen Zellextrakten. Die Wahl der Isolation fiel auf die Dichtegradientenzentrifugation, wobei der Nycodenz-Gradient dem Saccharose-Gradienten vorgezogen wurde, aufgrund seiner minimalen Auswirkungen auf die mitochondriale Struktur. Die Reinheit der isolierten Mitochondrien konnte mittels Western-Blot-Analyse validiert werden. Vitalitätsfärbung mittels MitoTracker und Kryo-Elektronenmikroskopie (cryo-EM) durch die die ATP-Synthase auf den mitochondrialen Membranen visualisiert wurde, lieferte zusätzliche Beweise zur Validierung der isolierten Mitochondrien. Die Studie konzentrierte sich auch auf die Gewinnung des Pyruvat-Dehydrogenase-Komplexes (PDHc) aus den Mitochondrien und zeigte eine funktionelle PDHc-Aktivität in sauberen Mitochondrien, die durch den Isolierungsprozess gewonnen wurden. PDHc wurde an die Mitochondrienmembranen gebunden gefunden. Insgesamt haben diese Studien erfolgreich Hefe-Mitochondrien gewonnen und isoliert, die Reinheit und Vitalität der Organellen bestätigt und Einblicke in die Interaktion von spezifischen Proteinen mit mitochondrialen Membranen gewonnen.

Im zweiten Teil, der Dissertation wurde die Kryo-EM in Kombination mit Modellierungstechniken eingesetzt, um Hefezelleextrakte zu analysieren, wobei der Schwerpunkt auf dem L-A-Viruskapsid und seinen Wechselwirkungen lag. Die Kryo-EM-Karte des L-A-Helfervirus bestand aus 120 Protomeren, die in einer ikosaedrischen Proteinhülle mit konsistenter Architektur angeordnet sind und das sich daraus ergebende Modell passte gut zu dieser Architektur. Die Studie identifizierte bisher unbekannte Kationen- $\pi$ -Wechselwirkungen innerhalb und zwischen den Kapsiduntereinheiten, die für die Faltung und den Zusammenbau des Kapsids entscheidend sind. Flexibilität und Konformationsänderungen wurden in verschiedenen Regionen des Kapsids beobachtet, insbesondere in Bereichen, die für den Zusammenbau und die Stabilität wichtig sind und von der Kristallstruktur abweichen. In der Studie wird ein Modell für die Stabilität von Kapsomeren postuliert, das auf energetischen Berechnungen beruht, und auf einen

mehrstufigen Prozess für Assoziation hindeutet. Darüber hinaus untersuchte die Studie den Ort und die Eigenschaften der mRNA-Decapping-Stelle innerhalb des Kapsids, wobei sich die äußere Oberfläche als Decapping-Stelle, die vor dem Angriff der Exoribonuklease des Wirts schützt, herausstellte. Flexible Schleifen in dieser Region beeinflussten die Form des Decapping-Grabens, wobei die Kryo-EM-Struktur einen stärker verschlossenen Zustand im Vergleich zu der Kristallstruktur zeigt. Die Lokalisierung der RNA-abhängigen RNA-Polymerase innerhalb der Cryo-EM-Karte und AI-basierter Modellierung visualisiert, die Anordnung der dsRNA innerhalb des Kapsids als röhrenförmige Verdichtungen, die durch eine mit einem flexiblen Linker verbundenen Polymerase an der Kapsidstruktur verbunden ist. Der geladene Linker dient der Positionierung der Pol-Domäne relativ zur Gag Domäne, wobei physikalisch-chemische Berechnungen auf unterschiedliche Abstände zwischen ihren Endpunkten hinweisen. Weitere Kryo-EM-Analysen zeigten das Vorhandensein von Ribosomen in unmittelbarer Nähe des L-A-Virus, was auf ihre Beteiligung an Virus-assoziierten Translationsprozesse hinweist. In der Studie wurden auch pleomorphe L-A-Viren und die Bildung von kleinen Gruppen von Viren, die als Nanokompartimente bezeichnet wurden entdeckt. Insgesamt hat diese Studie zu einem besseren Verständnis der Struktur des L-A Viruskapsid, Wechselwirkungen, Flexibilität und Funktionsstellen beigetragen. Sie wirft dabei ein Licht auf dessen Aufbau, seiner Stabilität und seiner molekularen Mechanismen. Die in dieser Dissertation durchgeführten Studien zeigten die Stärken der Kryo-EM in Kombination mit verschiedenen Techniken zur Analyse von mitochondrialen und zellulären Extrakten aus *S. cerevisiae*. Die Herausforderungen und die Reproduzierbarkeit der biochemischen und Kryo-EM Analyse von Hefezelleextrakten wurden ebenfalls diskutiert. Insgesamt lieferte die Kryo-EM von nativen Zellextrakten wertvolle Einblicke in die zelluläre Organisation und die Struktur biomolekularer Komplexe, mit dem Potenzial, in Zukunft noch nie dagewesene Informationen aufzudecken.

## 8. References

- 1 Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 546, 563-547 (1996).  
<https://doi.org:10.1126/science.274.5287.546>
- 2 Botstein, D., Chervitz, S. A. & Cherry, J. M. Yeast as a model organism. *Science* **277**, 1259-1260 (1997). <https://doi.org:10.1126/science.277.5330.1259>
- 3 Cazzanelli, G. *et al.* The Yeast *Saccharomyces cerevisiae* as a Model for Understanding RAS Proteins and their Role in Human Tumorigenesis. *Cells* **7** (2018).  
<https://doi.org:10.3390/cells7020014>
- 4 Hayles, J. & Nurse, P. Introduction to Fission Yeast as a Model System. *Cold Spring Harb Protoc* **2018** (2018). <https://doi.org:10.1101/pdb.top079749>
- 5 Kawashima, S. A., Takemoto, A., Nurse, P. & Kapoor, T. M. Analyzing fission yeast multidrug resistance mechanisms to develop a genetically tractable model system for chemical biology. *Chem Biol* **19**, 893-901 (2012).  
<https://doi.org:10.1016/j.chembiol.2012.06.008>
- 6 Bartlett, R. & Nurse, P. Yeast as a model system for understanding the control of DNA replication in Eukaryotes. *Bioessays* **12**, 457-463 (1990).  
<https://doi.org:10.1002/bies.950121002>
- 7 Glick, B. S. & Pon, L. A. Isolation of highly purified mitochondria from *Saccharomyces cerevisiae*. *Methods Enzymol* **260**, 213-223 (1995). [https://doi.org:10.1016/0076-6879\(95\)60139-2](https://doi.org:10.1016/0076-6879(95)60139-2)
- 8 Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147 (2002).  
<https://doi.org:10.1038/415141a>
- 9 Burley, S. K. *et al.* RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res* **49**, D437-D451 (2021).  
<https://doi.org:10.1093/nar/gkaa1038>
- 10 Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235-242 (2000).  
<https://doi.org:10.1093/nar/28.1.235>
- 11 Anderson, C. M., Stenkamp, R. E. & Steitz, T. A. Sequencing a protein by x-ray crystallography. II. Refinement of yeast hexokinase B co-ordinates and sequence at 2.1 Å resolution. *J Mol Biol* **123**, 15-33 (1978). [https://doi.org:10.1016/0022-2836\(78\)90374-1](https://doi.org:10.1016/0022-2836(78)90374-1)
- 12 Kuhlbrandt, W. Biochemistry. The resolution revolution. *Science* **343**, 1443-1444 (2014). <https://doi.org:10.1126/science.1251652>
- 13 Stewart, G. G. *Brewing and Distilling Yeasts (The Yeast Handbook)*. 1 edn, Vol. 343 (2019).
- 14 Claude, A. The Constitution of Protoplasm. *Science* **97**, 451-456 (1943).  
<https://doi.org:10.1126/science.97.2525.451>
- 15 Chance, B., Hess, B. & Betz, A. DPNH oscillations in a cell-free extract of *S. carlsbergensis*. *Biochem Biophys Res Commun* **16**, 182-187 (1964).  
[https://doi.org:10.1016/0006-291x\(64\)90358-4](https://doi.org:10.1016/0006-291x(64)90358-4)

- 16 Chance, B., Schoener, B. & Elsaesser, S. Control of the Waveform of Oscillations of the Reduced Pyridine Nucleotide Level in a Cell-Free Extract. *Proc Natl Acad Sci U S A* **52**, 337-341 (1964). <https://doi.org:10.1073/pnas.52.2.337>
- 17 Hess, B., Chance, B. & Betz, A. Isolierung eines oszillierenden Systems aus S. Carlsbergensis. *Berichte der Bunsengesellschaft für physikalische Chemie* **68**, 768-773 (1964). <https://doi.org:https://doi.org/10.1002/bbpc.19640680823>
- 18 Lynen, F. Biosynthesis of saturated fatty acids. *Fed Proc* **20**, 941-951 (1961).
- 19 Gunenc, A. N., Graf, B., Stark, H. & Chari, A. Fatty Acid Synthase: Structure, Function, and Regulation. *Subcell Biochem* **99**, 1-33 (2022). [https://doi.org:10.1007/978-3-031-00793-4\\_1](https://doi.org:10.1007/978-3-031-00793-4_1)
- 20 Heil, C. S., Wehrheim, S. S., Paithankar, K. S. & Grininger, M. Fatty Acid Biosynthesis: Chain-Length Regulation and Control. *Chembiochem* **20**, 2298-2321 (2019). <https://doi.org:10.1002/cbic.201800809>
- 21 Duve, C. Exploring cells with a centrifuge. *Science* **189**, 186-194 (1975). <https://doi.org:10.1126/science.1138375>
- 22 Nirenberg, M. & Leder, P. Rna Codewords and Protein Synthesis. The Effect of Trinucleotides Upon the Binding of Srna to Ribosomes. *Science* **145**, 1399-1407 (1964). <https://doi.org:10.1126/science.145.3639.1399>
- 23 Claude, A. & Potter, J. S. Isolation of Chromatin Threads from the Resting Nucleus of Leukemic Cells. *J Exp Med* **77**, 345-354 (1943). <https://doi.org:10.1084/jem.77.4.345>
- 24 Medicine., T. N. P. i. P. o. *The Nobel Prize in Physiology or Medicine 1974 was awarded jointly to Albert Claude, Christian de Duve and George E. Palade "for their discoveries concerning the structural and functional organization of the cell"*, <<https://www.nobelprize.org/prizes/medicine/1974/summary/>> (1974).
- 25 Kyrilis, F. L., Meister, A. & Kastritis, P. L. Integrative biology of native cell extracts: a new era for structural characterization of life processes. *Biol Chem* **400**, 831-846 (2019). <https://doi.org:10.1515/hsz-2018-0445>
- 26 Daum, B., Walter, A., Horst, A., Osiewacz, H. D. & Kuhlbrandt, W. Age-dependent dissociation of ATP synthase dimers and loss of inner-membrane cristae in mitochondria. *Proc Natl Acad Sci U S A* **110**, 15301-15306 (2013). <https://doi.org:10.1073/pnas.1305462110>
- 27 Kyrilis, F. L. *Cryo-electron microscopy studies of native cell extracts – Elucidating an active pyruvate dehydrogenase complex from Chaetomium thermophilum* Ph.D. thesis, Martin-Luther University Halle-Wittenberg, (2022).
- 28 Dignam, J. D. Preparation of extracts from higher eukaryotes. *Methods Enzymol* **182**, 194-203 (1990). [https://doi.org:10.1016/0076-6879\(90\)82017-v](https://doi.org:10.1016/0076-6879(90)82017-v)
- 29 McBride, H. M., Neuspiel, M. & Wasiak, S. Mitochondria: more than just a powerhouse. *Curr Biol* **16**, R551-560 (2006). <https://doi.org:10.1016/j.cub.2006.06.054>
- 30 Pucadyil, T. J., Chipuk, J. E., Liu, Y., O'Neill, L. & Chen, Q. The multifaceted roles of mitochondria. *Mol Cell* **83**, 819-823 (2023). <https://doi.org:10.1016/j.molcel.2023.02.030>
- 31 Wallace, D. C. Mitochondria and cancer. *Nat Rev Cancer* **12**, 685-698 (2012). <https://doi.org:10.1038/nrc3365>
- 32 Ernster, L. & Schatz, G. Mitochondria: a historical review. *J Cell Biol* **91**, 227s-255s (1981). <https://doi.org:10.1083/jcb.91.3.227s>

