Visualization of the interaction between precursor proteins and the outer mitochondrial translocase using cryoEM

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich
Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von

Pamela Ornelas

aus Guadalajara, Mexiko

Frankfurt am Main, 2023
(D30)

Diese Arbeit wurde in der Abteilung Strukturbiologie von Prof. Werner Kühlbrandt des Max-Planck-Instituts für Biophysik durchgeführt und vom Fachbereich Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe–Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Prof. Dr. Klaas Martinus Pos

Prof. Dr. Werner Kühlbrandt

Datum der Disputation:

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe.

Pamela Ornelas

Frankfurt am Main

Table of Contents

List of Abbreviations

ALDH aldehyde dehydrogenase

ATP adenosine triphosphate

CCD charge-coupled device

cryoEM electron cryo-microscopy

cryoET electron cryo-tomography

CTF contrast transfer function

DHFR dihydrofolate reductase

DNA deoxyribonucleic acid

DSSO disuccinimidyl sulfoxide

DDM dodecylmaltoside

EDC 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride

FEG field-emission gun

FSC Fourier shell correlation

GDN glyco-diosgenin

IMAC immobilized-metal affinity chromatography

IMS intermembrane space

IR infrared

LB lysogeny broth

LILBID laser-induced liquid bead ion-desorption

LMNG Lauryl maltose neopentyl glycol

MBP maltose binding protein

MIA mitochondrial intermembrane space import and assembly

MIM mitochondrial import machinery

MPP mitochondrial processing peptidase

MS mass spectrometry

nanoDSF nano differential scanning fluorimetry

NiNTA Nickel-nitrilotriacetic acid

NMR nuclear magnetic resonance

OMV outer membrane vesicle

PAM presequence translocase-associated motor

pALDH presequence of aldehyde dehydrogenase

PAGE polyacrylamide gel electrophoresis

PC phosphatidylcholine

PE posphatidylethanolamine

pSu9 presequence of ATP synthase subunit 9

SAM sorting and assembly machinery

sT small Toms

TEM transmission electron microscopy

TIM translocase of the inner mitochondrial membrane

TOM translocase of the outer mitochondrial membrane

TPR tetratrico peptide repeat

VDAC voltage-dependent anion channel

XL-MS crosslinking mass spectrometry

List of Figures

Summary

Mitochondria are important organelles in eukaryotic cells. Popularly known as the *power house of the cell*, their relevance comes from the oxidation process that happens within their inner membrane, which converts the energy that fuels most activities in living cells. Mitochondria are only able to produce a small fraction of the proteins they require to function. Ribosomes in their matrix synthesize proteins from mitochondrial deoxyribonucleic acid (DNA), such as those involved in oxidataive phosphorilation. The remaining mitochondrial proteins are encoded in the cellular nucleus and synthesised in the cytoplasm in the form of precursors. Precursor proteins contain a mitochondrial targeting signal, and, with help of chaperones, are able to find their way into mitochondria, whether they belong to the intermembrane space, the matrix or either one of the membranes.

The main gate into mitochondria, encountered by most precursor proteins, is the translocase of the outer membrane (TOM) complex. Structural studies show that in its most basic form, TOM *core* complex is a membrane-embedded homodimer formed by two β-barrel translocation pores (Tom40), two copies of a cytoplasmic receptor (Tom22), and six small transmembrane TOM subunits that stabilize the complex (Tom7, Tom6 and Tom5). Two additional subunits have been identified: Tom70 and Tom20, which along with the core complex, are part of TOM *holo* complex. Tom70 and Tom20 tend to dissociate from the core complex during isolation from mitochondria, therefore their stoichiometry and location within the complex is not fully understood. However, they are known to aid in precursor recognition because they distinguish a wider range of targeting signals than Tom22.

For years, biology has relied on physical methods to study life at a cellular level. Light microscopy allowed the observation of countless microorganisms and to understand cellular processes in human bodies. Crystallography has given us a window into the molecular world, becoming a key in our early understanding of DNA. In recent years,

electron cryo-microscopy (cryoEM) has become a powerful method in structural analysis of biomolecules. Through sample vitrification and cooling of microscopes, cryoEM has revealed the structure of proteins at atomic resolution. This has resulted in accurate models of thousands of proteins in a near-native environment.

This thesis focuses on studying the structure of TOM holo complex through single-particle cryoEM. Results of this work are supported by other structural biology techniques, such as mass spectrometry and computational structure prediction. This way, we aimed to learn more about the complex and its function, as well as its precursor protein translocation mechanism. Purifying TOM holo complex directly from *Neurospora crassa* mitochondria presented a great challenge, which we solved by isolating TOM from outer mitochondrial membrane vesicles instead. In addition, we produced two precursor protein constructs and incubated them with the complex prior to our analysis: the presequence of ATP synthase subunit 9 coupled to maltose binding protein (pSu9-MBP), and the presequence of aldehyde dehydrogenase (pALDH).

The structural analysis of TOM complex through cryoEM provided valuable insights into its architecture. An initial data set, collected from a TOM holo sample incubated with precursor pSu9-MPB, revealed a large feature on the cytoplasmic side of the complex, likely corresponding to Tom20 interacting with MBP. Incubation of the complex with the shorter precursor, pALDH, resulted in a more stable data set. At 3.3 Å, the structure of *N. crassa* TOM core complex presented high-resolution features of its five subunits. The map showed that both Tom40 pores interact with each other only at the cytoplasm-membrane interface, and are separated by a phospholipid that forces them to tilt $\sim 20^{\circ}$ relative to the membrane plane. This cryoEM density map made it possible to build an atomic model of its structure and to compare it with TOM complexes from other organisms. Furthermore, a 4 Å resolution map of the complex indicated binding of pAlDH as it crosses Tom40 from cytoplasm to intermembrane space. By comparison with the core model, possible contact sites between the pore and pALDH were identified, mainly with the internal α-helices of Tom40.

At 6-7 Å resolution, two maps of *N. crassa* TOM holo complex were obtained. Besides five subunits corresponding to the core domain, these maps showed a clear cytoplasmic domain that resembles an α-helix protruding from the edge of the micelle towards the

pores, in two different conformations. Mass spectrometry helped identify this new density. Various subcomplexes formed by TOM holo subunits were identified through laser-induced liquid bead ion desorption (LILBID). Of particular interest were subcomplexes formed by Tom70 and the small TOM subunits, as well as receptor subcomplexes involving Tom70, Tom22 and Tom20. Moreover, multiple interactions involving Tom40, Tom22, Tom20 and the small TOMs, point at a possible conformation of a TOM holo protomer. Crosslinking mass spectrometry (XL-MS) indicated a great number of interactions between Tom70 and Tom20. XL-MS further confirmed the cooperation between Tom70 and small TOMs, as well as between Tom20 and Tom40 and Tom22.

Based on these findings, prediction models of Tom70 and Tom20 were generated using AlphaFold software, and rigid-body fitted into both TOM holo maps. The predicted Tom20 structure matched the map's density both in size and shape, as it interacts with Tom40 and Tom22. Additional subunit interactions predicted by AlphaFold match those observed ithrough mass spectrometry. The $Tom20_1Tom22_1Tom40_1$ subcomplex, in particular, resulted in a conformation that agreed with both TOM holo maps. Hence, the identity of cytoplasmic density in our maps was confirmed as Tom20, present in two conformations. In one conformation, Tom20 hovers above both translocation pores, while connected only to Tom22. In the second, Tom20 moves over a single pore and interacts with the longest cytosolic loop in Tom40.

These structures show that Tom20 is a flexible subunit, docked to the core complex at a single point, the N-terminal of Tom22. Tom20 uses this as a pivot point in order to move its receptor domain from one pore to the other. At its N-terminal, Tom20 has a transmembrane helix that is only partially resolved. This helix does not interact with other TOM subunits inside the membrane, and is likely to be flexible in order to allow for the movement in the cytoplasm. This single binding point of Tom20 explains why it tends to easily dissociate from the complex during purification. However, the interaction between these subunits appears to be stabilized by the disordered N-terminal of Tom22, which wraps around Tom20's helix at the docking site.

The recorded cryoEM data set also contained particles larger than the TOM core dimer, that were initially discarded. The density maps generated from these particles have a strongly oval micelle and exhibit a cytoplasmic domain, indicating the presence of a second copy of Tom20 or the presence of Tom70. However, due to the small number and low image quality of these particles, the density maps generated from them only reached 16 Å resolution, making them difficult to interpret. It can be suggested that some of these larger particles contain Tom70. Structural predictions show that Tom70 consists of a soluble receptor domain connected to a single transmembrane helix via a long unstructured region. This indicates that Tom70 is highly flexible and may have multiple interaction sites with the TOM complex. Thus, the larger particles may represent the TOM complex to which Tom70 is bound at multiple sites. This heterogeneity in the particles complicates the reconstruction of a high-resolution cryoEM density map.

In summary, the results of this thesis indicate a translocation mechanism in which Tom20 and Tom70 are the main receptors of the TOM complex. After reception of a signal sequence, Tom20 approaches a free pore of the TOM complex, with the help of Tom22, and deposits the precursor protein into the Tom40 barrel. We suggest that Tom40 mediates translocation of the precursor by means of electrostatic interactions with the positive signal sequence, while keeping it stable by hydrophobic interactions between its two inner helices.

Zusammenfassung

Für die Erhaltung der elementaren und überlebenswichtigen Prozesse in eukaryotischen Zellen sind Mitochondrien von entscheidender Bedeutung. Die wesentliche Rolle dieser Organellen für die Zelle ergibt sich unter anderem aus dem Prozess der oxidativen Phosphorylierung, der in der inneren Mitochondrienmembran abläuft und die eukaryotische Zelle mit der überlebensnotwendigen Energie versorgt. Mitochondrien sind in verschiedene Kompartimente unterteilt, die durch eine innere und eine äußere Membran voneinander getrennt sind.

Nach der Endosymbiontentheorie sind Mitochondrien ursprünglich durch den Einschluss einer prokaryotischen Zelle in einen anderen Zelle entstanden. In der darauffolgenden Symbiose zwischen beiden Zellen und im Laufe der Evolution gingen immer mehr Teile des Genoms der eingeschlossenen Zelle verloren, wodurch die heutige mitochondriale Desoxyribonukleinsäure (DNS) nur noch eine geringe Anzahl an Genen besitzt. Aus diesem Grund sind Mitochondrien nur noch in der Lage einen geringen Anteil an Proteinen zu produzieren, welche zum größten Teil an der oxidativen Phosphorylierung beteiligt sind. Alle restlichen Proteine der Mitochondrien sind im Zellkern kodiert und werden von zytoplasmatischen Ribosomen in Form von Vorläuferproteinen synthetisiert. Die Vorlaüferproteine sind zunächst nicht funktionsfähig und enthalten eine mitochondriale Signalsequenz. Diese Signalsequenz stellt sicher, dass die Proteine mit Hilfe von Chaperonen ihren Weg in die Mitochondrien finden und dort an ihren endgültigen Platz transportiert werden können, sei es in den Intermembranraum, die Matrix oder eine der beiden Membranen der Mitochondrien.

Das Tor in die Mitochondrien, welches die meisten Vorläuferproteine passieren müssen, ist der TOM (translocase of the outer membrane) -Komplex. Dieser Komplex ist mit mehreren Rezeptoren ausgestattet und agiert als eine Art Maschine, die Vorläuferproteine für den Import sortiert und verteilt. In der äußeren Mitochondrienmembran arbeitet TOM beispielsweise mit der sorting and assembly machinery (SAM) zusammen, um

Vorläuferproteine mit β-barrels in den Intermembranraum zu transportieren und dort zusammenzubauen. Im Intermembranraum übergibt schließlich TOM die Vorstufen der Proteine an die Translokasen der inneren Membran, welche als TIM22 und TIM23 bezeichnet werden, für den endgültigen Import in die mitochondriale Matrix.

Strukturanalysen des TOM-Komplexes zeigen, dass dieser in seiner grundlegendsten Form ein in die Membran eingebettetes Homodimer darstellt, das aus zwei β-barrel (Tom40), zwei Kopien eines zytoplasmatischen Rezeptors (Tom22) und sechs kleinen Transmembranhelices zur Stabilisierung des Komplexes (Tom7, Tom6 und Tom5) besteht. All diese Untereinheiten bilden den sogenannten TOM Kernkomplex. Weiterhin wurden die zwei Untereinheiten Tom70 und Tom20 identifiziert, die zusammen mit dem Kernkomplex einen Teil des sogenannten TOM-holo-Komplexes bilden. Da Tom70 und Tom20 stark dazu tendieren, während der Isolierung des TOM-Komplexes zu dissoziieren, ist ihre Stöchiometrie und ihre Position innerhalb des Komplexes noch nicht vollständig geklärt. Es ist jedoch bekannt, dass sie ein breiteres Spektrum an Signalsequenzen als Tom22 erkennen können.

Seit Jahren stützt sich die Biologie auf physikalische Methoden zur Untersuchung der Zellen. Mithilfe der Lichtmikroskopie wurde es möglich, unzählige Mikroorganismen zu beobachten und die zellulären Prozesse im menschlichen Körper besser zu verstehen. Die Strukturaufklärung von Biomolekülen mittels Röntgenkristallographie öffnete ein Fenster in die molekulare Welt und war bspw. ein entscheidendes Element für das Verständnis des strukturellen Aufbaus der DNS. In den letzten Jahren hat sich die Kryo-Elektronenmikroskopie (KryoEM) zu einer leistungsfähigen Methode für die Strukturanalyse von Biomolekülen entwickelt. Durch die Vitrifizierung der Proben und die Kühlung der Mikroskope hat KryoEM die Struktur von Proteinen mit atomarer Auflösung enthüllt. Dies hat zu genauen Modellen von Tausenden von Proteinen in einer nahezu natürlichen Umgebung geführt.

Diese Arbeit konzentriert sich auf die Untersuchung der Struktur des TOM-Holokomplexes mittels Einzelteilchen-KryoEM. Die Ergebnisse dieser Arbeit werden durch weitere strukturbiologische Techniken, wie Massenspektrometrie und computergestützte Strukturvorhersagen unterstützt. Zielsetzung war es, ein besseres Verständnis über die Struktur

und Funktion des TOM-Komplexes zu gewinnen und seinen Mechanismus der Proteintranslokation zu entschlüsseln. Im Rahmen dieser Arbeit wurde der TOM-Holokomplex aus Vesikeln der äußeren Mitochondrienmembran isoliert, welche aus dem Pilz *Neurospora crassa* gewonnen wurden. Darüber hinaus wurden die zwei Vorläuferproteinkonstrukte pSu9-MBP und pALDH hergestellt, mit dem Komplex inkubiert und ihre Interaktion mit diesem untersucht.

Die durchgeführten Strukturanalysen des TOM-Komplexes mittels KryoEM liefern wertvolle Einblicke in seine Architektur. Analysen des TOM-Holokomplexes, welcher mit pSu9-MPB inkubiert wurde, enthüllten Dichten auf der zytoplas- matischen Seite des Komplexes, welche auf die Interaktion von Tom20 mit pSu9-MBP zurückzuführen sind. Die KryoEM-Analyse des Komplexes, inkubiert mit dem etwas kürzeren Vorläuferprotein pALDH, ergab hingegen eine deutliche höhere Auflösung und die Struktur des *N. crassa* TOM-Kernkomplexes mit seinen fünf Untereinheiten konnte mit einer Auflösung von 3.3 Å gelöst werden. Die Struktur zeigt, dass die Tom40-Poren nur an der Grenzfläche zwischen Zytoplasma und Membran miteinander interagieren und durch ein Phospholipid voneinander getrennt sind. Dies führt zu einer Neigung beider Poren von etwa 20° zueinander. Die hochauflösende KryoEM-Dichtekarte ermöglichte es weiterhin, ein atomares Modell der Struktur zu erstellen und mit TOM-Komplexen aus anderen Organismen zu vergleichen. Darüber hinaus weist Dichte mit einer Auflösung von etwa 4 Å auf die Bindung von pALDH hin, welches Tom40 vom Zytoplasma zum Intermembranraum durchquert. Im Vergleich zum Kernmodell konnten potentielle Kontaktstellen zwischen der Pore und pALDH identifiziert werden, welche sich hauptsächlich innerhalb der α-Helices in Tom40 befinden.