- 33 Altmann, R. *Die Elementarorganismen und ihre Beziehungen zu den Zellen*. (Veit, Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, 1890).
- 34 Benda, C. Über die Spermatogenese der Vertebraten und höherer Evertibraten, II. Theil:Die Histogenese der Spermien. *Archiv für Anatomie und Physiologie* **73**, 393-398 (1898).
- 35 Claude, A. Particulate Components of Normal and Tumor Cells. *Science* **91**, 77-78 (1940). <https://doi.org:10.1126/science.91.2351.77>
- 36 Kennedy, E. P. & Lehninger, A. L. Oxidation of fatty acids and tricarboxylic acid cycle intermediates by isolated rat liver mitochondria. *J Biol Chem* **179**, 957-972 (1949).
- 37 Friedkin, M. & Lehninger, A. L. Esterification of inorganic phosphate coupled to electron transport between dihydrodiphosphopyridine nucleotide and oxygen. *J Biol Chem* **178**, 611-644 (1949).
- 38 Loomis, W. F. & Lipmann, F. Reversible inhibition of the coupling between phosphorylation and oxidation. *J Biol Chem* **173**, 807 (1948).
- 39 Palade, G. E. The Organization of Living Matter. *Proc Natl Acad Sci U S A* **52**, 613-634 (1964). <https://doi.org:10.1073/pnas.52.2.613>
- 40 Palade, G. E. An electron microscope study of the mitochondrial structure. *J Histochem Cytochem* **1**, 188-211 (1953). <https://doi.org:10.1177/1.4.188>
- 41 Frey, T. G. & Mannella, C. A. The internal structure of mitochondria. *Trends Biochem Sci* **25**, 319-324 (2000). [https://doi.org:10.1016/s0968-0004\(00\)01609-1](https://doi.org:10.1016/s0968-0004(00)01609-1)
- 42 Gammage, P. A. & Frezza, C. Mitochondrial DNA: the overlooked oncogenome? *BMC Biol* **17**, 53 (2019). <https://doi.org:10.1186/s12915-019-0668-y>
- 43 Taanman, J. W. The mitochondrial genome: structure, transcription, translation and replication. *Biochim Biophys Acta* **1410**, 103-123 (1999). [https://doi.org:10.1016/s0005-2728\(98\)00161-3](https://doi.org:10.1016/s0005-2728(98)00161-3)
- 44 Kafkia, E. *et al.* Operation of a TCA cycle subnetwork in the mammalian nucleus. *Sci Adv* **8**, eabq5206 (2022). <https://doi.org:10.1126/sciadv.abq5206>
- 45 Kuhlbrandt, W. Structure and function of mitochondrial membrane protein complexes. *BMC Biol* **13**, 89 (2015). <https://doi.org:10.1186/s12915-015-0201-x>
- 46 Zhao, Q. *et al.* A mitochondrial specific stress response in mammalian cells. *EMBO J* **21**, 4411-4419 (2002). <https://doi.org:10.1093/emboj/cdf445>
- 47 Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci* **41**, 211-218 (2016). <https://doi.org:10.1016/j.tibs.2015.12.001>
- 48 Chistiakov, D. A., Sobenin, I. A., Revin, V. V., Orekhov, A. N. & Bobryshev, Y. V. Mitochondrial aging and age-related dysfunction of mitochondria. *Biomed Res Int* **2014**, 238463 (2014). <https://doi.org:10.1155/2014/238463>
- 49 Razvi, A. & Scholtz, J. M. Lessons in stability from thermophilic proteins. *Protein Sci* **15**, 1569-1578 (2006). <https://doi.org:10.1110/ps.062130306>
- 50 Kastiris, P. L. *et al.* Capturing protein communities by structural proteomics in a thermophilic eukaryote. *Mol Syst Biol* **13**, 936 (2017). <https://doi.org:10.15252/msb.20167412>
- 51 Kyrilis, F. L. *et al.* Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts. *Cell Rep* **34**, 108727 (2021). <https://doi.org:10.1016/j.celrep.2021.108727>
- 52 Tuting, C. *et al.* Cryo-EM snapshots of a native lysate provide structural insights into a metabolon-embedded transacetylase reaction. *Nat Commun* **12**, 6933 (2021). <https://doi.org:10.1038/s41467-021-27287-4>



- 53 Skalidis, I. *et al.* Structural analysis of an endogenous 4-megadalton succinyl-CoA-generating metabolon. *Commun Biol* **6**, 552 (2023). <https://doi.org:10.1038/s42003-023-04885-0>
- 54 Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737-741 (2003). <https://doi.org:10.1038/nature02046>
- 55 UniProt, C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523-D531 (2023). <https://doi.org:10.1093/nar/gkac1052>
- 56 Madadlou, A., O'Sullivan, S. & Sheehan, D. Fast Protein Liquid Chromatography. *Methods Mol Biol* **1485**, 365-373 (2017). [https://doi.org:10.1007/978-1-4939-6412-3\\_19](https://doi.org:10.1007/978-1-4939-6412-3_19)
- 57 Kristensen, A. R., Gsponer, J. & Foster, L. J. A high-throughput approach for measuring temporal changes in the interactome. *Nat Methods* **9**, 907-909 (2012). <https://doi.org:10.1038/nmeth.2131>
- 58 Chemistry., T. N. P. i. *The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel "for the development of multiscale models for complex chemical systems"*, <nobelprize.org/prizes/chemistry/2013/summary/> (2013).
- 59 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). <https://doi.org:10.1038/s41586-021-03819-2>
- 60 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022). <https://doi.org:10.1093/nar/gkab1061>
- 61 Rapp, M., Shapiro, L. & Frank, J. Contributions of single-particle cryoelectron microscopy toward fighting COVID-19. *Trends Biochem Sci* **47**, 117-123 (2022). <https://doi.org:10.1016/j.tibs.2021.10.005>
- 62 Piersimoni, L., Kastiris, P. L., Arlt, C. & Sinz, A. Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein-Protein Interactions horizontal line A Method for All Seasons. *Chem Rev* **122**, 7500-7531 (2022). <https://doi.org:10.1021/acs.chemrev.1c00786>
- 63 Joyce, A. R. & Palsson, B. O. The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol* **7**, 198-210 (2006). <https://doi.org:10.1038/nrm1857>
- 64 Przybyla, L. & Gilbert, L. A. A new era in functional genomics screens. *Nat Rev Genet* **23**, 89-103 (2022). <https://doi.org:10.1038/s41576-021-00409-w>
- 65 Tian, L., Chen, F. & Macosko, E. Z. The expanding vistas of spatial transcriptomics. *Nat Biotechnol* **41**, 773-782 (2023). <https://doi.org:10.1038/s41587-022-01448-2>
- 66 Kelly, R. T. Single-cell Proteomics: Progress and Prospects. *Mol Cell Proteomics* **19**, 1739-1748 (2020). <https://doi.org:10.1074/mcp.R120.002234>
- 67 Schubert, O. T., Rost, H. L., Collins, B. C., Rosenberger, G. & Aebersold, R. Quantitative proteomics: challenges and opportunities in basic and applied research. *Nat Protoc* **12**, 1289-1294 (2017). <https://doi.org:10.1038/nprot.2017.040>
- 68 Liu, X. & Locasale, J. W. Metabolomics: A Primer. *Trends Biochem Sci* **42**, 274-284 (2017). <https://doi.org:10.1016/j.tibs.2017.01.004>
- 69 Skalidis, I. *et al.* Cryo-EM and artificial intelligence visualize endogenous protein community members. *Structure* **30**, 575-589 e576 (2022). <https://doi.org:10.1016/j.str.2022.01.001>

- 70 Wu, X. & Rapoport, T. A. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc Natl Acad Sci U S A* **118** (2021). <https://doi.org:10.1073/pnas.2115001118>
- 71 Danev, R. & Baumeister, W. Expanding the boundaries of cryo-EM with phase plates. *Curr Opin Struct Biol* **46**, 87-94 (2017). <https://doi.org:10.1016/j.sbi.2017.06.006>
- 72 Wentinck, K., Gogou, C. & Meijer, D. H. Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins. *Curr Res Struct Biol* **4**, 332-337 (2022). <https://doi.org:10.1016/j.crstbi.2022.09.005>
- 73 Chemistry., T. N. P. i. *The Nobel Prize in Chemistry 2017 was awarded jointly to Jacques Dubochet, Joachim Frank and Richard Henderson "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution"*, <<https://www.nobelprize.org/prizes/chemistry/2017/summary/>> (2017).
- 74 McMullan, G., Faruqi, A. R. & Henderson, R. Direct Electron Detectors. *Methods Enzymol* **579**, 1-17 (2016). <https://doi.org:10.1016/bs.mie.2016.05.056>
- 75 Scheres, S. H. A Bayesian view on cryo-EM structure determination. *J Mol Biol* **415**, 406-418 (2012). <https://doi.org:10.1016/j.jmb.2011.11.010>
- 76 Rohou, A. & Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015). <https://doi.org:10.1016/j.jsb.2015.08.008>
- 77 Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12 (2016). <https://doi.org:10.1016/j.jsb.2015.11.003>
- 78 Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem J* **478**, 4169-4185 (2021). <https://doi.org:10.1042/BCJ20210708>
- 79 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017). <https://doi.org:10.1038/nmeth.4169>
- 80 Zhang, X. *et al.* An Atomic Structure of the Human Spliceosome. *Cell* **169**, 918-929 e914 (2017). <https://doi.org:10.1016/j.cell.2017.04.033>
- 81 Fica, S. M. & Nagai, K. Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. *Nat Struct Mol Biol* **24**, 791-799 (2017). <https://doi.org:10.1038/nsmb.3463>
- 82 Piper, S. J., Johnson, R. M., Wootten, D. & Sexton, P. M. Membranes under the Magnetic Lens: A Dive into the Diverse World of Membrane Protein Structures Using Cryo-EM. *Chem Rev* **122**, 13989-14017 (2022). <https://doi.org:10.1021/acs.chemrev.1c00837>
- 83 Croce, R. & van Amerongen, H. Light harvesting in oxygenic photosynthesis: Structural biology meets spectroscopy. *Science* **369** (2020). <https://doi.org:10.1126/science.aay2058>
- 84 Fromm, S. A. *et al.* The translating bacterial ribosome at 1.55 Å resolution generated by cryo-EM imaging services. *Nat Commun* **14**, 1095 (2023). <https://doi.org:10.1038/s41467-023-36742-3>
- 85 Drenth, J. *Principles of Protein X-Ray Crystallography*. (Springer New York, NY, 2007).
- 86 Garcia-Nafria, J. & Tate, C. G. Structure determination of GPCRs: cryo-EM compared with X-ray crystallography. *Biochem Soc Trans* **49**, 2345-2355 (2021). <https://doi.org:10.1042/BST20210431>

- 87 Bax, A. & Clore, G. M. Protein NMR: Boundless opportunities. *J Magn Reson* **306**, 187-191 (2019). <https://doi.org:10.1016/j.jmr.2019.07.037>
- 88 Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., Rance, M. & Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*. (Elsevier Inc, 2007).
- 89 Kwan, A. H., Mobli, M., Gooley, P. R., King, G. F. & Mackay, J. P. Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J* **278**, 687-703 (2011). <https://doi.org:10.1111/j.1742-4658.2011.08004.x>
- 90 Bieri, M. *et al.* Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination. *FEBS J* **278**, 704-715 (2011). <https://doi.org:10.1111/j.1742-4658.2011.08005.x>
- 91 Wagenknecht, T., Grassucci, R. & Frank, J. Electron microscopy and computer image averaging of ice-embedded large ribosomal subunits from *Escherichia coli*. *J Mol Biol* **199**, 137-147 (1988). [https://doi.org:10.1016/0022-2836\(88\)90384-1](https://doi.org:10.1016/0022-2836(88)90384-1)
- 92 Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* **21**, 129-228 (1988). <https://doi.org:10.1017/s0033583500004297>
- 93 Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438-449 (2015). <https://doi.org:10.1016/j.cell.2015.03.050>
- 94 Weissenberger, G., Henderikx, R. J. M. & Peters, P. J. Understanding the invisible hands of sample preparation for cryo-EM. *Nat Methods* **18**, 463-471 (2021). <https://doi.org:10.1038/s41592-021-01130-6>
- 95 Weber, M. S., Wojtynek, M. & Medalia, O. Cellular and Structural Studies of Eukaryotic Cells by Cryo-Electron Tomography. *Cells* **8** (2019). <https://doi.org:10.3390/cells8010057>
- 96 Turk, M. & Baumeister, W. The promise and the challenges of cryo-electron tomography. *FEBS Lett* **594**, 3243-3261 (2020). <https://doi.org:10.1002/1873-3468.13948>
- 97 Wu, M. & Lander, G. C. Present and Emerging Methodologies in Cryo-EM Single-Particle Analysis. *Biophys J* **119**, 1281-1289 (2020). <https://doi.org:10.1016/j.bpj.2020.08.027>
- 98 Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. *J Biol Chem* **294**, 5181-5197 (2019). <https://doi.org:10.1074/jbc.REV118.005602>
- 99 Casanal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci* **29**, 1069-1078 (2020). <https://doi.org:10.1002/pro.3791>
- 100 Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018). <https://doi.org:10.1107/S2059798318006551>
- 101 Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* **28**, 171-193 (1995). <https://doi.org:10.1017/s003358350000305x>
- 102 Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461 (2013). <https://doi.org:10.7554/eLife.00461>
- 103 Renaud, J. P. *et al.* Cryo-EM in drug discovery: achievements, limitations and prospects. *Nat Rev Drug Discov* **17**, 471-492 (2018). <https://doi.org:10.1038/nrd.2018.77>

- 104 Chari, A. & Stark, H. Prospects and Limitations of High-Resolution Single-Particle Cryo-Electron Microscopy. *Annu Rev Biophys* **52**, 391-411 (2023). <https://doi.org:10.1146/annurev-biophys-111622-091300>
- 105 Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat Methods* **11**, 63-65 (2014). <https://doi.org:10.1038/nmeth.2727>
- 106 Danev, R., Yanagisawa, H. & Kikkawa, M. Cryo-EM performance testing of hardware and data acquisition strategies. *Microscopy (Oxf)* **70**, 487-497 (2021). <https://doi.org:10.1093/jmicro/dfab016>
- 107 Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253-267 (2020). <https://doi.org:10.1107/S2052252520000081>
- 108 Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157-161 (2020). <https://doi.org:10.1038/s41586-020-2833-4>
- 109 Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152-156 (2020). <https://doi.org:10.1038/s41586-020-2829-0>
- 110 Sorzano, C. O. S. *et al.* On bias, variance, overfitting, gold standard and consensus in single-particle analysis by cryo-electron microscopy. *Acta Crystallogr D Struct Biol* **78**, 410-423 (2022). <https://doi.org:10.1107/S2059798322001978>
- 111 Cianfrocco, M. A. & Kellogg, E. H. What Could Go Wrong? A Practical Guide to Single-Particle Cryo-EM: From Biochemistry to Atomic Models. *J Chem Inf Model* **60**, 2458-2469 (2020). <https://doi.org:10.1021/acs.jcim.9b01178>
- 112 van Heel, M. & Schatz, M. Fourier shell correlation threshold criteria. *J Struct Biol* **151**, 250-262 (2005). <https://doi.org:10.1016/j.jsb.2005.05.009>
- 113 Sheldrick, G. Phase annealing in SHELX-90: direct methods for larger structures. *Acta Crystallographica Section A* **46**, 467-473 (1990). <https://doi.org:doi:10.1107/S0108767390000277>
- 114 Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **333**, 721-745 (2003). <https://doi.org:10.1016/j.jmb.2003.07.013>
- 115 D'Imprima, E. *et al.* Protein denaturation at the air-water interface and how to prevent it. *Elife* **8** (2019). <https://doi.org:10.7554/eLife.42747>
- 116 Ho, C. M. *et al.* Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nat Methods* **17**, 79-85 (2020). <https://doi.org:10.1038/s41592-019-0637-y>
- 117 Su, C. C. *et al.* A 'Build and Retrieve' methodology to simultaneously solve cryo-EM structures of membrane proteins. *Nat Methods* **18**, 69-75 (2021). <https://doi.org:10.1038/s41592-020-01021-2>
- 118 Su, C. C. *et al.* High-resolution structural-omics of human liver enzymes. *Cell Rep* **42**, 112609 (2023). <https://doi.org:10.1016/j.celrep.2023.112609>
- 119 Gelderblom, H. R. & Krüger, D. H. in *Advances in Imaging and Electron Physics* Vol. 182 (ed Peter W. Hawkes) 1-94 (Elsevier, 2014).
- 120 von Borries, B., Ruska, E. & Ruska, H. Bakterien und Virus in Übermikroskopischer Aufnahme. *Klinische Wochenschrift* **17**, 921-925 (1938). <https://doi.org:10.1007/BF01775798>

- 121 Ruska, H., Borries, B. v. & Ruska, E. Die Bedeutung der Übermikroskopie für die Virusforschung. *Archiv für die gesamte Virusforschung* **1**, 155-169 (1939). <https://doi.org:10.1007/BF01243399>
- 122 Kausche, G. A. & Ruska, H. Die Sichtbarmachung der Adsorption von Metallkolloiden an Eiweißkörper. *Kolloid-Zeitschrift* **89**, 21-26 (1939). <https://doi.org:10.1007/BF01518802>
- 123 De Rosier, D. J. & Klug, A. Reconstruction of three dimensional structures from electron micrographs. *Nature* **217**, 130-134 (1968). <https://doi.org:10.1038/217130a0>
- 124 Adrian, M., Dubochet, J., Lepault, J. & McDowell, A. W. Cryo-electron microscopy of viruses. *Nature* **308**, 32-36 (1984). <https://doi.org:10.1038/308032a0>
- 125 Mancini, E. J., Clarke, M., Gowen, B. E., Rutten, T. & Fuller, S. D. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol Cell* **5**, 255-266 (2000). [https://doi.org:10.1016/s1097-2765\(00\)80421-9](https://doi.org:10.1016/s1097-2765(00)80421-9)
- 126 Sachse, C. *et al.* High-resolution electron microscopy of helical specimens: a fresh look at tobacco mosaic virus. *J Mol Biol* **371**, 812-835 (2007). <https://doi.org:10.1016/j.jmb.2007.05.088>
- 127 Enard, D., Cai, L., Gwennap, C. & Petrov, D. A. Viruses are a dominant driver of protein adaptation in mammals. *Elife* **5** (2016). <https://doi.org:10.7554/eLife.12469>
- 128 Caspar, D. L. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* **27**, 1-24 (1962). <https://doi.org:10.1101/sqb.1962.027.001.005>
- 129 Sevana, M., Klose, T. & Rossmann, M. G. in *Encyclopaedia of Virology* Vol. 1 (eds D. H. Bamford & M. Zuckerman) 257-277 (Elsevier Inc, 2021).
- 130 Ryu, W.-S. in *Molecular Virology of Human Pathogenic Viruses* (ed Wang-Shick Ryu) 31-45 (Elsevier Inc. , 2017).
- 131 Fernández de Castro, I., Tenorio, R. & Risco, C. Virus Factories. *Encyclopedia of Virology*, 495-500 (2021).
- 132 Barber, G. N. Host defense, viruses and apoptosis. *Cell Death Differ* **8**, 113-126 (2001). <https://doi.org:10.1038/sj.cdd.4400823>
- 133 Feschotte, C. & Gilbert, C. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat Rev Genet* **13**, 283-296 (2012). <https://doi.org:10.1038/nrg3199>
- 134 Roossinck, M. J. The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol* **9**, 99-108 (2011). <https://doi.org:10.1038/nrmicro2491>
- 135 Louten, J. Virus Structure and Classification. *Essential Human Virology*, 19-29 (2016).
- 136 Ahlquist, P. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nat Rev Microbiol* **4**, 371-382 (2006). <https://doi.org:10.1038/nrmicro1389>
- 137 Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **16**, 18-29 (2015). <https://doi.org:10.1038/nrm3920>
- 138 Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197-208 (2005). <https://doi.org:10.1038/nrm1589>
- 139 Mishra, P. M., Verma, N. C., Rao, C., Uversky, V. N. & Nandi, C. K. Intrinsically disordered proteins of viruses: Involvement in the mechanism of cell regulation and pathogenesis. *Prog Mol Biol Transl Sci* **174**, 1-78 (2020). <https://doi.org:10.1016/bs.pmbts.2020.03.001>

- 140 Xue, B. *et al.* Structural disorder in viral proteins. *Chem Rev* **114**, 6880-6911 (2014).  
<https://doi.org:10.1021/cr4005692>
- 141 Goh, G. K., Dunker, A. K. & Uversky, V. N. Protein intrinsic disorder toolbox for comparative analysis of viral proteins. *BMC Genomics* **9 Suppl 2**, S4 (2008).  
<https://doi.org:10.1186/1471-2164-9-S2-S4>
- 142 Goh, G. K., Dunker, A. K. & Uversky, V. N. A comparative analysis of viral matrix proteins using disorder predictors. *Virology* **5**, 126 (2008).  
<https://doi.org:10.1186/1743-422X-5-126>
- 143 Dunker, A. K., Silman, I., Uversky, V. N. & Sussman, J. L. Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* **18**, 756-764 (2008).  
<https://doi.org:10.1016/j.sbi.2008.10.002>
- 144 Uversky, V. N. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* **1834**, 932-951 (2013). <https://doi.org:10.1016/j.bbapap.2012.12.008>
- 145 Xue, B., Dunker, A. K. & Uversky, V. N. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J Biomol Struct Dyn* **30**, 137-149 (2012).  
<https://doi.org:10.1080/07391102.2012.675145>
- 146 Dinman, J. D. & Wickner, R. B. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J Virol* **66**, 3669-3676 (1992). <https://doi.org:10.1128/JVI.66.6.3669-3676.1992>
- 147 Kristin L. DeBord, D. M. A., Melanie M. Marketon, Katie A. Overheim, R. William DePaolo, Nancy A. Ciletti, Bana Jabri, and Olaf Schneewind. Immunogenicity and Protective Immunity against Bubonic Plague and Pneumonic Plague by Immunization of Mice with the Recombinant V10 Antigen, a Variant of LcrV *Infect Immun* **74** (2006). <https://doi.org:10.1128/IAI.01860-05>
- 148 McBride, R. C., Boucher, N., Park, D. S., Turner, P. E. & Townsend, J. P. Yeast response to LA virus indicates coadapted global gene expression during mycoviral infection. *FEMS Yeast Res* **13**, 162-179 (2013). <https://doi.org:10.1111/1567-1364.12019>
- 149 Mertens, P. The dsRNA viruses. *Virus Research* **101**, 3-13 (2004).  
<https://doi.org:https://doi.org/10.1016/j.virusres.2003.12.002>
- 150 Du, Y. *et al.* Global burden and trends of rotavirus infection-associated deaths from 1990 to 2019: an observational trend study. *Virology* **19**, 166 (2022).  
<https://doi.org:10.1186/s12985-022-01898-9>
- 151 Schwartz-Cornil, I. *et al.* Bluetongue virus: virology, pathogenesis and immunity. *Vet Res* **39**, 46 (2008). <https://doi.org:10.1051/vetres:2008023>
- 152 Dodds, J. A., Morris, T. J. & Jordan, R. L. Plant Viral Double-Stranded RNA. *Annual Review of Phytopathology* **22**, 151-168 (1984).  
<https://doi.org:10.1146/annurev.py.22.090184.001055>
- 153 Drolet, B. S., McGregor, B. L., Cohnstaedt, L. W., Wilson, W. C. & McVey, D. S. in *Veterinary Microbiology* 679-692 (2022).
- 154 Decroly, E., Ferron, F., Lescar, J. & Canard, B. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* **10**, 51-65 (2011).  
<https://doi.org:10.1038/nrmicro2675>
- 155 Chen, Y. G. & Hur, S. Cellular origins of dsRNA, their recognition and consequences. *Nat Rev Mol Cell Biol* **23**, 286-301 (2022). <https://doi.org:10.1038/s41580-021-00430-1>