Mit einer Auflösung von jeweils 6-7 Å konnten zusätzlich zwei Dichtekarten des TOM-Holokomplexes mittels KryoEM bestimmt werden. Neben den fünf Untereinheiten der Kerndomäne, zeigen diese Karten die Anwesenheit einer zytoplasmatischen Domäne, die einer α-Helix ähnelt. Diese Helix verläuft vom Rand der Detergenz-Mizelle in Richtung der Tom40-Poren und liegt in zwei verschiedenen Konformationen vor. Mit Hilfe von Massenspektrometrie konnte diese Dichte identifiziert werden und durch Laser-induzierter Flüssig-Ionen-Desorption (LILBID) konnten verschiedene Unterkomplexe identifiziert werden, die von den TOM-Holo-Untereinheiten gebildet werden. Von besonderem Interesse

waren die Unterkomplexe, die von Tom70 und den kleinen TOM-Untereinheiten gebildet werden, sowie die Rezeptor-Unterkomplexe, die durch Tom70, Tom22 und Tom20 gebildet werden. Die Daten weisen auf Wechselwirkungen zwischen Tom40, Tom22, Tom20 und den kleinen TOMs hin und liefern somit Einblicke in die potentiellen Konformationen des TOM-Holoprotomers. Analysen mittels Crosslinking-Massenspektrometrie (XL-MS) zeigten weiterhin Wechselwirkungen zwischen Tom70 und Tom20 auf und bestätigten die Interaktion zwischen Tom70 und den kleinen TOM-Untereinheiten, sowie zwischen Tom20 und Tom40 und Tom22.

Auf Grundlage dieser Erkenntnisse wurden Strukturen mittels der Software AlphaFold für Tom70 und Tom20 vorhergesagt und diese in die Dichtekarten des TOM-Holo-Komplexes eingefügt. Die vorhergesagte Struktur von Tom20 stimmte in Größe und in Form mit der Dichte überein und scheint mit Tom40 und Tom22 zu interagieren. Weitere von AlphaFold vorhergesagte Untereinheiten-Interaktionen stimmen mit denen überein, die durch Massenspektrometrie ermittelt wurden. Insbesondere der vorhergesagte Unterkomplex $Tom20_1Tom22_1Tom40_1$ zeigt eine Konformation, die mit den experimentell bestimmten KryoEM-Dichtekarten des Komplexes übereinstimmt. Die beobachtbare zytoplasmatischen Dichte in den KryoEM-Karten zeigt Tom20, welches in einer der beiden beobachtbaren Konformation über den Translokationsporen schwebt und mit Tom22 interagiert. In der zweiten Konformation scheint Tom20 über nur eine der beiden Poren zu liegen und dabei zusätzlich mit der zytosolischen Schleife von Tom40 zu interagieren.

Die Strukturanalysen zeigen, dass Tom20 eine hochflexible Untereinheit ist und lediglich über eine Schnittstelle am N-Terminus von Tom22 an den Kernkomplex bindet. Es scheint, dass Tom20 diese Schnittstelle als Dreh- und Angelpunkt nutzt, um seine Rezeptordomäne über beide Poren des TOM-Komplexes zu bewegen. Tom20 besitzt eine N-terminale Transmembranhelix, welche in den Kryo-EM-Dichtekarten nur teilweise aufgelöst ist. Diese Helix interagiert nicht mit anderen TOM-Untereinheiten innerhalb der Membran und ist wahrscheinlich hochflexibel, um Bewegungen der löslichen Proteindomäne im Zytoplasma zu kompensieren. Die nur kleine Schnittstelle zwischen Tom20 und Tom22 bietet eine Erklärung dafür, warum Tom20 im Laufe der Isolation des Komplexes leicht von diesem dissoziiert. Die Interaktion zwischen beiden Untereinheiten scheint jedoch durch den ungeordneten N-Terminus von Tom22 stabilisiert zu werden, welcher sich um die Helix

von Tom20 nahe der Bindestelle wickelt. Die KryoEM-Dichtekarten weisen nur Dichte für eine Kopie von Tom20 auf, die an den TOM-Komplex gebunden ist. Die LILBID-Messungen zeigen jedoch Subkomplexe, bei denen zwei Kopien von Tom20 gebunden sind. Da die Stöchiometrie des TOM-Holokomplexes noch unbekannt ist, wurde die Struktur des TOM-Komplexes zusätzlich mit zwei Kopien von Tom20 und Tom70 mittels AlphaFold vorhergesagt. Dabei zeigte sich, dass Tom20 und Tom70 um den Platz oberhalb von Tom40 konkurrieren und dass zwei Kopien von Tom20 theoretisch an den Komplex binden können, allerdings nur in einer anderen Konformation als der, die mit KryoEM aufgelöst wurde.

Die aufgenommenen KryoEM-Daten zeigen auch Partikel mit einer viel größeren Form im Vergleich zum TOM-Kernkomplex. Die aus diesen Partikeln generierten Dichtekarten besitzen eine stark ovale Mizelle und weisen eine zytoplasmatische Domäne auf, die auf das Vorhandensein einer zweiten Kopie von Tom20 hindeutet. Aufgrund der geringen Anzahl und geringen Bildqualität dieser Partikel, erreichten die daraus generierten Dichtekarten lediglich eine Auflösung von 16 Å, was die Interpretation erschwert. Jedoch lässt sich vermuten, dass einige dieser größeren Partikel die Untereinheit Tom70 enthalten. Die Strukturvorhersagen zur Untereinheit Tom70 zeigen, dass diese aus einer löslichen Rezeptordomäne besteht, welche über einen langen unstrukturierten Bereich mit einer einzelnen Transmembranhelix verbunden ist. Dies deutet darauf hin, dass Tom70 sehr flexibel ist und mehrere Interaktionsstellen mit dem TOM-Komplex haben kann. Somit stellen die größeren Partikel möglicherweise den TOM-Komplex dar, an den Tom70 an vielen verschiedenen Stellen gebunden ist. Diese Heterogenität in den Partikeln erschwert die Rekonstruktion einer hochauflösenden KryoEM-Dichtekarte.

Zusammenfassend weisen die Ergebnisse dieser Arbeit auf einen Translokationsmechanismus hin, bei dem Tom20 und Tom70 die Hauptrezeptoren des TOM-Komplexes darstellen. Nach dem Binden an die Signalsequenz eines Vorläuferproteins, nähert sich Tom20 mit Hilfe von Tom22 einer freien Pore des TOM-Komplexes und lagert das Vorläuferprotein in das Tom40-Fass ein. Das Vorläuferprotein wird schließlich über elektrostatische Wechselwirkungen zwischen Tom40 und der positiven Signalsequenz, durch die Pore transportiert und durch hydrophobe Wechselwirkungen zwischen seinen beiden inneren Helices stabil gehalten.

Chapter 1

Introduction

Proteins are essential building blocks for life and as such, play an important role in all cellular activities. Studying them helps us understand their correct function and the effects they have in the cells' behavior. In general, knowing the structure of biological macromolecules is advantageous because it allows us to predict the effects of their malfunction, diagnose disease and to develop drugs to fix them.

Since the prediction of the helical structure of deoxyribonucleic acid (DNA), our understanding of cellular biology keeps improving, leading to extraordinary advances in structural biology. Currently, structure biology can be used to study biological machines. From their atomic composition, protein structure, and even their interactions with other molecules inside and outside cells. Methods such as X-Ray crystallography, nuclear magnetic resonance (NMR), electron cryo-microscopy (cryoEM), and electron cryo-tomography (cryoET) are common methods that have resulted in published structures of thousands of molecular assemblies. In addition, techniques like fluorescence microscopy, mass spectrometry and molecular dynamics simmulations, have given us greater insights into the behaviour of proteins, lipids and nucleic acids. In general, the combination of biochemical, physical, and computational methods has allowed us to better understand the macromolecular world in our cells.

1.1 Biological framework

1.1.1 Mitochondria

Mitochondria are essential organelles for cell function. They take part in fundamental cellular processes such as energy metabolism, homeostasis, intracellular signaling and apoptosis (*1*). This dynamic organelle is formed by four main components. First, the outer mitochondrial membrane, which acts as a barrier between the cytoplasm and the intermembrane space (IMS). In this membrane we can find crucial machinery for mitochondrial communication with the cell. Its most abundant protein is the voltage-dependent anion channel (VDAC) (*2*).

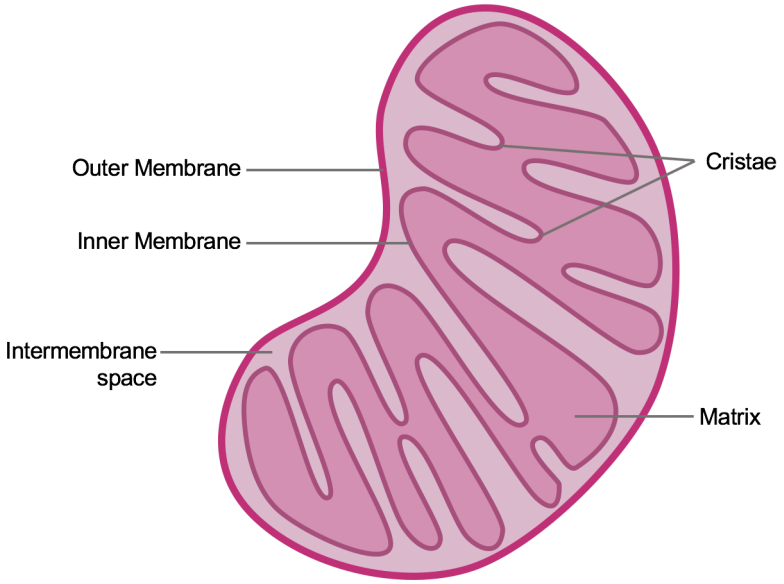

Figure 1.1: **Architecture of mitochondria.** Mitochondria are essential organelles for cellular activity. They are formed by two phospholipid membranes: the outer membrane separates the cytoplasm from the mitochondria and the inner membrane separates the intermembrane space from the matrix.

At the core of mitochondria, the matrix hosts mitochondrial DNA, mitochondrial ribosomes, and metabolic processes such as the citric acid cycle. The IMS is an aqueous medium delimited by the cristae-forming inner mitochondrial membrane and the outer mitochondrial membrane (*1*). Components of the mitochondrial respiratory chain are found in cristae. During respiration, electrons are transferred along four protein complexes embedded in the inner membrane, which results in a flow of protons from the matrix into the IMS. This proton gradient drives the production of adenosine

triphosphate (ATP), the energy source of all cellular processes, as ATP synthase pumps protons back into the matrix (*3*, *4*). This process, also known as oxydative phosphorilation, as many other mitochondrial mechanisms, relies on the inner membrane potential ($\Delta\psi$).

1.1.2 Protein translocation

The unusual structure of mitochondria can be understood by the endosymbiotic theory, which explains the similarity between mitochondria and bacteria. According to this theory, after many years of evolution, a bacteria-like prokaryote was engulfed by an archaeon (*5*, *6*). Subsequently, the prokraryote transferred the majority of its genes to the host and became the precursor of today's mitochondrion.

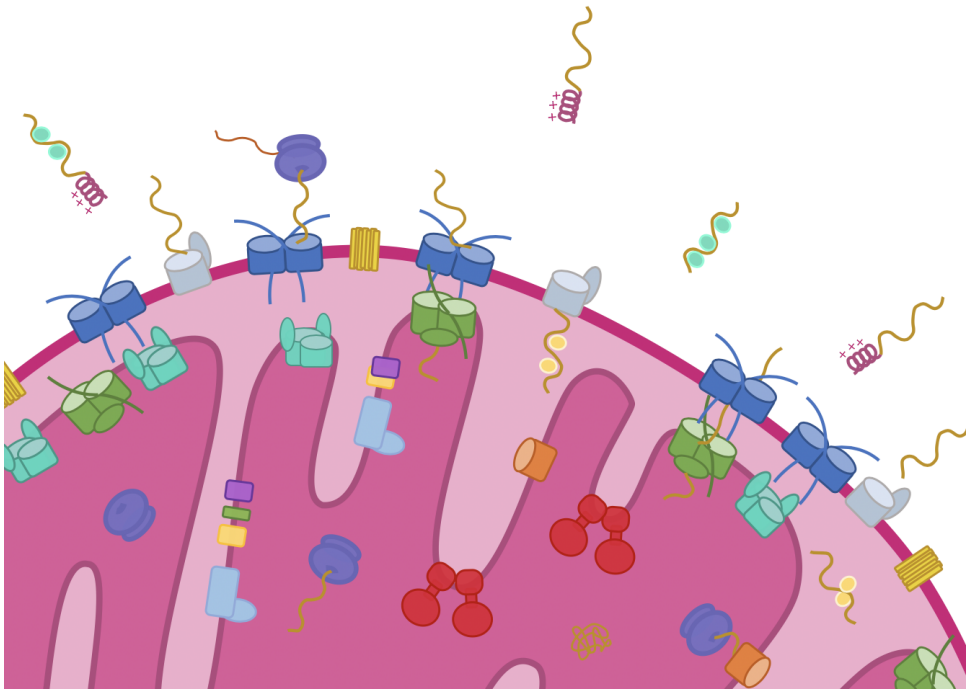

Figure 1.2: **Intricacies of mitochondria.** Multiple protein complexes are necessary for mitochondrial function. Because of their endosymbiotic origin, mitochondria are intricate organelles that, despite containing their own DNA and protein expression systems, must import nuclear-encoded proteins.

Currently, only close to 1% of mitochondrial proteins are synthesized in the matrix, while the rest are synthesized in the cell's cytoplasm. Therefore, specialized machines for protein expression and target signaling have developed, giving rise to target signaling and organelle protein import (*5*, *6*).

Precursor proteins

Most mitochondrial proteins are encoded in the nuclear genome as precursor proteins. After being synthesized by cytosolic ribosomes. Chaperones keep the precursors stable, as they are not synthesized in their functional form and are prone to degradation (*7*, *8*). These precursors contain targeting signals recognized by receptors in the outer membrane of mitochondria that direct them into their respective subcompartment in the organelle (*6*, *9*).

Several types of mitochondria targeting signals have been identified. The most abundant are matrix-targeted N-terminal signals, also known as presequences (*10*). They typically form 15-50 residue-long amphipathic helices with a positively charged side and a hydrophobic side, and are proteolytically cleaved inside the matrix (*7*). Proteins targeted to the inner membrane contain an additional hydrophobic signal, which allows them to be laterally inserted into the membrane during translocation (*11*). Many other precursor proteins contain non-cleavable signals. Metabolite carrier proteins, for example, are directed to the inner membrane through multiple internal targeting signals. These sequences remain in the mature protein once it has reached its final destination within mitochondria. Outer membrane and IMS proteins contain mostly internal targeting signals. Less abundant types have been observed, such as C-terminal and cysteine abundant signals (*6*, *11*).

Mitochondrial import

Mitochondrial protein import depends on a complex system spanning both outer and inner membranes; because of its endosymbiotic past, it is highly conserved across eukaryotic species (*12*). Several specialized machines are necessary to ensure the arrival of precursor proteins to their final location within the mitochondria in a functional conformation. The translocase of the outer mitochondrial membrane (TOM) complex is the first point of contact for most precursor proteins entering mitochondria. It contains several receptor subunits facing the cytoplasm, allowing it to interact with

different types of targeting signals (*13*). After importing them into the IMS, TOM distributes precursor proteins into distinct pathways that lead to various mitochondrial subcompartments (*9*). Figure 1.3 illustrates protein import routes into mitochondria.

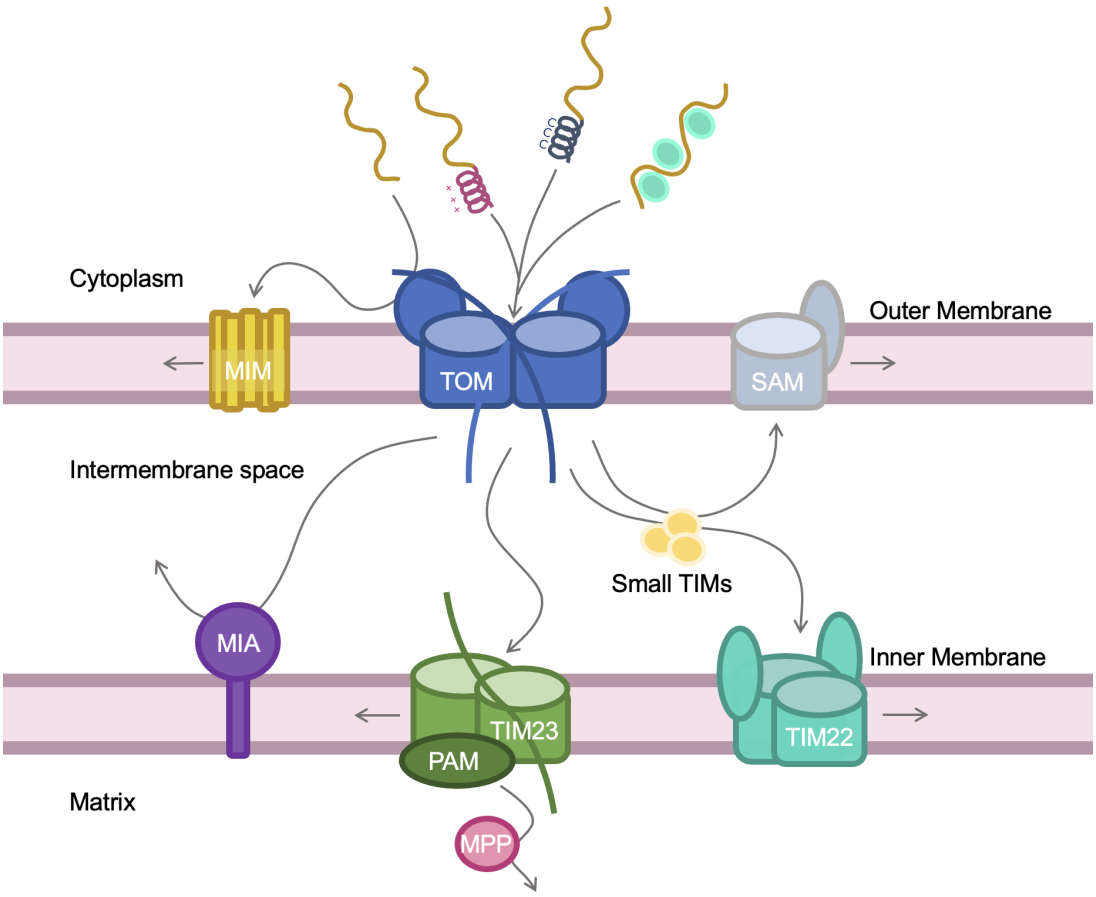

Figure 1.3: **Mitochondrial import.** Most precursor proteins enter mitochondria through the translocase of the outer membrane (TOM). Once in the intermembrane space, proteins can be translocated into the inner membrane by the translocase of the outer membrane (TIM) complexes. Matrix proteins are imported by the presequence translocase-associated motor (PAM) and then cleaved by the mitochondrial processing peptidase (MPP). Outer membrane proteins are encountered by small TIM proteins and delivered to the sorting assembly machinery (SAM). Alternatively, proteins can be inserted directly into the outer membrane by the mitochondrial import machinery (MIM) without TOM translocation. Simplified from Pfanner et al., 2019 (*13*).

Pecursor proteins headed to the inner membrane are handled by the translocase of the inner mitochondrial membrane (TIM) complexes. **Presequence-containing proteins** are transported by TIM23 in a membrane protential-dependent manner ($\Delta\psi$). TIM23 can either insert proteins directly into the inner membrane or hand them to the presequence translocase-associated motor (PAM) for import into the matrix via

ATP hydrolysis. Once exposed to the matrix, presequences are proteolitically removed by the mitochondrial processing peptidase (MPP). (*2*, *14*, *15*)

Once in the IMS, **non-cleavable precursors** of inner membrane mitochondrial carrier proteins are led by small TIM chaperones to the TIM22 complex, which inserts them into the membrane, dependent on potential ($\Delta\psi$). The small TIM proteins also support transfer of precursor proteins to the sorting and assembly machinery (SAM) complex. SAM then inserts **precursors with β-barrel fold** in the outer membrane (*2*, *14*, *15*). **Proteins with cystein motifs** are imported by TOM complex and handed into the mitochondrial intermembrane space import and assembly (MIA) machine. MIA functions as an oxidoreductase and mediates the formation of disulfide bonds into IMS proteins. Finally, with help of TOM receptors, **outer membrane α-helical proteins** avoid translocation by the TOM complex and are inserted directly into the outer membrane by the mitochondrial import machinery (MIM) (*2*, *14*, *15*).

1.1.3 Translocase of the outer membrane

The main gate into mitochondria has been thoroughly studied in fungi and human cells (*16*, *17*). TOM is a membrane-embedded protein complex with hydrophilic receptor domains facing the cytosol. These receptors detect mitochondrially-targeted precursor proteins and lead them into the translocation pores of the complex. Seven distinct subunits have been identified as part of the TOM complex, originally named after their approximate molecular masses: three receptors Tom70, Tom22 and Tom20, an import pore Tom40, and three small structural subunits Tom7, Tom6 and Tom5 (*18*).

The translocation mechanisms of TOM complex have been extensively studied by electrophysiological, biochemical and structural methods. In earlier years, electrophysiology investigations demonstrated channel-forming activity of the complex after reconstitution into lipid membranes, under application of potential at various voltages (*18*, *19*). The channels showed three levels of conductance: open, half open and closed (Figure 1.4). Compared to VDAC, TOM differed in ion permeability (*18*).

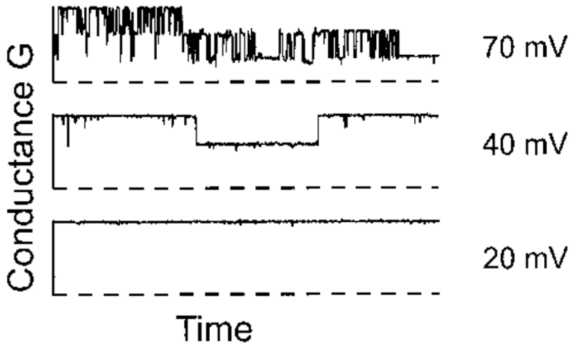

Figure 1.4: **Channel activity of TOM complex.** Electrophysiology studies on TOM complex in a lipidic membrane reveal three distinct levels of conductance: from top to bottom, open, half open and closed. Taken from Kunkele et al., 1998.

Precursor protein translocation

The size of the translocation pore was first characterized using electrophysiology in presence of differently sized polymer probes, and was reported to be blocked by molecules up of 6 kDa (*19*). Protease studies of the complex in vesicles demonstrated that during import of precursors, TOM is capable of transporting only the presequences along the channel, but unable to complete translocation of mature proteins (*20*).

Other studeis have shown that when recombinantly expressed precursor proteins containing a presequence, incubated with the isolated TOM complex, they co-elute in size exclussion chromatography (*19*). The binding affinity of the recombinant precursor protein pSu9-DHFR, formed by the presequence of ATP synthase subunit 9 (pSu9) and dihydrofolate reductase (DHFR), was determined by fluorescence correlation spectroscopy to be between 1.4 to 3.4 nM. Titration studies determined that approximately eight bound precursor proteins were required to saturate the TOM complex (*21*).

Further translocation studies have analyzed contact sites between TOM and TIM23 complexes. Using cryoET, Gold et al., (2014) studied import of matrix targeted pCb2-DHFR through both mitochondrial membranes (Figure 1.5). The recombinant precursor protein was labeled with a quantum dot for easy localization, and arrested in mid-translocation prior to freezing, resulting in the visualisation of clusters of TOM distributed on the outer membrane (*22*). Interestingly, these TOM-TIM23 clusters

were consistently found 30-60 nm away from cristae junctions. Subtomogram averaging revealed protein densities spanning the IMS and protruding out of the outer membrane in close proximity to the quantum dots. Figure 1.5B shows a representative tomogram slice and its corresponding reconstruction, highlighting two outer-membrane densities accompanied by two elongated densities in the IMS.

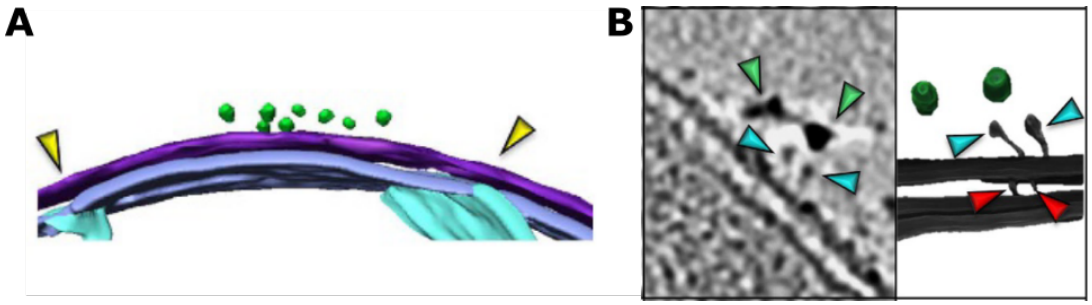

Figure 1.5: **Protein translocation observed by cryoET.** Precursor proteins arrested mid-translocation by TOM and TIM23 complexes. (A) Quantum dots (green) attached to precursor proteins were found in clusters 30-60 nm from cristae junctions (yellow arrows). (B) Tomogram slices (left) and volume segmentation (right) revealed protein densities spanning from outer membrane (blue arrows) to the IMS (red arrows) of mitochondria. Adapted from Gold et al., 2014.

TOM core complex

TOM core complex has been widely studied in *Neurospora crassa* (*18*, *19*, *23*). It is formed formed by the five subunits Tom40, Tom22, and the small Toms (sT), Tom5, Tom6 and Tom7. TOM core has been characterized as a symmetrical dimmer with length of 130 Å by 100 Å, spanning the outer membrane with height of roughly 30 Å (*23*). The total molecular mass of the dimer is around 150 kDa (*24*). Figure 1.6 shows a 6.8 Å resolution cryoEM structure of the core complex obtained by Bausewein et al. (2017), which allowed the first identification and assignment of its individual subunits.

Tucker et al., (2019) and Araiso et al., (2019), reported higher resolution structures of the core complex from *Saccharomyces cerevisiae* (*25*, *26*). Two maps of ScTOM, with resolutions between 3-4 Å, showed Tom40's β-strands, plus the transmembrane

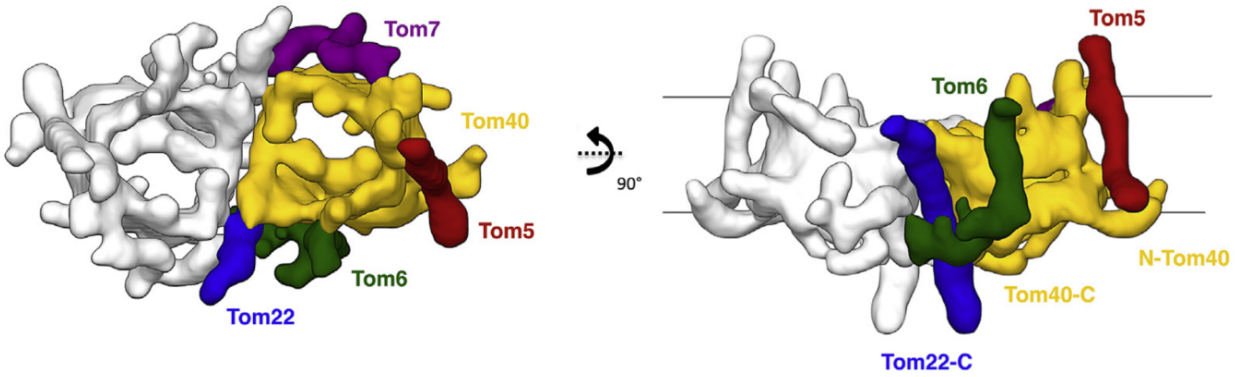

Figure 1.6: **Structure of TOM core complex.** First cryoEM map of TOM core complex structure from *Neurospora crassa*. The complex is shown as seen from the cytoplasm and the side. Tom40, yellow, Tom22, blue, Tom7, violet, Tom6, green and Tom5, red. N and C indicate the N and C-termini of Tom40 and Tom22. Adapted from Bausewein et al., 2017.

helical domains of Tom22, and the sTs, resulting in the first atomic models of the TOM core complex (Figure 1.7A). Later, two structures of TOM core complex from *Homo sapiens* were published by Wang et al., (2020) and Guan et al., (2021) at 3-3.5 Å resolution (Figure 1.7B) (*27*, *28*).

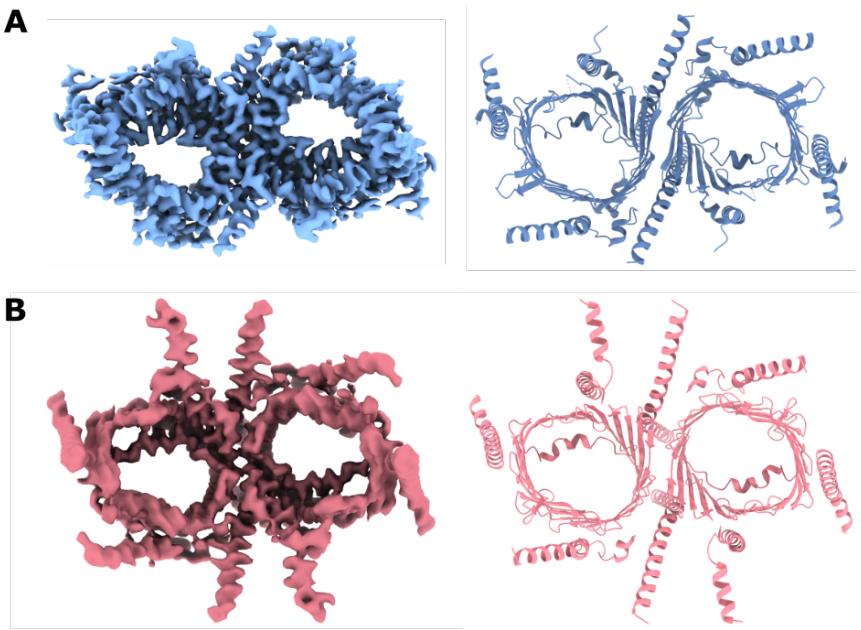

Figure 1.7: **Yeast and human TOM core complex.** CryoEM maps and atomic models of TOM core complex as viewed from the cytoplasmic side of A) yeast (EMDB-20728, PDB 6UCU) and B) human (EMDB-30382, PDB 7CK6) mitochondria.