- 156 Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *J Gen Virol* **93**, 1385-1409 (2012). <https://doi.org:10.1099/vir.0.042499-0>
- 157 Somers, J. M. & Bevan, E. A. The inheritance of the killer character in yeast. *Genet Res* **13**, 71-83 (1969). <https://doi.org:10.1017/s0016672300002743>
- 158 Breinig, F., Sendzik, T., Einfeld, K. & Schmitt, M. J. Dissecting toxin immunity in virus-infected killer yeast uncovers an intrinsic strategy of self-protection. *Proc Natl Acad Sci U S A* **103**, 3810-3815 (2006). <https://doi.org:10.1073/pnas.0510070103>
- 159 Schmitt, M. J. & Breinig, F. Yeast viral killer toxins: lethality and self-protection. *Nat Rev Microbiol* **4**, 212-221 (2006). <https://doi.org:10.1038/nrmicro1347>
- 160 Magliani, W., Conti, S., Gerloni, M., Bertolotti, D. & Polonelli, L. Yeast killer systems. *Clin Microbiol Rev* **10**, 369-400 (1997). <https://doi.org:10.1128/CMR.10.3.369>
- 161 Caston, J. R. *et al.* Structure of L-A virus: a specialized compartment for the transcription and replication of double-stranded RNA. *J Cell Biol* **138**, 975-985 (1997). <https://doi.org:10.1083/jcb.138.5.975>
- 162 Young, T. W. & Yagiu, M. A comparison of the killer character in different yeasts and its classification. *Antonie Van Leeuwenhoek* **44**, 59-77 (1978). <https://doi.org:10.1007/BF00400077>
- 163 Wickner, R. B., Fujimura, T. & Esteban, R. Viruses and prions of *Saccharomyces cerevisiae*. *Adv Virus Res* **86**, 1-36 (2013). <https://doi.org:10.1016/B978-0-12-394315-6.00001-5>
- 164 Schmitt, M. J. & Breinig, F. The viral killer system in yeast: from molecular biology to application. *FEMS Microbiol Rev* **26**, 257-276 (2002). <https://doi.org:10.1111/j.1574-6976.2002.tb00614.x>
- 165 Ribas, J. C. & Wickner, R. B. The Gag domain of the Gag-Pol fusion protein directs incorporation into the L-A double-stranded RNA viral particles in *Saccharomyces cerevisiae*. *J Biol Chem* **273**, 9306-9311 (1998). <https://doi.org:10.1074/jbc.273.15.9306>
- 166 Liu, G. L. *et al.* Yeast killer toxins, molecular mechanisms of their action and their applications. *Crit Rev Biotechnol* **35**, 222-234 (2015). <https://doi.org:10.3109/07388551.2013.833582>
- 167 Bourbonnais, Y., Germain, D., Latchinian-Sadek, L., Boileau, G. & Thomas, D. Y. Prohormone processing by yeast proteases. *Enzyme* **45**, 244-256 (1991). <https://doi.org:10.1159/000468899>
- 168 Martinac, B. *et al.* Yeast K1 killer toxin forms ion channels in sensitive yeast spheroplasts and in artificial liposomes. *Proc Natl Acad Sci U S A* **87**, 6228-6232 (1990). <https://doi.org:10.1073/pnas.87.16.6228>
- 169 Carroll, S. Y. *et al.* A yeast killer toxin screen provides insights into a/b toxin entry, trafficking, and killing mechanisms. *Dev Cell* **17**, 552-560 (2009). <https://doi.org:10.1016/j.devcel.2009.08.006>
- 170 Hutchins, K. & Bussey, H. Cell wall receptor for yeast killer toxin: involvement of (1 leads to 6)-beta-D-glucan. *J Bacteriol* **154**, 161-169 (1983). <https://doi.org:10.1128/jb.154.1.161-169.1983>
- 171 Sturley, S. L., Elliot, Q., LeVitre, J., Tipper, D. J. & Bostian, K. A. Mapping of functional domains within the *Saccharomyces cerevisiae* type 1 killer preprotoxin. *EMBO J* **5**, 3381-3389 (1986). <https://doi.org:10.1002/j.1460-2075.1986.tb04654.x>

- 172 Brown, J. T., Bai, X. & Johnson, A. W. The yeast antiviral proteins Ski2p, Ski3p, and  
Ski8p exist as a complex in vivo. *RNA* **6**, 449-457 (2000).  
<https://doi.org/10.1017/s1355838200991787>
- 173 Schmidt, C. *et al.* The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase  
complex. *Science* **354**, 1431-1433 (2016). <https://doi.org/10.1126/science.aaf7520>
- 174 Halbach, F., Reichelt, P., Rode, M. & Conti, E. The yeast ski complex: crystal structure  
and RNA channeling to the exosome complex. *Cell* **154**, 814-826 (2013).  
<https://doi.org/10.1016/j.cell.2013.07.017>
- 175 Tomecki, R., Drzkwowska, K., Kobylecki, K. & Tudek, A. SKI complex: A multifaceted  
cytoplasmic RNA exosome cofactor in mRNA metabolism with links to disease,  
developmental processes, and antiviral responses. *Wiley Interdiscip Rev RNA*, e1795  
(2023). <https://doi.org/10.1002/wrna.1795>
- 176 Marquina, D., Santos, A. & Peinado, J. M. Biology of killer yeasts. *Int Microbiol* **5**, 65-  
71 (2002). <https://doi.org/10.1007/s10123-002-0066-z>
- 177 Steensels, J. *et al.* Improving industrial yeast strains: exploiting natural and artificial  
diversity. *FEMS Microbiol Rev* **38**, 947-995 (2014). [https://doi.org/10.1111/1574-  
6976.12073](https://doi.org/10.1111/1574-6976.12073)
- 178 in *Virus Taxonomy* (eds Andrew M. Q. King, Michael J. Adams, Eric B. Carstens, &  
Elliot J. Lefkowitz) 639-650 (Elsevier, 2012).
- 179 Poulos, B. T., Tang, K. F. J., Pantoja, C. R., Bonami, J. R. & Lightner, D. V. Purification  
and characterization of infectious myonecrosis virus of penaeid shrimp. *J Gen Virol*  
**87**, 987-996 (2006). <https://doi.org/10.1099/vir.0.81127-0>
- 180 Lovoll, M. *et al.* A novel totivirus and piscine reovirus (PRV) in Atlantic salmon (*Salmo*  
*salar*) with cardiomyopathy syndrome (CMS). *Virology* **7**, 309 (2010).  
<https://doi.org/10.1186/1743-422X-7-309>
- 181 Unterholzner, S. J., Poppenberger, B. & Rozhon, W. Toxin-antitoxin systems: Biology,  
identification, and application. *Mob Genet Elements* **3**, e26219 (2013).  
<https://doi.org/10.4161/mge.26219>
- 182 Rodriguez-Cousino, N. & Esteban, R. Relationships and Evolution of Double-Stranded  
RNA Totiviruses of Yeasts Inferred from Analysis of L-A-2 and L-BC Variants in Wine  
Yeast Strain Populations. *Appl Environ Microbiol* **83** (2017).  
<https://doi.org/10.1128/AEM.02991-16>
- 183 Naitow, H., Tang, J., Canady, M., Wickner, R. B. & Johnson, J. E. L-A virus at 3.4 Å  
resolution reveals particle architecture and mRNA decapping mechanism. *Nat Struct  
Biol* **9**, 725-728 (2002). <https://doi.org/10.1038/nsb844>
- 184 Janssen, M. E. *et al.* Three-dimensional structure of a protozoal double-stranded  
RNA virus that infects the enteric pathogen *Giardia lamblia*. *J Virol* **89**, 1182-1194  
(2015). <https://doi.org/10.1128/JVI.02745-14>
- 185 Stevens, A., Muratore, K., Cui, Y., Johnson, P. J. & Zhou, Z. H. Atomic Structure of the  
*Trichomonas vaginalis* Double-Stranded RNA Virus 2. *mBio* **12** (2021).  
<https://doi.org/10.1128/mBio.02924-20>
- 186 Tang, J. *et al.* Infectious myonecrosis virus has a totivirus-like, 120-subunit capsid,  
but with fiber complexes at the fivefold axes. *Proc Natl Acad Sci U S A* **105**, 17526-  
17531 (2008). <https://doi.org/10.1073/pnas.0806724105>
- 187 Dunn, S. E. *et al.* Three-dimensional structure of victorivirus HvV190S suggests coat  
proteins in most totiviruses share a conserved core. *PLoS Pathog* **9**, e1003225  
(2013). <https://doi.org/10.1371/journal.ppat.1003225>