In addition, Tucker et al., 2019 reported complexes in yeast of larger dimensions than the TOM core dimer. The ∼4 Å resolution structure of tetrameric TOM complex shows a dimer of two dimeric TOM core complexes (Figure 1.8). The dimers are arranged in an stepped, parallel fashion, allowing possible assembly of even higher-order TOM oligomers, but whether these arragements exist in nature is unknown. Tom22, Tom6 and Tom5 are involved in the interaction between the individual dimers. Cluster formation of the complex agrees with Gold's findings. This could lead into higher precursor protein import efficiency, however, tetramers could also be derived from aggregation of isolated proteins in detergent.

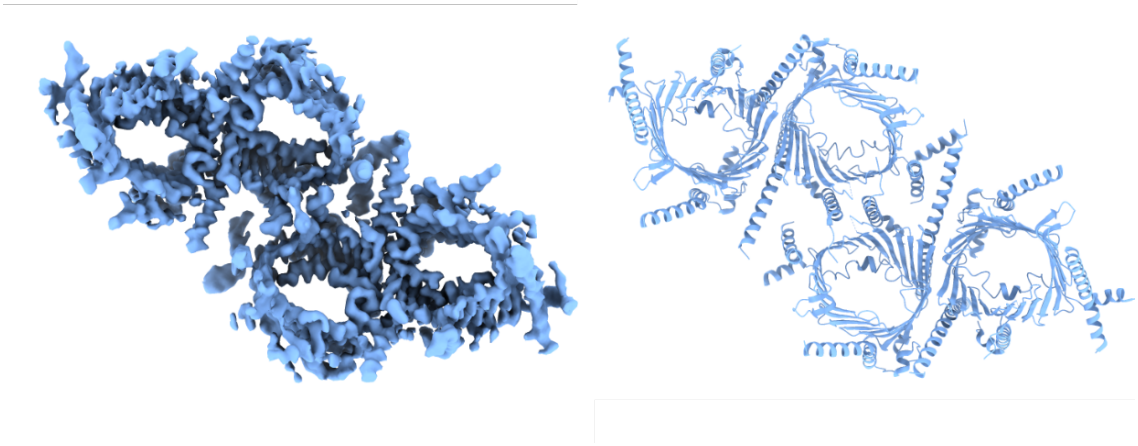

Figure 1.8: **TOM core complex tetramer.** CryoEM map and the corresponding atomic model of tetrameric yeast TOM core complex, viewed from the cytoplasm (EMDB-20729, PDB 6UCV).

A recent study found that freely diffusing TOM core molecules reconstituted in lipid membranes interact with their environment and switch between three states of activity (*29*). Single-molecule tracking revealed that a Ca^{+}-sensitive fluorescent dye (Fluo-8) shone at different intensities in correlation between stop-and-go movement, and open-closed channel activity (Figure 1.9A). The observed fluorescence (Figure 1.9B,C) suggests that the complex has higher mobility when two pores are open, slower mobility when only one pore is open, and no diffusion when both pores are closed.

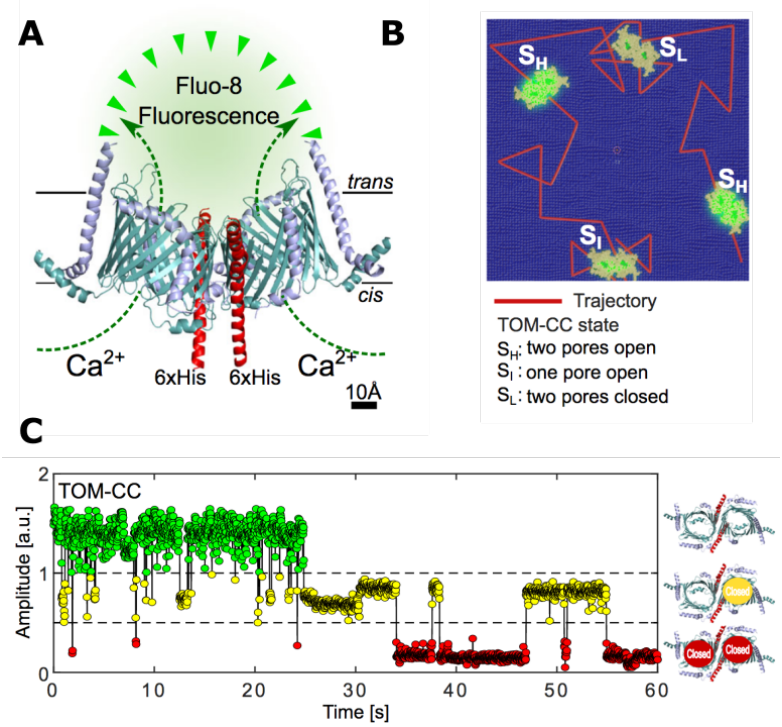

Figure 1.9: **Two-pore channel activity of the TOM core complex.** Correlation between diffusion and channel activity of the TOM core complex in a lipid membrane. (A) Model of TOM core embedded into a membrane separating Ca^{+} ions from a Ca^{+}-sensitive fluorescent dye used for single-molecule tracking. (B) Scheme of the trajectory and channel activity of the complex showing spots of high (S_H), intermediate (S_I), and low (S_L) intensity. (C) Fluorescence amplitude reveals three levels of activity: S_H in green, S_I in yellow, and S_L in red. Adapted from Wang et al., 2022.

The Tom40 translocation pore

At the most basic level, TOM might consist of only the translocation barrel, Tom40. Recombinantly expressed Tom40 is able to form short-lived dimers, suggesting that the remaining TOM subunits are not necessary for dimer formation, but rather for stability (*30*). Similar to VDAC, Tom40 is a β-barrel of 19 β-strands, containing an internal helix (*31*). Additionally, Tom40 has helical segments in both its N and C-termini (*23*). Its N-terminal helix interacts with Tom5 in the IMS domain and is suggested to help transfer incoming presequences to small TIM chaperones for further translocation into the matrix by the TIM complexes (*32*). After β-strand 19, the C-terminal helix of Tom40 is located at the IMS, ending with an extended loop that interacts with the inner pore. This helix is thought to help thread precursor proteins through the barrel (Figure 1.10).

With an internal opening of roughly 13 Å by 19 Å, Tom40 is large enough to fit one or two helices at once along its vertical translocation axis (*23*, *25*). The oblong shape of the pore is likely determined by its internal N-terminal helix. On the inner wall, a line of six highly conserved residues crosses the pore from the cytoplasm to the IMS, forming an anchor for the internal helix and restricting its mobility (*31*, *33*). In contrast, VDAC does not have an equivalent polar patch, which strongly suggests that the patch in TOM crucial for presequence translocation (*31*).

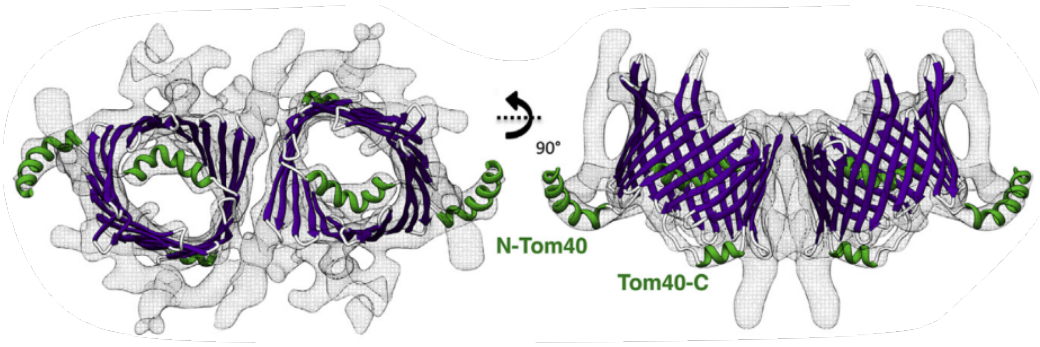

Figure 1.10: **Tom40: translocation pore.** Model of Tom40, based the structure og the *N. crassa* TOM core complex. 19 β-strands are highlighted in purple, and helical segments in green. Taken from Bausewein et al., (2017).

The inner pore contains multiple acidic patches, giving it an overall negative potential, increasingly negative towards the IMS (*25*). This is thought to promote protein translocation, by attracting positively charged residues in presequences. In the dimer, two Tom40s interact only at the cytoplasmic side, close to the pore seam where β-strands 1 and 19 meet (*23*). They interact at a 20° angle relative to the vertical axis, which would result in a slight curvature of the outer membrane.

Tom22: main receptor

Tom22 has an essential role in the correct function of TOM complex. Lack of Tom22 is lethal, as it causes the dissociation of the complex (*34*). Two copies of Tom22 stabilize the interaction between two Tom40s (*23*). Besides its role as receptor, the presence of Tom22 appears to regulate the likelihood of Tom40 to be in its open state, although the exact gating mechanism remains unknown (*29*, *34*).

Tom22 is anchored to the outer membrane by a hydrophobic transmembrane segment, flanked by two exposed acidic termini (*35*). Its N-terminal, facing the cytoplasm, is highly negatively charged and serves as a receptor for incoming proteins. Its C-terminal, in the IMS, is thought to anchor precursor proteins as they exit the import channel and serve as a contact site for TIM23 (*23*, *34*, *36*).

The small Toms Tom7, 6 & 5

Three smaller TOM subunits have been identified: Tom7, Tom6 and Tom5. Their roles are not yet fully understood, but they appear to be involved presequence recognition, and in assembly and disassembly of the complex (*35*). Tom6 and Tom5 are suggested to promote tetramer formation in the core complex (*25*). Tom7 and Tom6 seem to have opposite roles, as Tom6 appears to help stabilize the receptors in the cytoplasmic side, while Tom7 supports the dissociation of Tom22 (*37*).

Tom7 interacts with strands 3 to 6 of Tom40 and has an unusual topology (*23*, *25*). Its N-terminal forms a helix facing the cytoplasm, while its C-terminal is partially unstructured on the IMS side of the outer membrane. This unstructured end is stabilized by hydrogen bonds between Tom7 and Tom40. In *S. cerevisiae*, cells lacking Tom7 were viable but had a reduced growth rate (*37*). The same study also determined that Tom7 is involved in accumulation of precursor proteins in the outer membrane, suggesting that Tom7 forms part of a presequence sorting site prior to translocation.

Tom6 is positioned next to strands 13 to 15 of Tom40 and to one side of Tom22. It has been proposed to modulate cooperation between receptors, mainly of Tom22 and the general import pore (*23*, *38*). It has a C-terminal anchor and its N-terminal faces the cytosol. Lastly, Tom5 interacts with strands 10 and 11 of Tom40. Its N-terminal has a negative net charge, making it relevant for protein import. It is proposed that this domain might serve as a docking platform for incoming sequences and to help their insertion into the import pore. On the IMS side of the membrane, the C-terminal of Tom5 functions as an anchor and holds Tom40's N-terminal helix in place (*23*, *39*).

TOM holo complex

The TOM core complex is considered to be the simplest form of the translocase, a fragment of the more intricate TOM holo complex. With two additional receptors, Tom20 and Tom70, the holo complex has proven to be a difficult target of study (Figure1.11). These subunits easily dissociate from the core complex even in presence of even mild detergents (*19*). In blue native electrophoresis of wild-type mitochondria, Tom20 and Tom70 have been found to be released from the core complex (*34*). The soluble domains of Tom70 and Tom20 increase precursor protein binding, but they are not essential for cell viability (*19*). There is no consensus on the exact position of these receptors relative to the core subunits. However, Tom22 has been suggested as possible docking platform of Tom70 and Tom20 (*34*). Moreover, genetic depleting of Tom7 has proven an effect in the dissociation of Tom20 and Tom22 from Tom40, but that no effect on Tom70 (*37*).

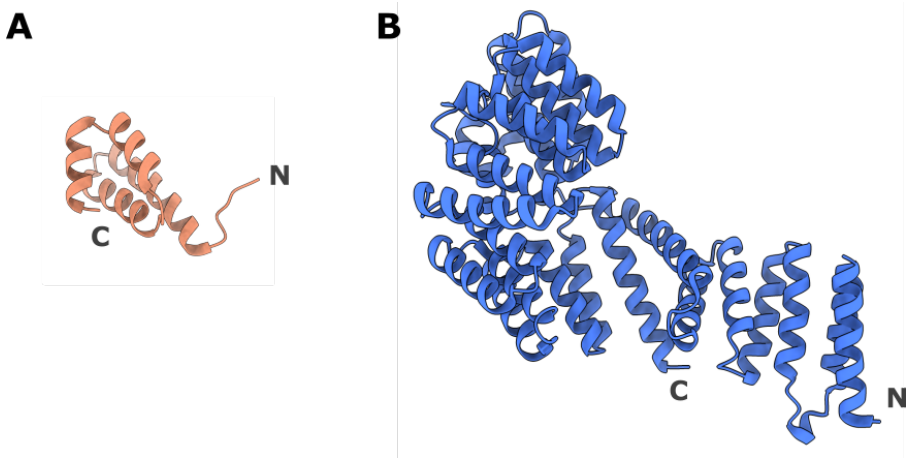

Figure 1.11: **Tom20 and Tom70 receptor domains.** A) Crystal structure of the presequence binding pocket of Tom20 (Obita et al., 2007, PDB 2V1T). B) CryoEM structure of the receptor domain of Tom70 (Gordon et al., 2020, PDB 7KDT). N and C-termini of the sequences are denoted by N and C respectively.

Earlier structural studies in *S. cerevisiae* and *N. crassa* resulted in lower resolution maps of TOM complex, containing two or three translocation pores (*18*, *40*). A possible explanation of these different oligomeric structures is the severance of Tom70 and Tom20 in part of the population. Figure 1.12 shows two structures of TOM holo suggesting a third translocation pore.

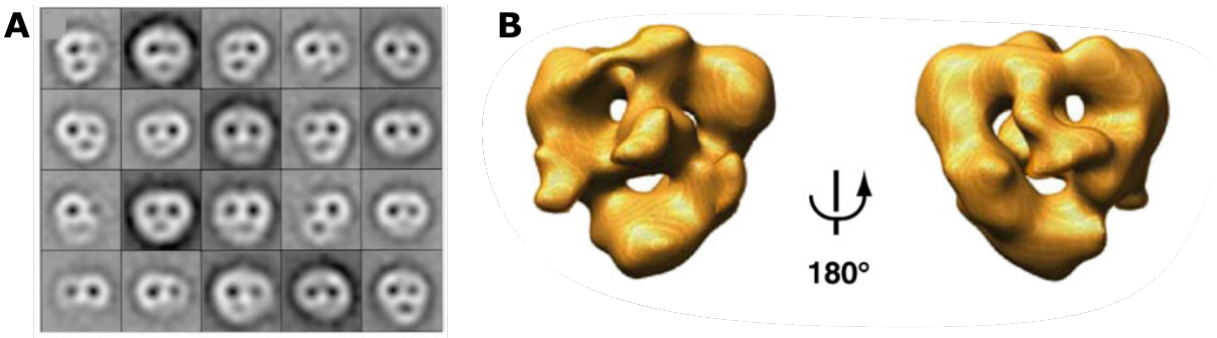

Figure 1.12: **Structure of the TOM complex trimer.** Electron-microscopy structures of the TOM holo complex. A) Negative stained images of the TOM holo complex from *N. crassa*. Taken from Künkele et al. 1998. B) cryoEM map of the TOM holo complex from *S. cerevisiae*. Taken from Model et al. 2008.

The existence of a TOM complex trimer is backed up by site-specific crosslinking studies (*41*). These suggest rearrangement of subunits between the dimer and trimer, with Tom22 rotating 90° on its axis in order to interact with the third pore, stabilizing the trimer conformation (*42*). Interestingly, according to these studies, Tom20 interacts with two hairpin loops of Tom40 on the cytoplasmic side, but not with its transmembrane region (Figure 1.13).

There is no consensus on the stoichiometric composition of TOM holo complex. A recent publication showed the cryoEM structure of TOM core complex crosslinked with Tom20. This study suggested the presence of two copies of Tom20 per TOM core dimer (*43*). These different oligomeric and subunit conformations of TOM complex may be related to the many roles that the complex plays in protein import. Every distinct assembly form of TOM might take part in a different precursor protein import pathway (*42*).

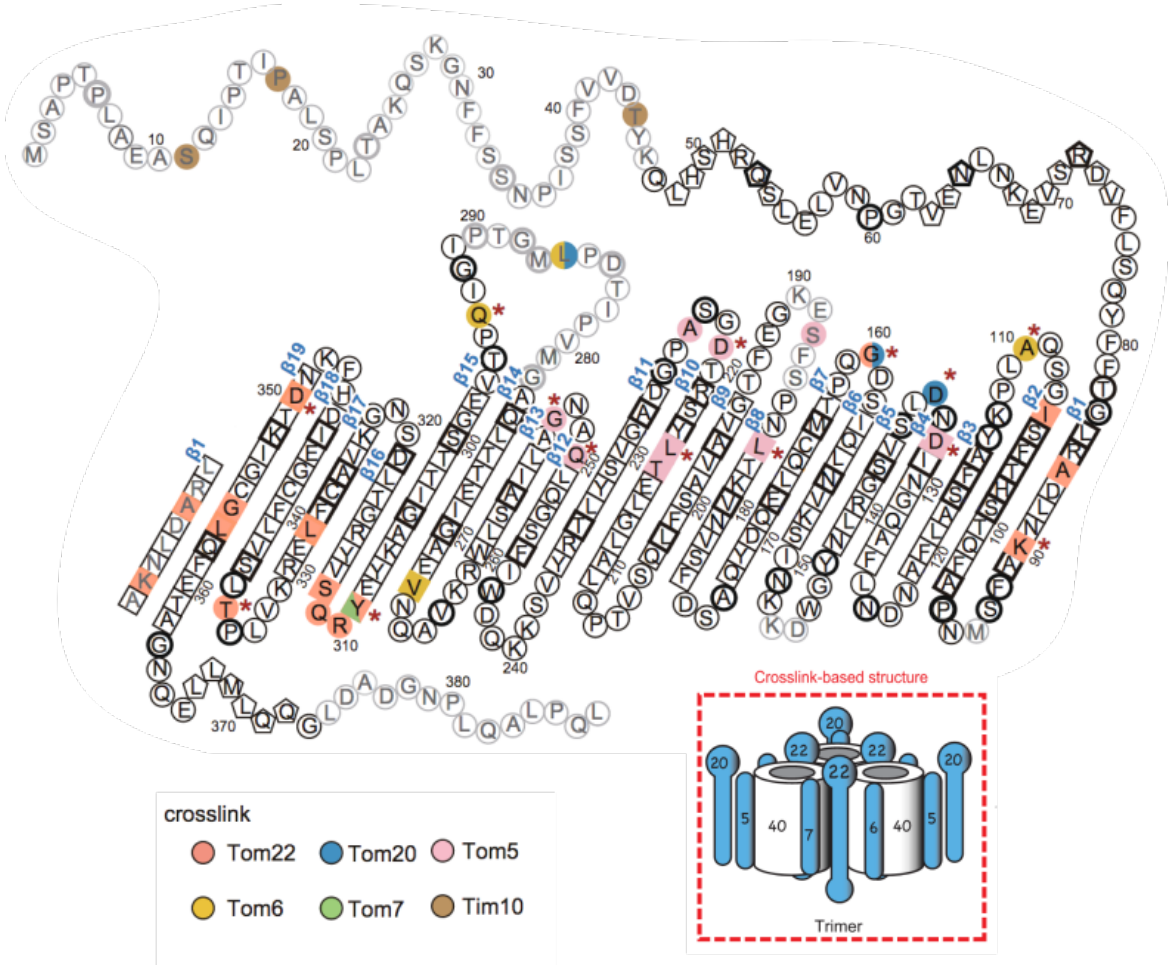

Figure 1.13: **Crosslinking studies of Tom40 suggest a trimeric configuration.** Site-specific crosslinking revealed close interactions between Tom40 and other TOM subunits. A cartoon model based on these interactions is shown. Adapted from Araiso et al., 2019.

Tom70

With its soluble domain measured at ∼100 Å long, Tom70 is the largest mitochondrial protein receptor in TOM (*44*). It is anchored to the outer membrane by an N-terminal helix and has a large C-terminal soluble domain in the cytosol, connected by a long, unstructured chain. Published structures show only the soluble domain of Tom70: a crystal dimer from yeast (*44*) and two monomeric cryoEM models from human TOM (*45*, *46*). The receptor domain of Tom70 consists of 26 tightly packed helices, forming 11 tetratrico peptide repeat (TPR) motifs, which can be classified in two parts. The first part, composed of helices 1 to 7, interacts with protein-carrying-chaperones. The second part, formed by helices 8 to 26, forms the precursor protein binding pocket, which binds internal hydrophobic targeting sequences (*44*, *47*). Künkele et al., (1998)

reported purifying the holo complex, including Tom70, in one single peak, indicating its stable association with other TOM components (*18*). However, they noted that Tom70 tends to fall apart more easily during the purification of the complex from mitochondria rather than from outer membrane vesicle (OMV), suggesting an abundance of Tom70 contact sites between the outer and inner membranes.

Tom20

Tom20 has a transmembrane N-terminal domain and a soluble C-terminal domain. Its cytosolic domain contains five α-helices, including a TPR motif involved in protein-protein interactions (*48*). Through NMR and X-Ray crystallography, two structures have been determined for the presequence bound rat Tom20 (*48*, *49*). It was found that the presequence was bound by hydrophobic interactions, and its multiple locations near the the TPR motif indicated its mobility within the Tom20 binding groove.

Despite its easy dissociation from the core complex, Tom20 has was found to be a necessary component for efficient protein translocation. Tom20 and Tom22 cooperate in presequence recognition, guiding them into the translocation pore (*36*, *41*, *50*). Their interaction appears to be mediated by electrostatic force and has been associated to the acidic patch in the C-terminal of Tom22. Site-specific crosslinking of Tom20 and Tom22 contributed to this suggestion, but also revealed that their interaction decreases when a precursor protein is involved. This indicates that Tom22 and the presequence might compete for interaction with Tom20 (*51*). Tom22 and Tom20 offer a large binding area for precursors on the outer membrane (*20*, *51*). It is then possible that when a helical presequence arrives near the mitochondria, Tom22 might recognise its positively charged side, while Tom20 recognizes its hydrophobic side.

1.2 Physical framework

Biology uses microscopy to study cellular life at different scales. Since the 17th century, scientists have been using light microscopy to analyse complex biological materials. There are, however, problems inherent to light microscopy. Its spatial resolution is limited, meaning the minimum distance between two points that appear separate in an image (*52*). The smaller the distance that can be resolved, the higher resolution. By the end of the 19th century, it was found that the wavelength of incoming light is directly proportional to the maximum resolvable distance between two points. This set a resolution limit of visible-light microscopy to about 300 nm, known as the Abbe diffraction limit (*53*). The following years saw a wide variety of methods in microscopy attempting to break its resolution limit (*54*). The use of fluorescent labels in biological samples improved image resolution to near 100 nm. Nowadays, super-resolution microscopy techniques, such as stimulated emission depletion (STED), rely on laser excitation of fluorophores and depletion of surrounding areas to produce images of ∼70 nm resolution.

During the 19th century, scientists discovered X-rays a potential alternative for high-resolution characterisation of materials. Electromagnetic radiation consists of particle-like photons traveling at speed of light. By measuring diffraction of X-ray beams, scientists can produce reconstructions of atoms within crystalls (*55*). In the case of biological samples, X-ray crystallography has allowed the determination of over 170 thousand protein structures (*56*). However, significant effort is required to crystallize a protein, requiring addition of salts, additives and even modifying the protein to the point where it is highly diverged compared to its native state. One century later, NMR offered biochemists another mechanism to study small molecules based on the distance between their atoms. NMR depends on the observable spin of nuclei when subject to a static magnetic field (*57*). This means that in some cases, data collection requires labeling of protein samples with isotopes.

1.2.1 Electron microscopy

Another imaging alternative flourished during the 20th century thanks to the advancement of quantum theory: electron microscopy. The wave-particle duality of matter offered electrons as a new substitute for photons. Electrons can be accelerated by electrical fields. As they travel, their momentum, and hence their energy, is inversely proportional to their wavelength. Electric fields of hundreds of thousands of volts can therefore produce electron beams with wavelengths that are much shorter than the distances between atoms in protein structures (*53*). This technology has resulted in record resolutions of 0.39 Å for material sciences (*58*) and 1.22 Å for a biological sample (*59*).

The process of transmission electron microscopy (TEM) starts with a coherent beam of electrons, meaning electrons with the same frequency and wavelength. The incident electron beam hits the specimen and results in different kinds of scattering. Coherently scattered electrons remain in phase with the incoming beam, while incoherently scattered electrons share no phase relationship (*53*). Scattering events can result in energy transfer to the specimen, known as inelastic scattering (Figure 1.14). Elastic scattering, however, results in no energy loss for the electron. Scattering can also happen in two directions: forwards and backwards. Forward-scattered electrons are used in TEM. The amount of transmitted electrons depends of the thickness of the sample and on how heavy its atoms are: the ticker the specimen, the more electrons get backscattered (*60*, *61*). Nevertheless, TEM comes with its disadvantages (*53*, *61*). Compared to light-microscopy, TEM cannot image living organisms due to the effects of radiation damage. Also, TEM generates 2D projections of 3D objects, which can lead to misinterpretation of artifacts due to a lack of depth perception. Tomography deals with this issue by tilting the specimen during data acquisition, recording a single 3D object from different angles. Finally, a detrimental factor, especially in structural biology, is radiation damage to the sample. Inelastic scattered electrons deposit energy into the specimen, changing it's structure and limiting high-resolution imaging.

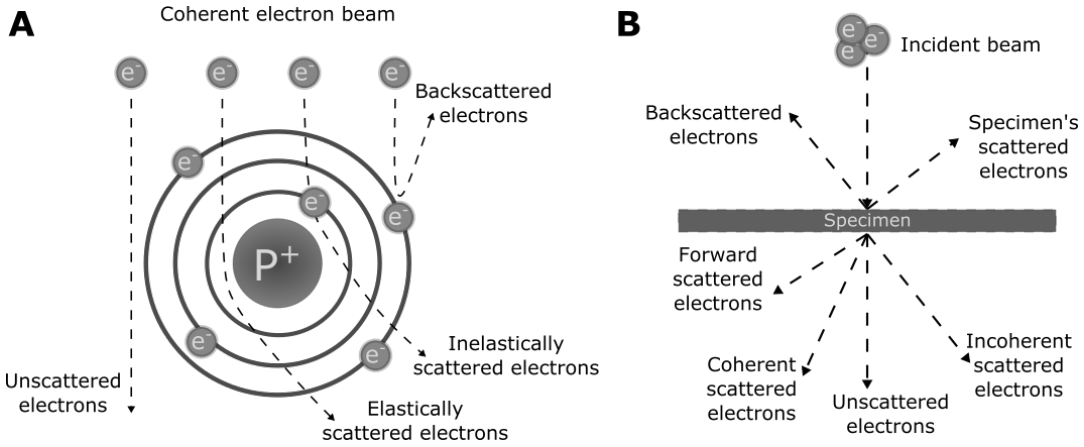

Figure 1.14: **Electron scattering.** Incident electrons can interact with atoms in a sample elastically and inelastically, leading to scattering events. A) Interactions of incident electrons with an atom. B) Different kinds of electron scattering from a thin specimen.

Microscopes

TEMs are equipped with an electron source, a series of lenses, apertures and a detection system. There are two types or electron sources or **electron guns** (*53*). In a **thermionic gun**, a material, such as LaB_6 crystals, is heated to high temperatures to produce electrons. The **field-emission gun (FEG)** produces electrons when subject to an electric potential. FEGs rely on increased strength of an electric field at very sharp points (Figure 1.15). By charging an anode with a high voltage, the electrons are extracted from the tip, while a second anode accelerates the electron beam.

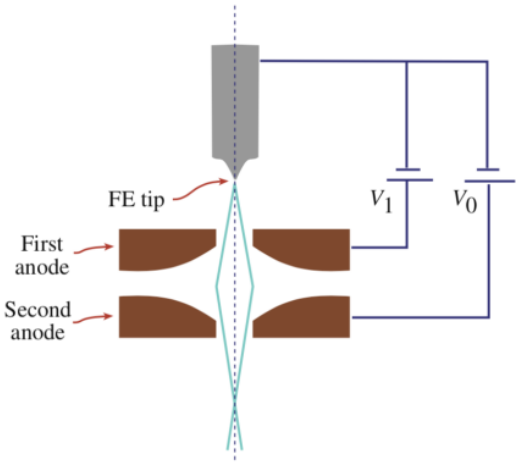

Figure 1.15: **Field-emission electron gun.** FEGs are formed by a very sharp tip that acts as a cathode, and anodes at different voltages that extract electrons from the tip. Taken from Williams and Carter, 2008.

Below the electron gun, a series of magnetic electron lenses deflect the beam within the microscope column (Figure 1.16). They are equivalent to conventional lenses in light microscopy: they focus parallel rays into a single point in a plane, and recreate an object as an image (*53*). The more sophisticated the TEM is, the more lenses it will have. The **condenser lenses** control strength and size of the beam. They can be adjusted to illuminate the specimen with a parallel beam of electrons.

The **objective lens** is close to the specimen and, because it forms the first image of the object, it is the most important lens in the microscope (*60*). The quality of this lens determines the quality of any resulting information and is the main source of aberration. A series of **magnifying lenses** follow the objective lens, amplifying the image to desired size. The **projector lens** is used to align the final image into a viewing screen or detector (*53*). These lenses modify the phase of electron waves behind the specimen, limiting contrast transfer of high-spatial frequencies. This means that electrons with lower energy are strongly deflected, resulting in a blurred disk in the Gaussian image plane (*60*). Usually, the electron beam is not completely parallel. Electrons that are scattered at a higher-than-desired angle can be removed with an **aperture**. The objective aperture is located at the back-focal plane of the objective lens (Figure 1.17), and controls the resolution of the image formed by this lens, its image contrast, and depth of field (*53*, *61*).