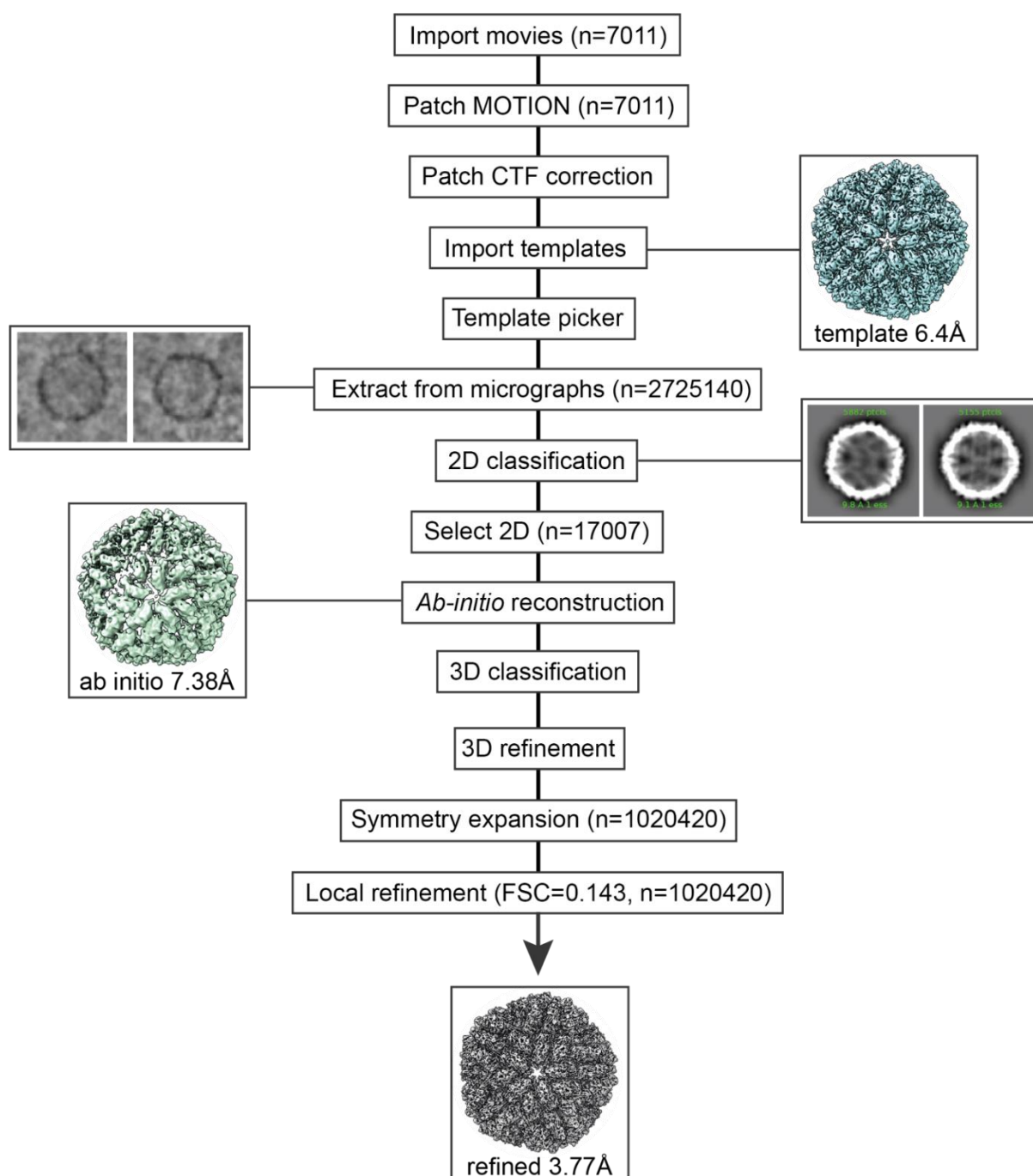


- 188 Parent, K. N. *et al.* Structure of a protozoan virus from the human genitourinary parasite *Trichomonas vaginalis*. *mBio* **4** (2013). <https://doi.org/10.1128/mBio.00056-13>
- 189 Prochazkova, M. *et al.* Capsid Structure of Leishmania RNA Virus 1. *J Virol* **95** (2021). <https://doi.org/10.1128/JVI.01957-20>
- 190 Okamoto, K. *et al.* The infectious particle of insect-borne totivirus-like Omono River virus has raised ridges and lacks fibre complexes. *Sci Rep* **6**, 33170 (2016). <https://doi.org/10.1038/srep33170>
- 191 Shao, Q. *et al.* Cryo-EM reveals a previously unrecognized structural protein of a dsRNA virus implicated in its extracellular transmission. *PLoS Pathog* **17**, e1009396 (2021). <https://doi.org/10.1371/journal.ppat.1009396>
- 192 Grybchuk, D. *et al.* Structures of L-BC virus and its open particle provide insight into Totivirus capsid assembly. *Commun Biol* **5**, 847 (2022). <https://doi.org/10.1038/s42003-022-03793-z>
- 193 Prasad, B. V. & Schmid, M. F. Principles of virus structural organization. *Adv Exp Med Biol* **726**, 17-47 (2012). [https://doi.org/10.1007/978-1-4614-0980-9\\_3](https://doi.org/10.1007/978-1-4614-0980-9_3)
- 194 Cheng, R. H. *et al.* Fungal virus capsids, cytoplasmic compartments for the replication of double-stranded RNA, formed as icosahedral shells of asymmetric Gag dimers. *J Mol Biol* **244**, 255-258 (1994). <https://doi.org/10.1006/jmbi.1994.1726>
- 195 Rowley, P. A., Ho, B., Bushong, S., Johnson, A. & Sawyer, S. L. XRN1 Is a Species-Specific Virus Restriction Factor in Yeasts. *PLoS Pathog* **12**, e1005890 (2016). <https://doi.org/10.1371/journal.ppat.1005890>
- 196 Fujimura, T. & Esteban, R. Cap-snatching mechanism in yeast L-A double-stranded RNA virus. *Proc Natl Acad Sci U S A* **108**, 17667-17671 (2011). <https://doi.org/10.1073/pnas.1111900108>
- 197 Chu, C. *et al.* Structure of the guanylyltransferase domain of human mRNA capping enzyme. *Proc Natl Acad Sci U S A* **108**, 10104-10108 (2011). <https://doi.org/10.1073/pnas.1106610108>
- 198 He, Y. *et al.* In situ structures of RNA-dependent RNA polymerase inside bluetongue virus before and after uncoating. *Proc Natl Acad Sci U S A* **116**, 16535-16540 (2019). <https://doi.org/10.1073/pnas.1905849116>
- 199 Fujimura, T., Esteban, R. & Wickner, R. B. In vitro L-A double-stranded RNA synthesis in virus-like particles from *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **83**, 4433-4437 (1986). <https://doi.org/10.1073/pnas.83.12.4433>
- 200 Fujimura, T. & Esteban, R. Cap snatching in yeast L-BC double-stranded RNA totivirus. *J Biol Chem* **288**, 23716-23724 (2013). <https://doi.org/10.1074/jbc.M113.490953>
- 201 Mertens, P. P. *et al.* Bluetongue virus replication, molecular and structural biology. *Vet Ital* **40**, 426-437 (2004).
- 202 de la Rosa-Trevin, J. M. *et al.* Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol* **195**, 93-99 (2016). <https://doi.org/10.1016/j.jsb.2016.04.010>
- 203 Sigworth, F. J. Principles of cryo-EM single-particle image processing. *Microscopy (Oxf)* **65**, 57-67 (2016). <https://doi.org/10.1093/jmicro/dfv370>
- 204 Rubinstein, J. L. & Brubaker, M. A. Alignment of cryo-EM movies of individual particles by optimization of image translations. *J Struct Biol* **192**, 188-195 (2015). <https://doi.org/10.1016/j.jsb.2015.08.007>

- 205 Scheres, S. H. Beam-induced motion correction for sub-megadalton cryo-EM  
particles. *Elife* **3**, e03665 (2014). <https://doi.org:10.7554/eLife.03665>
- 206 Wu, J. G. *et al.* Machine Learning for Structure Determination in Single-Particle Cryo-  
Electron Microscopy: A Systematic Review. *IEEE Trans Neural Netw Learn Syst* **33**,  
452-472 (2022). <https://doi.org:10.1109/TNNLS.2021.3131325>
- 207 de la Rosa-Trevin, J. M. *et al.* Xmipp 3.0: an improved software suite for image  
processing in electron microscopy. *J Struct Biol* **184**, 321-328 (2013).  
<https://doi.org:10.1016/j.jsb.2013.09.015>
- 208 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure  
determination in RELION-3. *Elife* **7** (2018). <https://doi.org:10.7554/eLife.42166>
- 209 Sharov, G., Morado, D. R., Carroni, M. & de la Rosa-Trevin, J. M. Using RELION  
software within the Scipion framework. *Acta Crystallogr D Struct Biol* **77**, 403-410  
(2021). <https://doi.org:10.1107/S2059798321001856>
- 210 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta  
Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).  
<https://doi.org:10.1107/S0907444904019158>
- 211 Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res* **50**,  
W210-W215 (2022). <https://doi.org:10.1093/nar/gkac387>
- 212 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research  
and analysis. *J Comput Chem* **25**, 1605-1612 (2004).  
<https://doi.org:10.1002/jcc.20084>
- 213 Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers,  
educators, and developers. *Protein Sci* **30**, 70-82 (2021).  
<https://doi.org:10.1002/pro.3943>
- 214 Gotze, M., Iacobucci, C., Ihling, C. H. & Sinz, A. A Simple Cross-Linking/Mass  
Spectrometry Workflow for Studying System-wide Protein Interactions. *Anal Chem*  
**91**, 10236-10244 (2019). <https://doi.org:10.1021/acs.analchem.9b02372>
- 215 Iacobucci, C. *et al.* A cross-linking/mass spectrometry workflow based on MS-  
cleavable cross-linkers and the MeroX software for studying protein structures and  
protein-protein interactions. *Nat Protoc* **13**, 2864-2889 (2018).  
<https://doi.org:10.1038/s41596-018-0068-8>
- 216 Rueden, C. T. *et al.* ImageJ2: ImageJ for the next generation of scientific image data.  
*BMC Bioinformatics* **18**, 529 (2017). <https://doi.org:10.1186/s12859-017-1934-z>
- 217 Gregg, C., Kyryakov, P. & Titorenko, V. I. Purification of mitochondria from yeast  
cells. *J Vis Exp* (2009). <https://doi.org:10.3791/1417>
- 218 Blum, T. B., Hahn, A., Meier, T., Davies, K. M. & Kuhlbrandt, W. Dimers of  
mitochondrial ATP synthase induce membrane curvature and self-assemble into  
rows. *Proc Natl Acad Sci U S A* **116**, 4250-4255 (2019).  
<https://doi.org:10.1073/pnas.1816556116>
- 219 Schmidt, L. *et al.* Delineating organizational principles of the endogenous L-A virus by  
cryo-EM and computational analysis of native cell extracts. *bioRxiv*,  
2022.2007.2015.498668 (2022). <https://doi.org:10.1101/2022.07.15.498668>
- 220 Zhou, Y. *et al.* Structural impact of K63 ubiquitin on yeast translocating ribosomes  
under oxidative stress. *Proc Natl Acad Sci U S A* **117**, 22157-22166 (2020).  
<https://doi.org:10.1073/pnas.2005301117>
- 221 Holm, L. DALI and the persistence of protein shape. *Protein Sci* **29**, 128-140 (2020).  
<https://doi.org:10.1002/pro.3749>

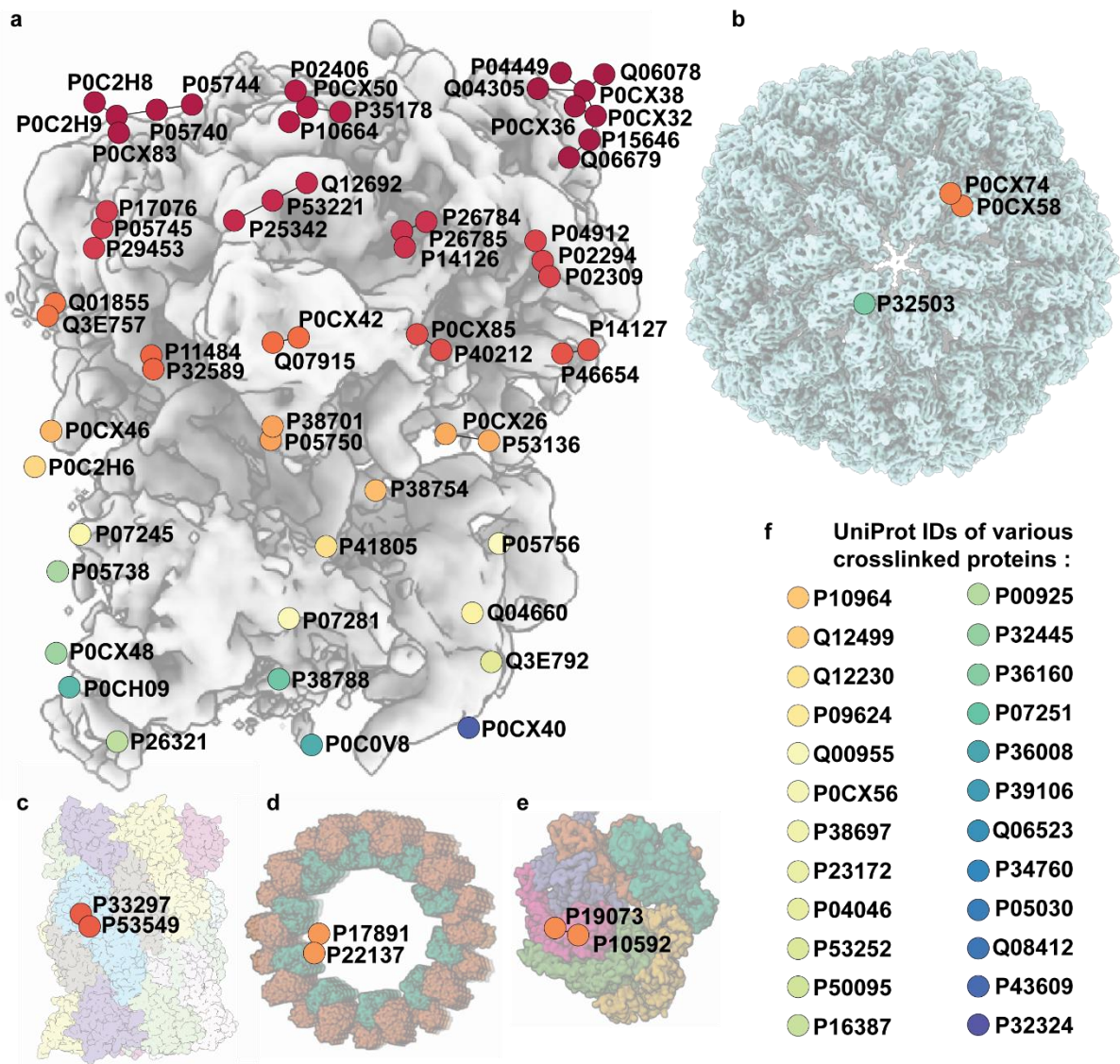
- 222 Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-D361 (2017). <https://doi.org:10.1093/nar/gkw1092>
- 223 Gallivan, J. P. & Dougherty, D. A. Cation- $\pi$  interactions in structural biology. *Proc Natl Acad Sci U S A* **96**, 9459-9464 (1999). <https://doi.org:10.1073/pnas.96.17.9459>
- 224 Huet, A., Duda, R. L., Boulanger, P. & Conway, J. F. Capsid expansion of bacteriophage T5 revealed by high resolution cryoelectron microscopy. *Proc Natl Acad Sci U S A* **116**, 21037-21046 (2019). <https://doi.org:10.1073/pnas.1909645116>
- 225 Matsuo, E., Yamazaki, K., Tsuruta, H. & Roy, P. Interaction between a Unique Minor Protein and a Major Capsid Protein of Bluetongue Virus Controls Virus Infectivity. *J Virol* **92** (2018). <https://doi.org:10.1128/JVI.01784-17>
- 226 van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modelling of Biomolecular Complexes. *J Mol Biol* **428**, 720-725 (2016). <https://doi.org:10.1016/j.jmb.2015.09.014>
- 227 Papadakos, G., Wojdyla, J. A. & Kleanthous, C. Nuclease colicins and their immunity proteins. *Q Rev Biophys* **45**, 57-103 (2012). <https://doi.org:10.1017/S0033583511000114>
- 228 Tesina, P. *et al.* Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts. *EMBO J* **39**, e103365 (2020). <https://doi.org:10.15252/emj.2019103365>
- 229 Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J Struct Biol* **213**, 107702 (2021). <https://doi.org:10.1016/j.jsb.2021.107702>
- 230 Behrmann, E. *et al.* Structural snapshots of actively translating human ribosomes. *Cell* **161**, 845-857 (2015). <https://doi.org:10.1016/j.cell.2015.03.052>
- 231 Bamford, D. H. & Wickner, R. B. Assembly of double-stranded RNA viruses: bacteriophage  $\phi$ 6 and yeast virus L-A. *Seminars in Virology* **5**, 61-69 (1994). <https://doi.org:https://doi.org/10.1006/smvv.1994.1007>
- 232 Poranen, M. M. & Bamford, D. H. Assembly of large icosahedral double-stranded RNA viruses. *Adv Exp Med Biol* **726**, 379-402 (2012). [https://doi.org:10.1007/978-1-4614-0980-9\\_17](https://doi.org:10.1007/978-1-4614-0980-9_17)
- 233 Kresze, G. B. & Ronft, H. Pyruvate dehydrogenase complex from baker's yeast. 1. Purification and some kinetic and regulatory properties. *Eur J Biochem* **119**, 573-579 (1981). <https://doi.org:10.1111/j.1432-1033.1981.tb05646.x>
- 234 Stoops, J. K. *et al.* On the unique structural organization of the *Saccharomyces cerevisiae* pyruvate dehydrogenase complex. *J Biol Chem* **272**, 5757-5764 (1997). <https://doi.org:10.1074/jbc.272.9.5757>
- 235 Baker, R. E. *et al.* Infectious disease in an era of global change. *Nat Rev Microbiol* **20**, 193-205 (2022). <https://doi.org:10.1038/s41579-021-00639-z>
- 236 Aboul-ela, F., Karn, J. & Varani, G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol* **253**, 313-332 (1995). <https://doi.org:10.1006/jmbi.1995.0555>

## 9. Appendix



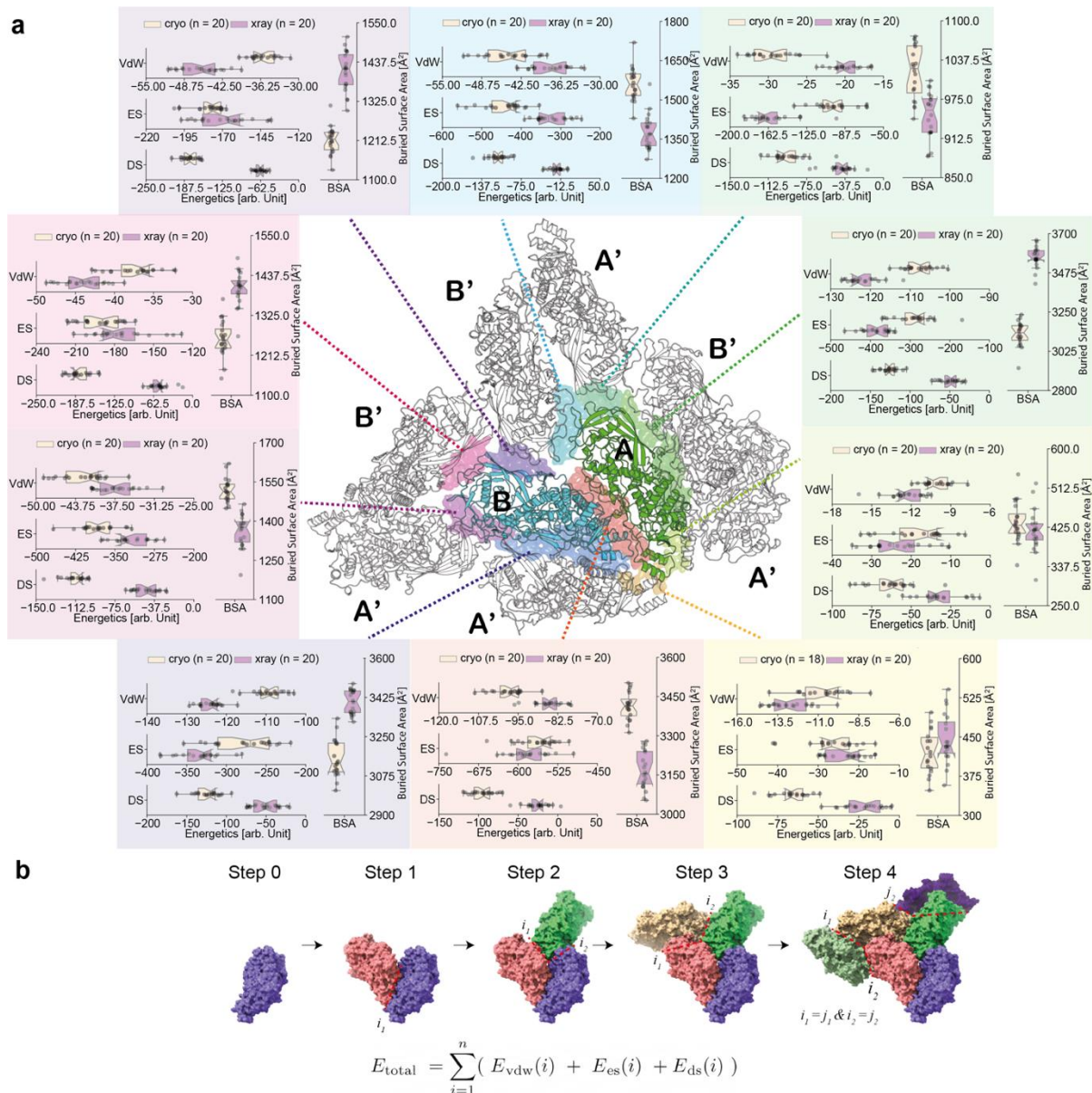
### Supplementary Figure 1. Detailed image analysis workflow.

Shown is a workflow of image analysis, specifically followed for the high-resolution cryo-EM structure of the L-A virus. Data is imported into the workspace, followed by motion correction and CTF correction procedures. After these processes finished, particles can be picked, and a de novo lower resolution reconstruction is derived after 3D classification and refinement. For template picking this reconstruction was chosen. The picked particles are then extracted and 2D classified. Selected 2D class averages were used for another ab initio reconstruction, followed by 3D classification and 3D refinement of chosen particles. The map quality was improved after using symmetry expansion and particles were, lastly, locally refined.



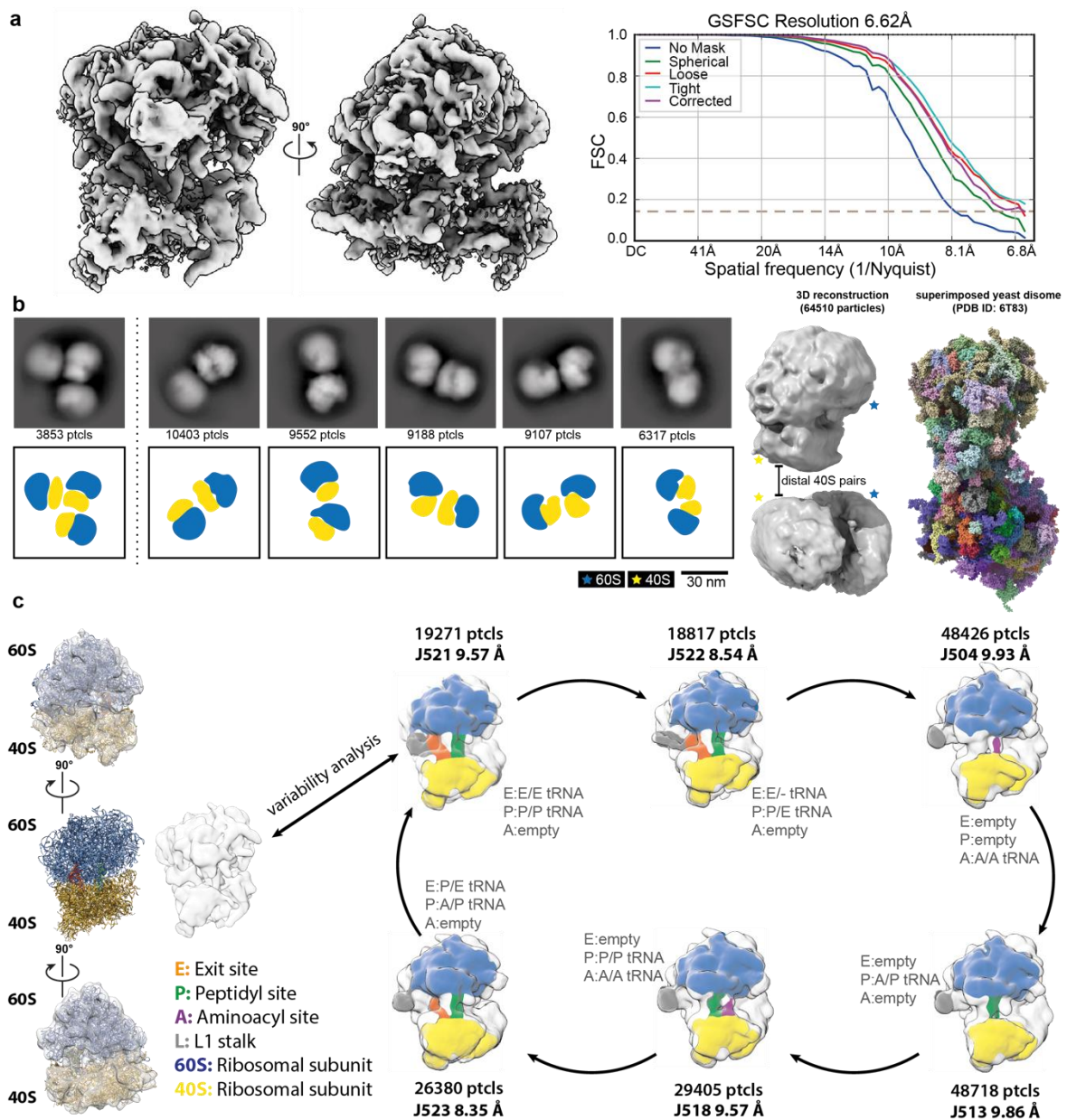
**Supplementary Figure 2. XL-MS of fractions 5 and 6 identified enzymatic pathways and protein complexes.**

Most inter and intra molecular crosslinks were found in ribosomes and between ribosomes and other translation related complexes (see uniprot IDs of involved proteins in **a**; *map created for this thesis*). For the yeast L-A helper virus intramolecular crosslinks (UniProt ID P32503) could be identified as well as between the Ty1-PR1 Gag and Ty1-JR1 Gag transposon (**b**; *map of the L-A helper virus created for this thesis*). The subunits RPT4 and RPT5 of the proteasome could be found crosslinked (**c**; *PDB ID 3DY4*). Clathrin heavy chain (UniProt ID P22137) and light chain (UniProt ID P17891) showed crosslinks (**d**; *PDB ID 5AHV*). (**e**) is showing crosslinks between the cell division control protein 42 and heat shock protein SSA2 (*PDB ID 6AHF*). Next to inter molecular crosslinks also intra molecular (homo-inter) crosslinks can be found (see **f** UniProt IDs of various crosslinked proteins).



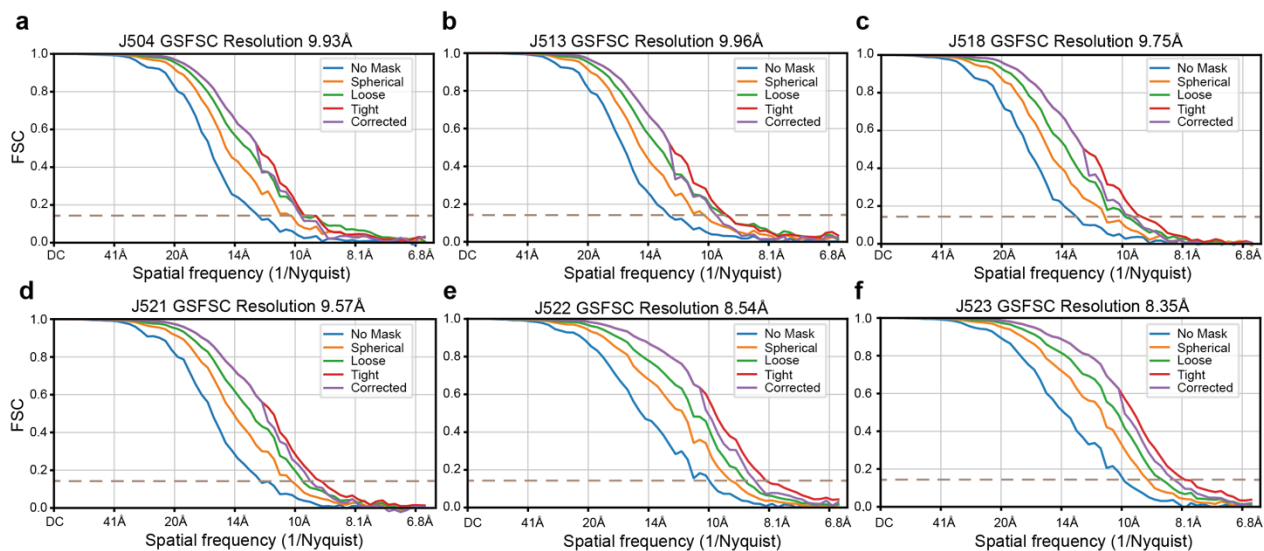
**Supplementary Figure 3. Comparison of interface energetics forming the crystallographically-resolved and cryo-EM resolved L-A virus capsid.**

(a) Shown in peach (chain A) and green (chain B) in comic representation is the yeast L-A virus capsomer with adjacent protomers. Subunits corresponding to chain A are labelled with A' and subunits corresponding to B are labelled B'. The colors for the calculated interfaces match the box-plot background color. Calculated were van der Waals forces (VdW), desolvation scores (DE), electrostatics scores (ES) in arbitrary units (a.u.), and the buried surface area (BSA) in Å<sup>2</sup>. (b) Proposed stability analysis (from monomer until step 4) of the capsid informed by the energetics calculations for derived interfaces.