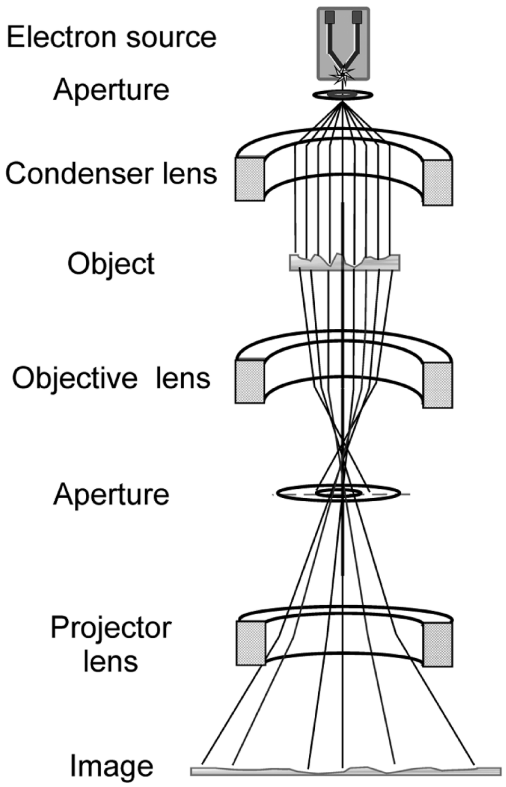

Figure 1.16: **Basic scheme of an electron microscope.** Electron microscopes are composed by an electron source, various lenses and apertures. Taken from Orlova and Saibil, 2011.

Just as glass lenses, magnetic lenses suffer from aberrations that limit the resolution. The most common defect is astigmatism, which derives from contamination, charging up and misalignment of condenser lenses, thus deflecting the beam (*53*). Spherical aberration arises from different strengths at which rays are refracted by lenses. Chromatic aberration depends on energy of the electron beam, in which images are formed at different planes relative to the object (*61*). Spherical and chromatic aberrations are usually mininized by selecting electrons near the optic axis by using apertures. However, this can limit the resolution.

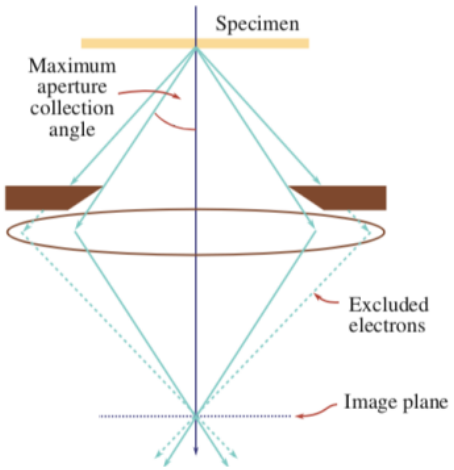

Figure 1.17: **Objective lens and aperture.** The most important lens in a microscope is the image-forming objective lens. Electrons scattered to high angles are removed with an aperture. Taken from Williams and Carter, 2008.

Contrast transfer function

High-resolution imaging requires information of electrons scattered at high-spatial frequencies, since they correspond to the smallest distances within the specimen. Electrons that reach the objective lens at high angles are not focused at the same point as the rest, because of spherical aberration. This results in degradation in image quality and a reduction in contrast of the fine details of the image (*53*). However, spherical aberration also shifts the phase shift between scattered and unscattered electrons, resulting in stronger image contrast (*61*).

Any two-dimensional function can be represented by a sum of sine waves. The Fourier transform expresses a function as a sum of all of its frequencies. Independent of the specimen, we can formulate the final projection of an object as a sum of distortions and focus. This is known as contrast transfer function (CTF) (*60*). The CTF is an oscillatory function with effects that can be seen as alternating Thon rings in the Fourier transform of the final image, depicted as bands of good transmission separated from gaps without transmission (Figure 1.18). Introducing an aperture limits high-frequency information contained in the image, eliminating the outer Thon rings in Fourier space.

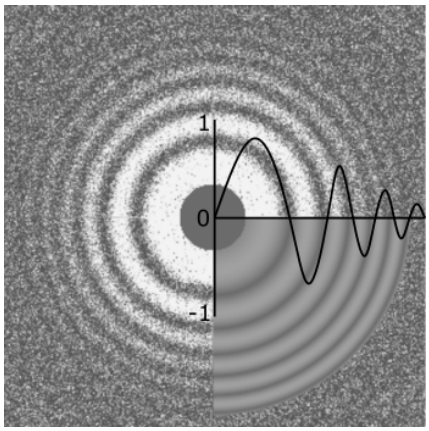

Figure 1.18: **Contrast transfer function.** Applying the Fourier transform to any image in real space reveals its contrast in terms of frequencies in a Thon ring. This can be formulated as the CTF of the image (black).

Detectors

Up until a few years back, photographic film was widely used to record microscope data. Film provided a large amount of "pixels", which allowed recording high-resolution information, although its quality was often reduced during processing and digitizing. Film was replaced by a digital alternative: charge-coupled device (CCD) detectors. CCD detectors are metal-insulator-silicon devices that store charge. They consist of millions of pixels that accumulate charge in proportion to incident radiation intensity (*53*). High-energy electrons damage the photosensitive wells in CCDs, so they use scintillators that convert electrons into photons, which are detected by the CCD chip.

This conversion results in multiple scattering, reducing the spatial resolution that detectors can provide, specially at higher voltages (*61*).

The recently developed direct detectors can withstand high-energy electrons. Direct detectors have smaller pixels, built with a semiconductor that enables local conversion from charge to voltage, and faster readout (*61*). This change in detection improved signal-to-noise ratio in comparison to CCDs. Because of faster readout, dozens of images of the same area can be taken in rapid succession, meaning that beam-induced movements can be easily detected and corrected during image processing (*62*).

1.2.2 Negative-stain electron microscopy

In negative-stain electron microscopy, isolated particles, such as proteins and macromolecules, are embedded in a heavy salt solution prior to imaging (*61*). Figure 1.19 shows a sample preparation scheme. Particles suspended in solution are deposited onto a continuous carbon support film on a copper gid. While excess liquid is blotted away, particles adsorb to the carbon layer. The sample is washed with water, and the grid is then covered with staining solution, commonly uranyl formate and uranyl acetate, blotted and dried.

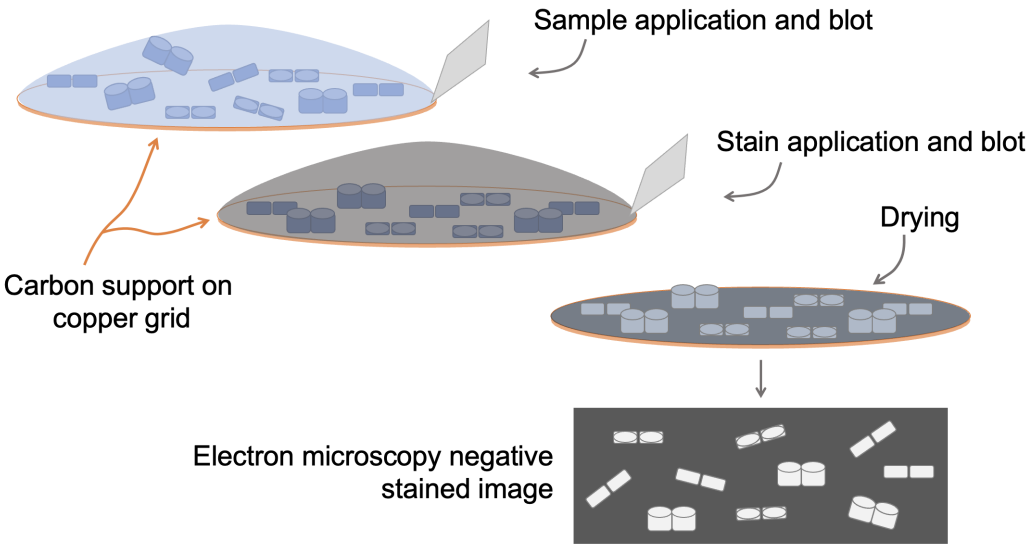

Figure 1.19: **Negative stain sample preparation.** Macromolecules are deposited on a carbon support film on a copper grid, and then stained using a heavy salt solution. Staining of isolated particles for electron microscopy results in high-contrast, low-resolution images.

Heavy-metal stains bind to the carbon support and outline the surface of particles, resulting in a dark background. During imaging, incident electrons are scattered by heavy atoms in the stain. Unscattered and forward-scattered electrons can reach the detector through the particles, generating negative images, where projections of macromolecules are visible by exclusion of stain (*61*). This method provides valuable information about size and shape of biological assemblies, as well as providing quality evaluation of samples prior to cryoEM. However, it cannot provide high-resolution information of particles, and tends to flatten or distort fragile assemblies.

1.2.3 Single-particle electron cryo-microscopy

In electron cryo-microscopy of isolated assemblies, macromolecules are frozen in vitrified ice and imaged at cryogenic temperatures. Ice protects biological assemblies from vacuum and the low temperature mitigates radiation damage (*63*). Crystalline ice expands and can alter the structure of target proteins, but vitrified ice forms a non-crystalline solid phase, in which proteins are deposited in a wide range of orientations (*64*, *65*). Figure 1.20 shows sample preparation for cryoEM.

A few microliters of protein in solution are applied to a grid lined with holey carbon support. Excess liquid is blotted away and the grid is plunge-frozen in liquid ethane (*65*). Samples are vitrified in liquid ethane at -182°C rather than in colder liquid nitrogen to avoid the Leidenfrost effect, which results in slow freezing. Because of its higher heat capacity, liquid ethane does not boil after contact with the sample, and does not produce an insulating gas layer on its surface, resulting in fast, non-crystalline freezing (*61*, *66*). After plunging, grids are quickly transferred into liquid nitrogen, where they are stored until used. Using this method, proteins are preserved in a near-native state, suitable for high-resolution cryoEM imaging.

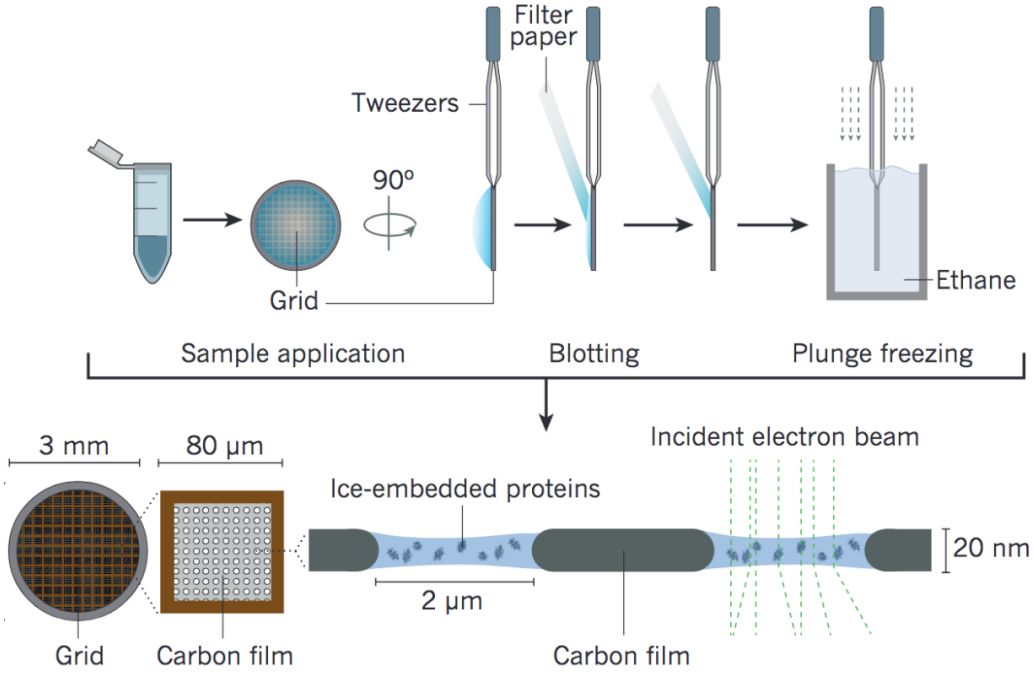

Figure 1.20: **Sample preparation for cryo-EM.** Macromolecules are deposited into a grid lined with a holey carbon support and then plunge frozen in ethane. This results in a thin layer of vitrified ice that protects proteins during imaging. Taken from Fernandez-Leiro et al., 2016.

1.2.4 Mass spectrometry

Another method frequently used in structural biology is mass spectrometry (MS). The principle of MS is generating ions and separating them by their atomic mass per number of elementary charges. Ions are detected and plotted according to their mass-to-charge ratio (m/z) and abundance (*67*). When only single-charged molecules are observed, m/z scale reflects its atomic mass. However, in some conditions, charges can double or triple, depending on the ionization method.

Laser-induced liquid bead ion desorption

Native-MS refers to the MS methods that employ electrospray ionization in nondenaturing solvents to analyze biological assemblies (*68*). A promising native-MS technique that was recently developed for the study of membrane proteins is laser-induced liquid bead ion-desorption (LILBID). LILBID employs a laser to release ions from a sample in aqueous solution, and can reveal the mass-to-charge ratio of intact membrane protein

complexes. Increasing laser power results in dissociation of individual components of a protein complex. Shinning the laser at different intensities into the sample, helps analyse the complex's composition and arrangement of its subunits (*69*).

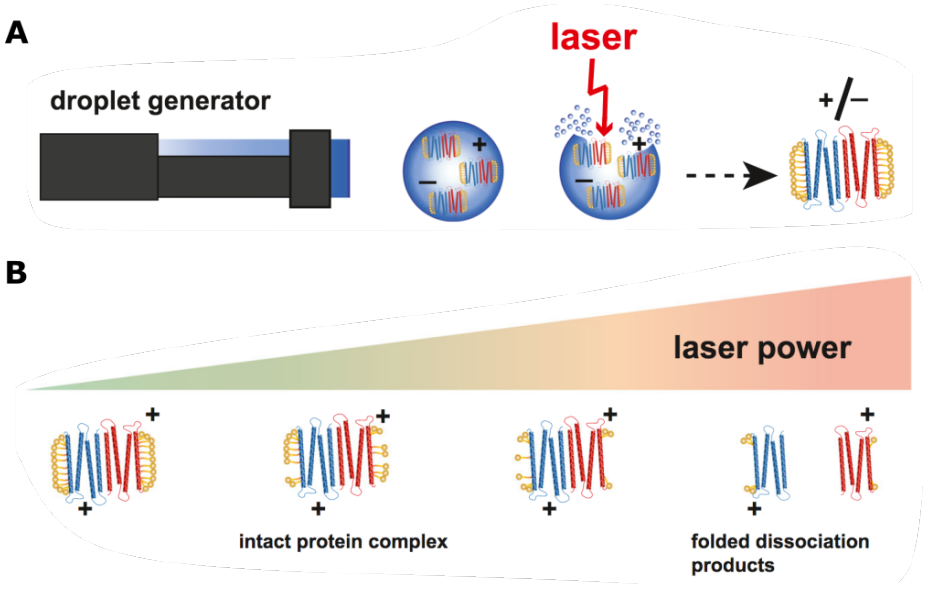

Figure 1.21: **Laser induced liquid bead ion desorption.** A) A laser is employed in LILBID to ionize a protein contained in liquid droplets. B) Laser intensity can be modified to study an entire protein complex or its subunits. Adapted from Peetz et al., 2019.

Crosslink mass spectromety

Crosslinking is a technique commonly used in biochemistry to artificially link two amino acids, thus providing information about protein-protein interactions. Depending on spatial constraints of the crosslinker, crosslinking mass spectrometry (XL-MS) can be used to study protein folding and protein interactions. XL-MS requires proteolytical digestion of crosslinked peptides (*70*). Peptide fragments attached by the linkers are then scanned, analysed and sequenced, revealing inter and intrapeptide interactions of the target proteins.