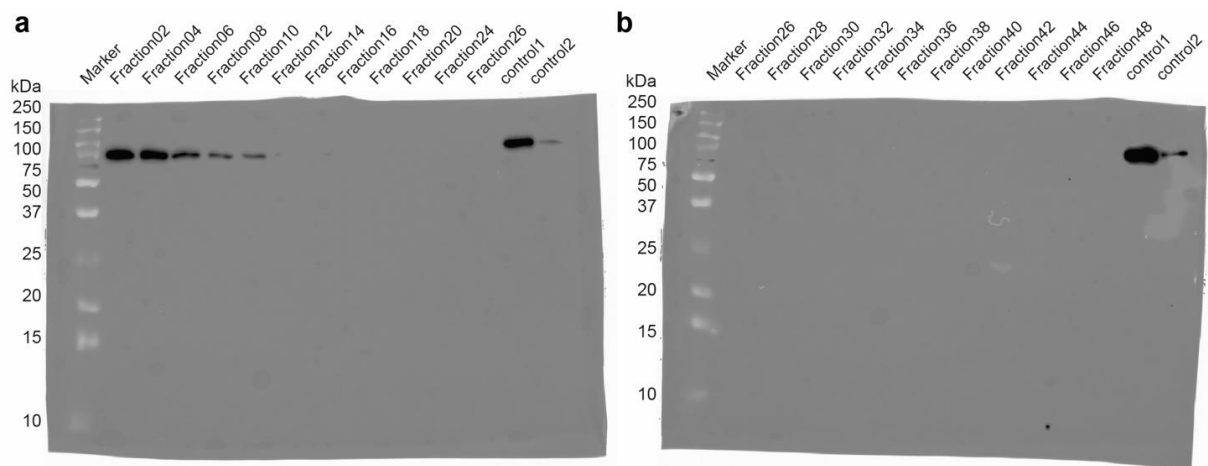
**Supplementary Figure 4. Cryo-EM reconstructions of translationally active ribosomes.**

(a) 3D reconstructed ribosome at 6.62 Å with an FSC=0.143. (b) 2D class averages and 3D reconstruction of identified polysomes within the native cell extract. Scale bar represents 30 nm. (c) Identified distinct translational states after variability analysis performed in cryoSPARC.



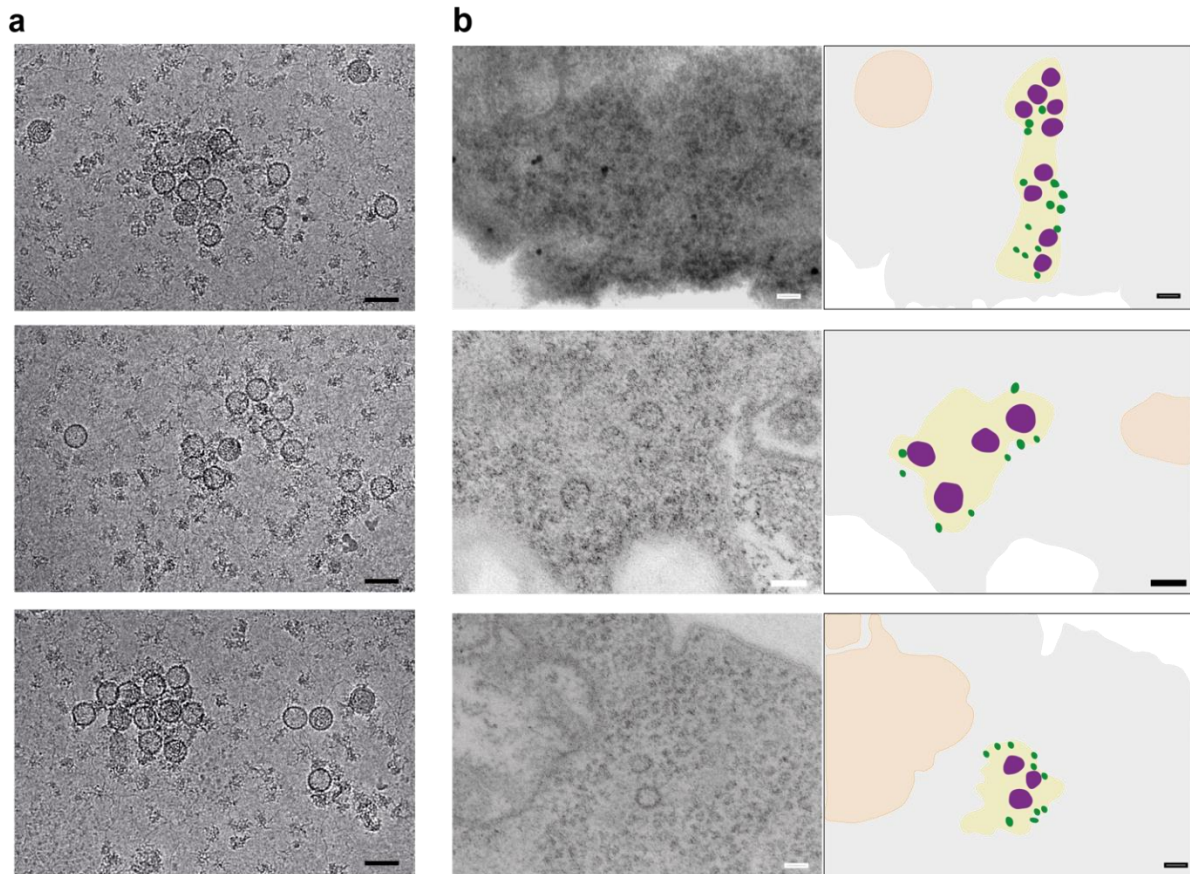
**Supplementary Figure 5.** (a-f) FSCs of the 6 distinct translational stages of reconstructed ribosomes after variability analysis shown in **Supplement Figure 4**.





**Supplementary Figure 6. Tracking of the abundance of the L-A virus via Western blot analysis.**

(a, b) Original uncropped membranes after Western blot analysis. Every second fraction was applied on the gel and L-A virus capsid was detected. Control 1 and 2 are positive controls and correspond to the recombinant protein and the fraction in which the LA virus was detected by mass spectrometry (1), and the yeast lysate used for injecting the SEC (2).



**Supplementary Figure 7. L-A virus communities.**

(a) More examples of viral particles grouping in the micrographs. Cryo-substituted, in resin embedded *Saccharomyces cerevisiae* sections of 30 nm in width and imaged on the EM 900. Shown are crops through yeast on the left and a comic representation on the right. Vesicles are colored in orange, ribosomes/adjacent proteins dark green, and viral particles in purple (b). Scale bar representing 50 nm.

**Supplementary Table 1.** Cryo-EM acquisition and reconstruction parameters of specimen reported in this study.

	Acquisition 1 (L-A virus) (EMD-15214)	Acquisition 2 (L-A virus) (EMD-15189) (PDB-8A5T)	Acquisition 3 (Ribosome*1) (EMD-15215)	Acquisition 3 (Inner density L-A virus)
<b>Data collection and processing</b>				
Magnification	45000X	92000X	45000X	
Voltage (kV)	200	200	200	
Microscope model	TFS Glacios	TFS Glacios	TFS Glacios	
Camera model	TFS Falcon III EC	TFS Falcon III EC	TFS Falcon III EC	
Number of frames	12*	12*	12*	
Electron exposure (e-/Å <sup>2</sup> )	30	30	30	
Per-frame exposure (e-/Å <sup>2</sup> )	2.3	2.3	2.3	
Defocus range (µm)	-1.0 to -3.0	-1.0 to -3.0	-1.0 to -3.0	
Pixel size (Å)	3.17	1.57	3.17	
Images acquired (no.)	10067	7011	2728	
Acquisition software	TFS EPU 2	TFS EPU 2	TFS EPU 2	
Symmetry imposed	I	I	C1	
Initial particle images (no.)	230,000	2,725,140	750,242	17,549
Final particle images (no.)	65,000	1,020,420	627,748	17,549
Map resolution (Å)	6.4	3.77	6.62	18.13
FSC threshold	0.143	0.143	0.143	0.143
Map B-factor	188.3	198.2	559.3	360.2
<b>Refinement</b>				
Initial model used (PDB code)	-	1m1c	-	-
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-	0 (not modified)	-	-
Model composition				
Non-hydrogen atoms	-	10302		
Protein residues	-	1302		
<i>B</i> factors				
Protein (min/max/mean)	-	76.63/187.82/113.84		
R.m.s. deviations				
Bond lengths (Å)	-	0.004 (0)		
Bond angles (°)	-	0.985 (0)		
Validation				
MolProbity score	-	2.12		
Clashscore	-	15.89		
Poor rotamers (%)	-	0.45		
Ramachandran plot				
Favored (%)	-	93.68		
Allowed (%)	-	6.32		
Disallowed (%)	-	0.00		
Map-CC	-	0.84		

\*Part of the dataset was acquired at 13 frames; the last frame was removed during image analysis.

**Supplementary Table 2.** Identification of structural homologues of L-A virus region 481-587: Structural alignment was performed by DALI (<http://ekhidna2.biocenter.helsinki.fi/dali/>). Z score (Z), root-mean-square-deviation (RMSD) between the two aligned structures, the length of the aligned residues (lali) and the sequence identity (%id) are displayed. The description is fetched from the PDB entry.

<b>Chain</b>	<b>Z</b>	<b>rmsd</b>	<b>lali</b>	<b>%id</b>	<b>Description</b>
<i>1m1c-B</i>	14.9	1.1	97	100	Major Coat Protein
<i>1m1c-A</i>	7.6	0.8	97	100	Major Coat Protein
<i>3e4g-A</i>	3.5	3.7	64	9	ATP Synthase Subunit S, Mitochondrial
<i>3e2j-A</i>	3.5	3.8	61	10	ATP Synthase Subunit S, Mitochondrial
<i>6kms-C</i>	3.4	3.2	46	9	Methyltransferase N6AMT1
<i>3dze-A</i>	3.3	3.7	65	9	ATP Synthase Subunit S, Mitochondrial
<i>6k0x-A</i>	3.3	3.5	45	9	Methyltransferase N6AMT1
<i>2dfx-I</i>	3.2	2.8	60	7	Colicin-E5
<i>2fhz-A</i>	3.2	2.8	60	7	Colicin-E5 Immunity Protein
<i>3qph-A</i>	3.1	3.3	61	7	TRMB, A Global Transcription Regulator
<i>3wvq-A</i>	3.1	3.8	65	17	PGM1
<i>3wvr-B</i>	3.1	3.8	65	17	PGM1
<i>6khs-A</i>	3.1	3.1	45	11	Methyltransferase N6AMT1

## 10. Acknowledgements

This section is one I left open until the very last moment of my thesis. At this point my thoughts are going everywhere at once and nowhere at all. First, I'll start with the official acknowledgements and later it will become a bit more personal and emotional.

I was happy to receive funding from the DFG and be part of the GRK2467. I learned many soft skills and met amazing PhD students which I hold dear. Special Thanks here goes to Andrea Sinz, the head of the research training group as well as Claudia Spielmann, Lolita Piersimoni and Jeanette Köppen.

I also want to say thank you to our administrators from the lab Ulla, Marie, Rosita and Fabienne for their work and help over the last 3+ years.