1.3 Computational framework

1.3.1 Single-particle analysis

In order to generate a high-resolution, three-dimensional structure of a protein by single-particle cryoEM, 2D projections must be pre-processed to compensate for microscopy defects. As mentioned in the previous section, incident electrons can interact inelastically with the specimen, leading to beam-induced sample movement. This, together with stage movement, is motion corrected by collecting several frames in a single area. Motion between frames is traced by considering specific patches or whole frames (*71*). This generates a single image, called a micrograph. The micrograph CTF is then estimated, making it possible to correct for lens aberrations during downstream processing and to measure defoci (*72*). Once this has been applied to every micrograph in the data set, coordinate positions of proteins 'single-particles' are ready to be selected and extracted (Figure 1.22).

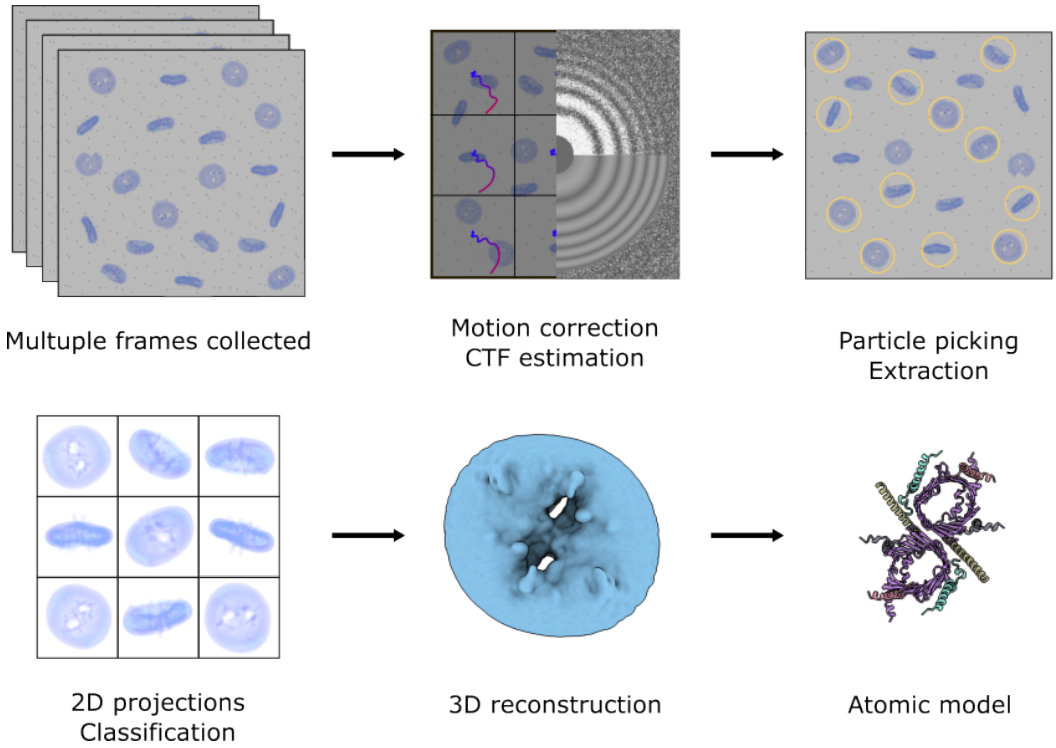

Figure 1.22: **Computational single-particle workflow.** High-resolution single-particle cryoEM involves image preprocessing, including motion correction and CTF estimation. Particles are picked and extracted from micrographs, then subjected to 2D and 3D classification, followed by refinement. The final cryoEM map can be used for model building.

Individual particles are centered in a box of specified size for extraction. Since these particles are projections of each protein contained in the sample, they are present in many orientations and must be classified and aligned. In reference-free 2D alignment, the particle data set is compared on basis of correlations between images (*61*). Most alignment information comes from low-frequency components of the images, although they can carry a lot of noise in their background. This noise can be attenuated by the use of masks and band-pass filters. Projections corresponding to same in-plane orientations are clustered together. Different classes are then used as references to bring similar images to the same position by cross-correlation. (*73*).

Three-dimensional reconstructions are generated from 2D projections. According to the projection theorem (Figure 1.23), the 2D projection of a 3D object in real space is equivalent to a central 2D section through the 3D Fourier transform of that object (*73*). This means that with multiple projections in various orientations, the original 3D structure can be computed from the inverse Fourier transform. A sufficiently large data set will prevent artifacts due to missing angular coverage. Multiple iterations will improve resolution during reconstruction by finding more suitable orientations (*74*).

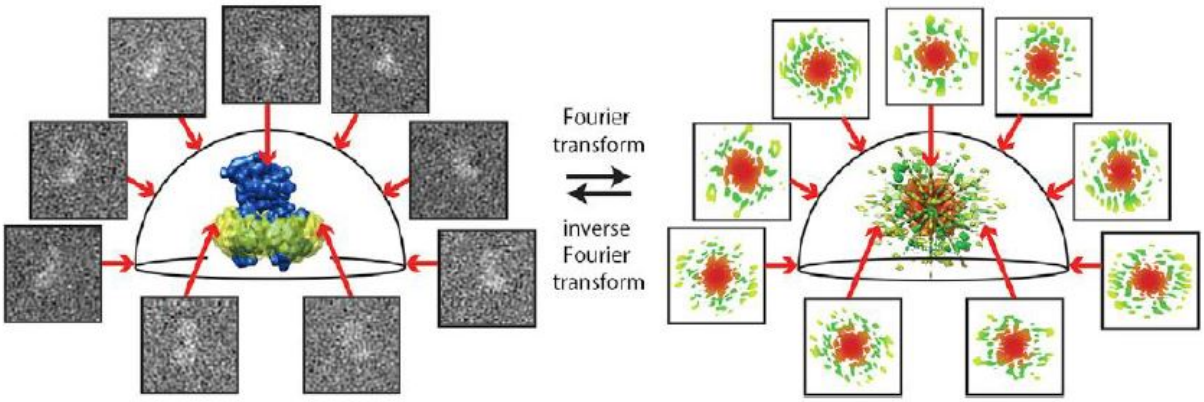

Figure 1.23: **Projection-slice theorem.** The projection-slice theorem states that a real space 3D reconstruction can be calculated from the inverse Fourier transform of aligned 2D projections of a protein in the Fourier space. Taken from Nogales et al., 2015.

An initial 3D reconstruction can be used to further classify particles and clean the data set from false positives. Through the iterative process of projection matching, the best matches between projections and a 3D reference can be selected for each given orientation. An improved cryoEM map is obtained after refinement, taking in consideration higher-frequency signals for alignment. Final particles contained in the map should be corrected from microscope aberrations using the initially estimated CTF values. This is followed by a measurement of the map's resolution through Fourier shell correlation (FSC). FSC curves are calculated by separating the particle stack in two halves, reconstructing a map for each half, and comparing their Fourier transforms by cross-correlation. Resolution is measured at the point where the correlation drops below 0.5 or 0.143, and reported accordingly (*61*).

1.3.2 Modelling

Depending on the resolution of the final cryoEM map, it might be possible to fit in an existing atomic model or to create one based on the protein sequence. Currently, the highest resolution cryoEM map has a resolution of 1.22 Å, which allows visualization of individual atoms in the best-resolved regions (*59*). However, this is not common, especially when it comes to membrane proteins. In either case, when a peptide sequence is known, prediction software can be used to assign amino acid chains to a map, based on secondary protein structure. Furthermore, modern software, such as AlphaFold, relies on deep learning techniques to predict tertiary and quaternary protein structure, based on statistical analysis of publicly available atomic models (*75*). Using model building software, protein structures can be built *de novo* or based on an existing structure, while considering angular restrictions of residues, steric clashes, electromagnetic forces governing atomic interactions and more (*76*, *77*).

Chapter 2

Motivation

In mammals, only thirteen mitochondrial proteins are encoded by mitochondrial DNA, including key components of the oxidative phosphorylation machinery (*78*). In contrast the vast majority of mitochondrial proteins are encoded in the nuclear DNA, requiring import into mitochondria. Understanding the translocation mechanism that facilitates import of nuclear-encoded mitochondrial proteins is fundamental to further understand mitochondrial function and disease. In humans, impaired mitochondrial function can lead to a wide variety of pathologies (*78*, *79*). It has been reported, for example, that translocases are often upregulated in cancer cells. Increased activity of protein translocation can result in imbalanced mitochondrial homeostasis and affect cellular response to stimuli. Mutations in proteins related to programmed cell death can lead to sequestration of the translocating system, hiding important signals of cellular stress and helping cancer cells to avoid apoptosis. Another example is deafness dystonia syndrome, a genetically inherited syndrome related to both TIM translocases, which leads to hearing loss and blindness.

More specifically, impairment of TOM complex in yeast has been linked in multiple studies to cell malfunction. Depletion of Tom40 leads to non-viable organisms, while depletion of other TOM subunits hinders cell growth and causes defective cells (*80*). In disease, TOM complex doesn't only act as a mediator, as impaired TOM activity has been associated with neurodegeneration (*78*, *79*). The PINK1-parkin pathway is

responsible for selected autophagy. Under stress conditions, PINK1 forms a complex with TOM to recruit parkin and signal autophagy. However, brain samples from patients suffering from Parkinson's disease show downregulation of Tom40, leading to reduced PINK1 signalling in stressed cells. Another example is Alzheimer's disease, in which an accumulation of amyloid β has been reported to block the TOM complex and to stop import of mitochondria-targeted proteins. Furthermore, it has been reported that α-Synuclein can be imported into mitochondria by TOM, where it targets ATP-synthase and upregulates it (*81*). In some cases, it was observed that α-Synuclein strongly binds Tom20, leading to impaired protein import, and therefore, deficient mitochondrial respiration.

2.1 *Neurospora crassa* and TOM complex

Neurospora crassa is a common type of mold. It is a multicellular, filamentous fungus with bright orange hyphae. Since the early 20th century, *Neurospora crassa* proved to be an effective model organism in the study of eukaryotic biology. Its genome has been fully sequenced and is almost four times larger than that of *S. cerevisiae* (*82*, *83*). Mitochondrial translocases are highly conserved throughout eukaryotic evolution and fungi is therefore often used to infer on mammalian import mechanisms (*78*).

TOM has been studied in several organisms for many years, but only recently was its core structure elucidated. The first structure of *N. crassa* TOM core complex was published by our group in 2017 at roughly 7 Å resolution (*23*). Now, the structure of the core complex at high-resolution is available from *S. cerevisiae* and *H. sapiens* (*25*, *27*, *28*, *35*). However, the high-resolution structure of *N. crassa* TOM holo complex was still missing. The elusive Tom20 and Tom70 subunits remain a mystery. They have been studied independently, and together with TOM core complex, however their stoichiometry within the complex, and mechanism of cooperation with other TOM

subunits is still an enigma. Only recently, a group published a cryoEM structure of TOM core, with the soluble domain of Tom20 crosslinked to Tom40, opening the way for TOM holo structural studies (*43*).

2.2 Aims of this thesis

In this project we aim at a comprehensive investigation of the conformation and structure of TOM complex from *Neurospora crassa*. We use single-particle cryoEM to resolve a high-resolution structure of TOM complex and to build an accurate atomic model that sheds light into the location of Tom70 and Tom20. Coupled with cutting-edge MS techniques, we planned to model interactions between TOM subunits, subcomplex formation and possible oligomerization states.

This project also aims to better understand the import mechanisms of mitochondrial precursor proteins. Structural insights into presequence interactions of TOM complex will enable us to build a model of translocation in agreement with our structures. We anticipate to observe the presequence-bound complex, so that we can analyze binding interactions inside the pore. Finally, the combination of all of the above, together with modern protein structure prediction software, we can propose a translocation mechanism for positively-charged, matrix targeted, precursor proteins as they cross the main gate into mitochondria.

Chapter 3

Methodology

Mitochondria isolation, preparation of outer membrane vesicles and purification of the core complex were carried out by our collaborators Beate Nitschke and Prof. Stephan Nussberger at the University of Stuttgart. LILBID-MS was carried out by Janosch Martin in the laboratory of Prof. Nina Morgner at the Goethe University Frankfurt. XL-MS was carried out by Dr. Julian Langer at the proteomics laboratory of the Max Planck Institute of Biophysics. Isolation of TOM holo complex, sample preparation for MS and cryoEM, image collection, and data processing were perfomed by myself under supervision of Prof. Dr. Werner Kühlbrandt. Visiting student Parinaz Lordifard assisted in preparation of precursor proteins.

3.1 Isolation of mitochondria

Strain GR-107 of *N. crassa* with a hexahistidinyl-tagged form of Tom22 was used for this study (*18*). Hyphae (1.5 kg wet weight) were homogenized with a mixer in buffer containing in 250 mM sucrose, 2 mM EDTA, 20 mM Tris pH 8.5, 1 mM PMSF at 4 °C. Cell walls were mechanically ruptured by passing the suspension, containing 1.5 kg of silica sand, through a corundum stone mill. Debris were pelleted and discarded in two centrifugation steps (4,000 × g) for 5 min at 4 °C. Mitochondria were sedimented twice

in 250 mM sucrose, 2 mM EDTA, 20 mM Tris pH 8.5, 1 mM PMSF at 17,000 × g for 80 min. Isolated mitochondria were suspended in 250 mM sucrose, 20 mM Tris pH 8.5, 1 mM PMSF to final protein concentration of 50 mg/ml.

3.2 Isolation and purification of TOM core complex

Mitochondria were solubilized in 20% glycerol, 20 mM Tris pH 8.5, 300 mM NaCl, 1 mM PMSF, 20 mM imidazole and 1% DDM (Glycon) for 30 minutes. Insoluble material was pelleted at 130,000 × g and the supernatant was loaded into an equilibrated Nickel-nitrilotriacetic acid (NiNTA) column (Cytiva). TOM core complex was isolated from mitochondria via immobilized-metal affinity chromatography (IMAC) through a His-tag at the C-terminus of Tom22 (*29*). The column was washed with the same buffer, containing 0.1% DDM. Bound TOM was eluted with 0.1% DDM, 10% glycerol, 20 mM HEPES pH 7.2, 1 mM PMSF and 300 mM imidazole. Glycerol was removed by anion exchange chromatography using a Resource-Q column (Cytiva). The complex was eluted with a 0–500 mM KCl gradient (Figure 3.1).

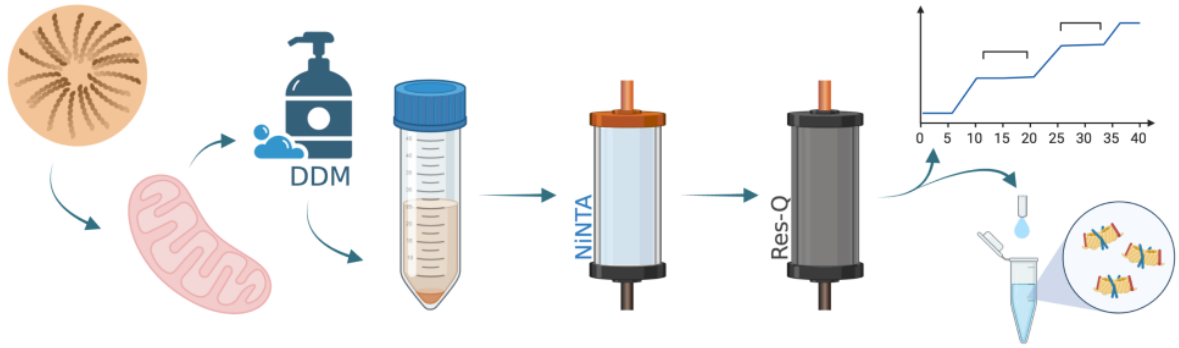

Figure 3.1: **Purification of TOM core complex.** *N. crassa* hyphae were grown at the laboratory of Prof. Stephan Nussberger. Mitochondria were isolated and solubilized in detergent. TOM core complex was purified using ion affinity (NiNTA) and anion exchange (Res-Q) chromatography. Created with BioRender.