I am honored and thankful to my TAC committee Milton Stubbs and Stefan Feller for providing me with valuable input, justified criticism and encouragement.

I am also grateful for Prof. Dr. Petra Wendler for agreeing to be my 3<sup>rd</sup> assessor for this thesis.

Lastly, I want to thank all my collaborators for the help with my research and the resulting publication.

Now for the more personal thank you. Let's start with my family. I will switch to German for this part. An meine Familie: Ich kann nicht in Worte fassen wie dankbar ich euch bin. Liebe Mama lieber Papa, vielen Dank für die Unterstützung. Seit meiner Kindheit habt ihr jede meiner Entscheidungen respektiert, unterstützt und seid da gewesen, wann immer ich Schwierigkeiten hatte. DANKE. An meine Schwester: Liebe Celina, danke dass du meine Kindheit in der ich dir eine ältere Schwester sein durfte zu einem sehr schönen Teil meines Lebens gemacht hast. An Tante Andrea: ja ich weiß du magst den Zusatz Tante nicht aber du bist nunmal meine Tante und zudem die Beste die man sich hätte wünschen können. Du warst immer für mich da, in jeder Situation und hast mich unterstützt wo es nur ging. Liebe Omi, auch dir bin ich unendlich dankbar, dass du immer für mich da bist. Immer ein offenes Ohr hast und mich gut mit unseren kleinen Computerspielrunden von stressigen Situationen abgelenkt hast. An dieser Stelle möchte ich auch dich erwähnen Björn, du bist als Freund meiner Schwester in mein Leben getreten und du bist ein sympatischer junger Mann, bitte pass weiter gut auf meine kleine Schwester auf. Vielen lieben Dank auch an Robi und Alex

dafür dass wir einfach ein so tolles Cousintrio abgeben und über alles reden können, sowie Tante Ingrid mit dir kann man sich einfach ewig über Stargate und ähnliches unterhalten.

Let's continue with my second family, my chosen family. Yeah, I mean to address you that way my dear friends. First of all: Steffi und Kevin, ich freue mich jedes Mal euch wieder zu sehen, auch wenn es nur sehr selten ist. Ihr gebt mir immer ein Gefühl von Heimat. For the next ones mentioned I do not know who to mention first I love you all equally, I guess I'll go chronologically. Dear Sarah, we met right in the beginning of our bachelors, and you quickly became a close friend. Since we were living together of course we went on each other's nerves here and there but over all you made studying a lot of fun and I 1000% know I can always count on you. The next one I also met during my bachelors is Katrin. Over the past years we became closer and closer, I can come to you with any problem, and I do not want to miss you in my live anymore. Then also Daniel, thank you for taking me with you to the PnP nights and always being encouraging and there to lend an open ear. The next person is Patrick, I have never met a man before that is as sensitive and empathic as you are. You usually know something is wrong before I even realized myself. I really love all our late-night gaming and talking sessions. Speaking about late night talking:

```
for std(time) in voice_chat:  
    print('Thank you Griesch')
```

The next part will feature my third family: My working family over the past 3+ years. When I had the interview Panos was honest and told me the position will most likely go to Yiannis. I'd never expected to be called a few days later to hear, that he had a second open position and whether I wanted to apply for it. After that I moved to Halle and started my PhD. When I visited the lab for the first time (everything was still wrapped up in boxes), I met Panos and Fotis. Fotis became the heart and soul of the lab. Being Panos' assistant, he was responsible for all of us, and instead of being stressed he made it a habit to go forward with such a good example that you had only one choice and follow it. Then also Annette, an experienced researcher that is always calm, straight forward and polite. The next person I met was Yiannis. 'This guy' as Panos would refer to later on. This guy sat next to me in the office and was a presence. Tall, loud, lovable and helpful in any way possible. Love you man we just shared all the emotions of the PhD together. The next PhD was Kevin, who in the beginning seemed a bit grumpy but coming to know him a bit better he is a sweetheart and going out of his way to help others in the lab. Then also Marija joined our ranks as PhD. We had some

really deep talks and I want to say thank you to you too for the nice time at the lab. I also appreciate our shared love for bouldering Kevin and Marija. Another now PhD is wanting to thank is Tonyyy. He came later, first as an intern, then master student, and now well appreciated member and PhD of the group. I'll continue now with the postdocs. Thank you, Dmitry, for all the help with setting up image analysis and teaching me what to look for. Thank you, Ferrie, for your dedication in teaching and becoming a close friend and conversation partner. I learned an enormous amount about electron microscopy. Christian you are one of the people without whom I would not write this acknowledgement. Thank you for all your support and patience. Also, thanks to Jaydeep. You had a high impact on a nice and fun working atmosphere. Finally, Panos or rather before I say my thanks to your wife Marta. You dedicated a lot of time for the LSM. I really appreciate your help and input. Now it is your turn. In German we have the term 'Doktorvater' and in my opinion this really hits the mark. You became somewhat of a scientific father figure for me. Without you I would not be at this point. You invest yourself to the fullest into the research as well as well-being of your PhD students and I cannot be more thankful to have been offered a position at your lab. This was an amazing but also stressful journey, which was a bit damped by the pandemic but nevertheless, I learned so much and your enthusiasm was always catching over to me. I had the feeling we learned and grew together, you in your position as junior professor and me as a researcher. Πάνο σε ευχαριστώ από τα βάθη της καρδιάς μου για το χρόνο σου, τη βοήθειά σου και την εμπιστοσύνη που μου δείχνεις ως επιστήμονα.

The final thanks go to the new family (WIP). Hendric. I am at a loss for words right now, sitting next to you, writing these few sentences. How can I ever show you how grateful I am that you stumbled into my life. From being the most understanding and supportive person, to giving your honest thoughts every time to keeping everything off my back while I was trying to finish my thesis. I must have been frustrating during the last few months, as I was overwhelmed with the new position while simultaneously writing my thesis. I just fall in love with you over and over again, even though I might sometimes not show it as much as you'd deserve. THANK YOU, I LOVE YOU!!!

## 11. Resume

<i>Personal data</i>	Name: Date of birth: Place of birth:	Lisa Schmidt 18.03.1991 Gera, Germany
<i>Professional Experience</i>	<b>2023-present</b>  <b>April-July 2016</b>	Scientific employee UFZ Leipzig ISOBIO-ProVIS MetaSCOPE Project – Group Dr. Stryhanyuk Scientific employee TU-Braunschweig Institute of Plant biology – Group Mendel
<i>Education</i>	<b>2019-2023</b>  <b>2013-2016</b> <b>2009-2013</b>	PhD candidate MLU Halle-Wittenberg Institute of Biochemistry and Biotechnology GRK2467 Group of Jun. Prof. Dr. P.L. Kastritis M.Sc. TU-Braunschweig B.Sc. TU-Braunschweig
<i>Non-scientific work experience</i>	<b>2017-2018</b> <b>2012-2017</b>	Customer representative Buchpartner Part timer for Buchpartner during university
<i>Research Training</i>	<b>14./21./28.06.2022</b> <b>12./19.01.2022</b> <b>29.11.-01.12.2021</b>  <b>20.-22.10.2021</b>  <b>04.10.2021</b>  <b>28.09.2021</b> <b>27.07.2021</b> <b>26.05./01./08./15.06.2021</b>  <b>27.05.2021</b>  <b>08.04.2021</b>  <b>04.02.2021</b>  <b>28.-30.09.2020</b> <b>09.09.2020</b>  <b>24.-27.08.2020</b>  <b>28.-29.05.2020</b> <b>13.11.2019</b>	Scope and vision of AlphaFold webinar Scientists in Leadership Workshop Recent Advances in Structural Biology of Membrane Proteins EMBO workshop virtual event EMBL Conference EMBL virtual event: Bringing Molecular Structure to Life: 50 Years of the PDB IPB MLU Halle-Wittenberg Conflict management in academia Science Communication Workshop VIAS Halle: Career coaching for females Introduction into Statistics and R- course I & II Research results, patents and more by Dr. Andreas Lauenroth, Steffen Ahrens, Dr. Robert Szczesny and Moritz Bradler Ateliers de Inserm Paris France: Workshop #262 Inserm (BFSSR) Mass spectrometry for structural biology IDP seminars Invited speakers: Dorothee Dormann and Sonia Longhi Statistics course by sum of squares Good scientific practice workshop by sum of squares Good manufacturing practice by sum of squares Thermo Fisher Glacios training 'in-house' JEOL training with Dr. Farzad Hamdi



	<b>20.09.-26.09.2019</b> <b>26.-30.08.2019</b> <b>20.07.2019</b> <b>16.-19.07.2019</b>	Chemical disposal seminar ThermoFisher cryo-EM training on Glacios EPR Spectroscopy seminar ThermoFisher cryo-EM training on Glacios
<i>Presentations</i>	<b>09.08.2022</b> <b>04-06.05.2022</b> <b>12.02.2022</b> <b>29.11.-01.12.2021</b> <b>15./16.09.2022</b> <b>06.09.2021</b> <b>22.07.2021</b> <b>28.05.2021</b> <b>19.-21.11.2020</b> <b>14.-16.10.2019</b>	TAC report Halle: Presentation Delineating organizational principles of the endogenous L-A virus by cryo-EM and computational analysis of native cell extracts GRK2467-seminar Weißenfels: Presentation Hidden structural states of the endogenous L-A virus revealed by cryo-EM and computational analysis of native cell extracts Biochemical, structural and functional characterization of enzymatic pathways GRK2467-seminar Nimbschen: Biochemical, structural and functional characterization of enzymatic pathways Advances in cryo-EM ZIKHALOmem conference MLU Halle-Wittenberg Cryo-EM of yeast extracts provides insights into endogenous virus-embedded protein disorder TAC report Halle Biochemical, structural and functional characterization of enzymatic pathways GRK2467-internal Conference: Discovery and structural analysis of the endogenous virus particles from native cell extracts - Poster GRK2467 internal conference: Methods of the RTG2467 (Presentation "Cryo-Electron-Microscopy International ZIK HALOmem conference (Poster: Towards the structural characterization of the Metabolic complexes across scales pyruvate oxidation metabolons with cryo-EM)
<i>Teaching</i>	<b>18.02.2021</b> <b>09.02.2021</b> <b>23.-27.11.2020</b>	Image Processing/Data analysis at Structural Biology Masters Course Structure elucidation of proteins and protein complexes by cryo-EM and proteomics Image Processing/Data analysis

## 12. Publication list

1. **Schmidt, L.**, Tüting, C., Kyrilis, F.L., Hamdi, F., Semchonok, D.A., Hause, G., Meister, A., Ihling, C., Shah, P.N.M., Stubbs, M.T., *et al.* (2022). Delineating organizational principles of the endogenous L-A virus by cryo-EM and computational analysis of native cell extracts. **bioRxiv**, 2022.2007.2015.498668.
2. Hamdi, F., Tuting, C., Semchonok, D.A., Visscher, K.M., Kyrilis, F.L., Meister, A., Skalidis, I., **Schmidt, L.**, Parthier, C., Stubbs, M.T., *et al.* (2020). 2.7 Å cryo-EM structure of vitrified *M. musculus* H-chain apoferritin from a compact 200 keV cryo-microscope. **PLoS One** 15, e0232540.
3. Piersimoni, L., Abd El Malek, M., Bhatia, T., Bender, J., Brankatschk, C., Calvo Sanchez, J., Dayhoff, G.W., Di Ianni, A., Figueroa Parra, J.O., Garcia-Martinez, D., *et al.* (2022). Lighting up Nobel Prize-winning studies with protein intrinsic disorder. **Cell Mol Life Sci** 79, 449.
4. Tuting, C., **Schmidt, L.**, Skalidis, I., Sinz, A., and Kastritis, P.L. (2023). Enabling cryo-EM density interpretation from yeast native cell extracts by proteomics data and AlphaFold structures. **Proteomics**, e2200096.

### 13. Affidavit /Eidesstaatliche Erklärung

I hereby declare to have written this thesis independently without help or assistance all external sources figuratively or literally presented are cited accordingly. I also declare that I have not submitted this thesis to any other university, College or else to obtain an academic degree.

Ich erkläre hiermit, dass ich diese Arbeit selbständig und ohne fremde Hilfe verfasst habe und dass alle wörtlich oder bildlich genutzten externen Quellen, entsprechend zitiert wurden. Ich erkläre auch, dass ich diese Arbeit an keiner anderen Universität, Hochschule oder sonstigen Einrichtung eingereicht habe, um einen akademischen Grad zu erlangen.

Halle (Saale), 28.07.2023

Lisa Schmidt