3.3 Preparation of outer membrane vesicles

Isolated *N. crassa* mitochondria were used to prepare OMVs (*18*). Mitochondria were swollen in 5 mM Tris pH 8.5, 1 mM EDTA, and 1 mM PMSF. Inner and outer membranes were separated from soluble proteins by centrifugation at 17,700 × g. Pellets were resuspended in the same buffer and homogenized in an automated glass-Teflon douncer for 60 min at 4°C to separate the outer membrane (Figure 3.2).

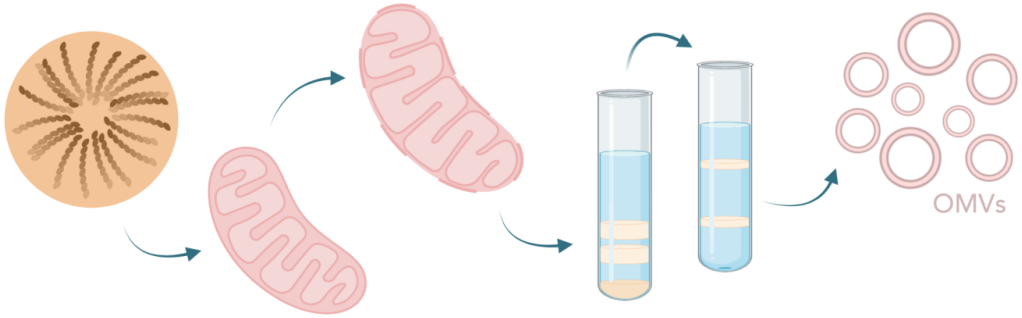

Figure 3.2: **Preparation of outer membrane vesicles.** *N. crassa* OMVs were prepared at the laboratory of Prof. Stephan Nussberger. Mitochondria were swollen, and their membranes separated through mechanical homogenization. OMVs were isolated by centrifugation. Created with BioRender.

Vesicles were isolated by sedimentation and flotation centrifugation of the homogenate in two sucrose step gradients, using 20 mM Tris pH 8.5, 1 mM EDTA and 1 mM PMSF. The outer membrane fraction was then diluted with 50 mM Tris pH 8.5, sedimented by centrifugation at 250,000 × g, and resuspended in 50 mM Tris pH 8.5 with a protein concentration of ∼1 mg/ml.

3.4 Isolation and purification of TOM holo complex

TOM holo complex was isolated by solubilizing OMVs in 20% glycerol, 10 mM MOPS pH 7.0, 50 mM KAc, 50 mM imidazole, 1 mM PMSF and 1% GDN. After 1 h incubation at 4°C, the lysate was centrifuged at 13,000 x g for 20 min. The supernatant was loaded onto a NiNTA gravity column (Figure 3.3). The column was washed with

the same buffer, containing 0.02% GDN. Bound TOM was eluted with the same buffer, containing 300 mM imidazole, and concentrated using an AmiconUltra 100 kDa cutoff tube.

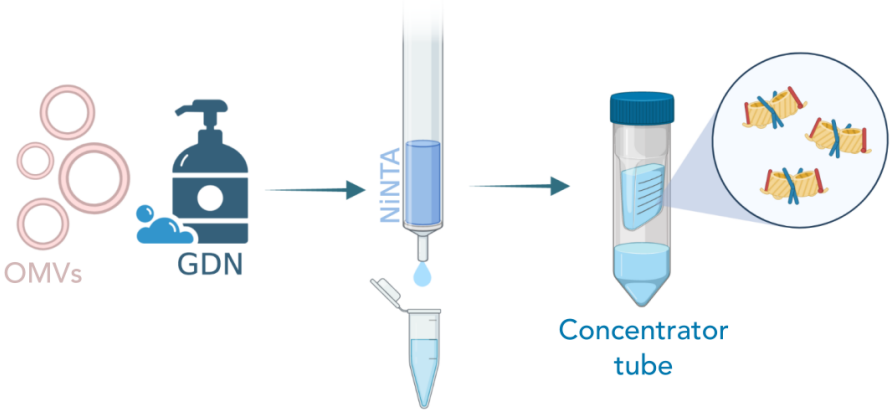

Figure 3.3: **Purification of TOM holo complex.** OMVs were solubilized in detergent and centrifuged. An ion affinity (NiNTA) column was used to purify the soluble fraction. Created with BioRender.

3.5 Nano-differential scanning fluorimetry

Stability of TOM core and holo complexes was tested by nano differential scanning fluorimetry (nanoDSF) using a Prometheus (NanoTemper) device (*84*). Both core and holo concentrated samples were incubated in the new buffer mix at a 1:15 old to new buffer mixture, respectively. Tested buffers were citrate (pH 5, 5.5 and 6), phosphate (pH 6.5, 7 and 7.5), and bicine (pH 8, 8.5 and 9). The buffer mix contained 50 mM buffer, 50 mM KCl, 1 mM TCEP, 1 mM EDTA, and 0.02% GDN (holo) or 0.08% DMM (core). Measurements were made in quadruplicate over 1°C increments ranging from 20°C to 90°C. All four data sets were averaged and processed and plotted using Python3.

3.6 Preparation of precursor proteins

3.6.1 Cloning and purification of pSu9-MBP

N. crassa pSu9(1-67) was amplified by PCR from a pSU9-BC2NB-pUC57 synthetic plasmid (Supplementary material 1) provided by Dr. Thomas Bausewein, with primers GAG ATA TAC CAT GGC TTC TAC CCG TGT TCT GG and CAG ATT ACC AGT TTA CCT TCT TCG ATT TTC ATA TGA CCA GAA GAG TAA GCA CGT TTC TG resulting in a 255 base pair long gene. The forward primer contained an N-terminal NcoI restriction site for cloning, and the reverse primer, a C-terminal serine-serine-glycine (SSG) linker. MBP's gene was amplified by PCR from a pMAL plasmid with primers CAG AAA CGT GCT TAC TCT TCT GGT CAT ATG AAA ATC GAA GAA GGT AAA CTG GTA ATC TG and GCT ATC CTC GAG TTA TTT TTC GAA CTG CGG GTG GCT CCA GCT CGA ATT AGT CTG CGC resulting in a 1206 base pair gene. The reverse primer contained a StrepII-tag for purification and a C-terminal XhoI restriction site for cloning.

pSu9(1-67) and MBP genes were fused by PCR into a single product using primers GAG ATA TAC CAT GGC TTC TAC CCG TGT TCT GG and GCT ATC CTC GAG TTA TTT TTC GAA CTG CGG GTG GCT CCA GCT CGA ATT AGT CTG CGC. The resulting gene had 1402 base pairs and was digested with restriction enzymes NcoI and XhoI (New England Biolabs). This gene was ligated into a NcoI and XhoI digested pET15b vector and transformed into DH5α *E. coli* cells for amplification. The resulting plasmid was purified using the Monarch Plasmid Miniprep Kit (New England Biolabs) and sequenced by Microsynth Seqlab. pSu9(1-67)-MBP-pET15b was then transformed into BL21 *E. coli* cells for expression.

Cells were grown in lysogeny broth (LB) media at 37°C and induced with 1 mM IPTG. After overnight incubation at 16°C, cells were pelleted by centrifugation at 10,000 rpm for 20 min. Pellets were resuspended in 100 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA and 1 mM PMSF, and sonicated for cell disruption. The suspension was centrifuged at 40,000 rpm for 30 min at 4°C. The supernatant was loaded into a

gravity Strep-Tactin column and washed with 100 mM Tris pH 8.0, 150 mM NaCl and 1 mM EDTA. The protein construct was eluted with the same buffer, containing 2.5 mM desthiobiotin. The resulting fractions were concentrated and loaded into a Superdex 200 Increase size-exclusion column (Cytiva). Precursor proteins were eluted with 10 mM Tris pH 7.0, 50 mM ammonium acetate, 1 mM EDTA, 1 mM TCEP, 0.02% GDN.

3.6.2 Synthesis of pALDH

The pALDH(1-19) presequence was taken from rat aldehyde dehydrogenase (ALDH) and synthetized by GenScript. A glycine-glycine-serine (GGS) linker and a StrepII-tag were added at its C-terminal. Upon delivery, lyophilized proteins were resuspended in water, aliquoted and frozen until required.

3.7 Laser-induced liquid bead ion desorption

The TOM holo sample was buffer-exchanged using a Superdex 200 Increase column (GE). The complex was eluted with 10 mM Tris pH 7.0, 15 mM ammonium acetate and 0.02% GDN. After SDS-PAGE assessment, peak fractions containing TOM were pooled and concentrated to 4 mg/ml using an AmiconUltra 100 kDa cutoff tube.

For LILBID-MS analysis, an infrared (IR) laser was used to generate ions from 50 µm microdroplets containing our proteins (*85*). Microdroplets were produced by a droplet generator (MD-K-130; Microdrop Technologies) and released ions were detected with a home-built time-of-flight analyzer, operating at a vacuum of 10^{-6} mbar. The IR laser operated at the absorption wavelength of water (2.94 µm) and droplets were produced and irradiated at a frequency of 10 Hz. Laser power was varied between a range of 10 mJ and 23 mJ. Each measurement was performed in negative ion mode with a sample volume of 4 µl. Spectra show the averaged signals of 1000 droplets that were normalized and analyzed using Massign (*86*). Peak assignment was based on predicted average molecular mass of individual subunits, calculated using ExPASy (*87*).

3.8 Crosslinking mass spectromety

TOM holo complex was further purified by size-exclusion chromatography using a Superdex 200 Increase column (GE). The complex was eluted with 50 mM KPO_4 pH 8.0 and 0.02% GDN. The best fraction was incubated with pSu9-MBP at a 1:8 ratio of TOM:pSu9-MBP, respectively, for 1 h at room temperature. Crosslinker DSSO was dissolved in DMSO and incubated with both samples at 8 mM concentration at room temperature. After 1 h, 1 M Tris pH 7.5 was added to the mixtures to quench the crosslinking reaction. A second TOM sample, in MES buffer at pH 5.5, was incubated in 12 mM EDC and 30 mM NHS. After 1 h, pH was increased to 7 using concentrated KPO_4. Then the activated sample was incubated with pSu9-MBP at a 1:8 ratio for 2 h at room temperature. The samples were assessed by SDS-polyacrylamide gel electrophoresis (PAGE) using Coomassie Brilliant Blue staining. Gel bands corresponding to crosslinked products were cut out and delivered to Dr. Langer's laboratory. Crosslinking results were visualized using xiNET (*88*).

3.9 Negative-stain

3 μl of TOM sample were incubated on a glow-discharged, carbon-coated, copper grid (3 mm, 400 mesh) for 2 mins. The grid was washed twice with ddH_2O and stained twice with uranyl formate. After every wash and stain step, the grid was blotted using Whatman paper no.50 (Figure 3.4). The grid was loaded into a Tecnai Spirit Biotwin electron microscope (FEI) at 120 kV. Micrographs were collected with a magnification of 49,000x, corresponding to a pixel size of 2.2 Å. All negative stain data was processed using open source software Relion-4.0 (*89*).

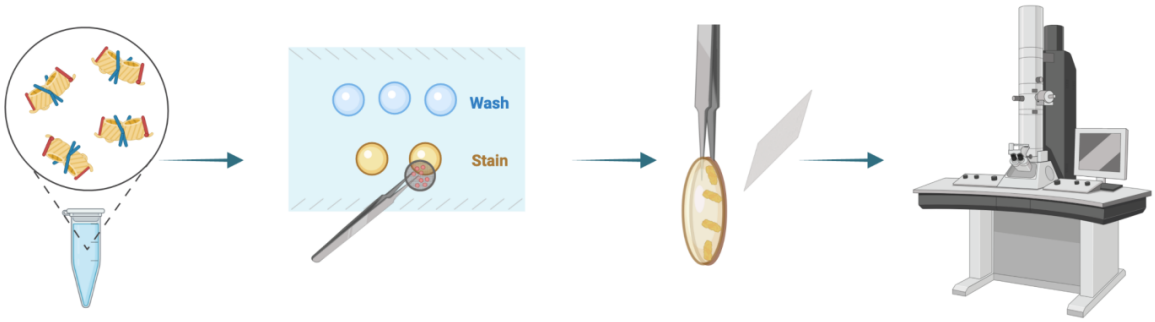

Figure 3.4: **Sample preparation for negative-stain electron microscopy.** A droplet of sample is placed on a carbon-coated copper grid. The grid is washed, stained, blotted and air-dried prior to loading into an electron microscope. Created with BioRender.

3.10 CryoEM

3.10.1 TOM holo + pSu9-MBP

For single-particle cryoEM, purified TOM holo complex in GDN was incubated with pSu9-MBP at a 1:8 ratio for 1 h at room temperature. The mix was loaded onto a Superdex 200 Increase size-exclusion column (GE) equilibrated with 10 mM Tris pH 7.0, 50 mM amonium acetate, 1 mM EDTA, 1 mM TCEP and 0.02% GDN (Figure 3.5). The peak fraction, eluting at 1.1 ml, with a concentration of 1.8 mg/ml, was applied to a glow-discharged C-Flat 1.2/1.3 Cu grid. Using a Vitrobot Mark IV (FEI), the grid was blotted for 3 s at 100% humidity at 4°C, and flash-frozen in liquid ethane.

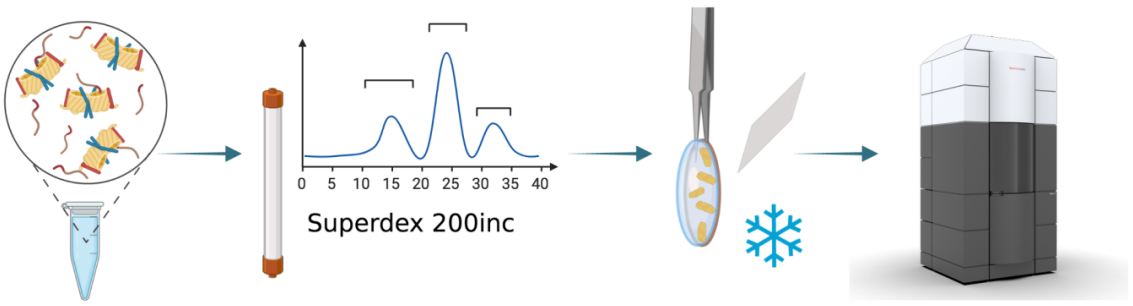

Figure 3.5: **Sample preparation for cryoEM.** After incubation with a precursor protein, TOM holo complex was further purified and flash frozen on a copper grid. Created with BioRender.

9,442 movies were collected using a Titan Krios G2 electron microscope (ThermoFisher Scientific) at 300 keV, equipped with a Gatan K3 camera in counting mode and a Gatan BioQuantum energy filter. Dose-fractionated movies were acquired at a 105,000x magnification, corresponding to 0.83 Å pixel size. Total accumulated dose was 55 e/A^2 during a 3 s exposure picture. Images were collected with defocus between -1.8 and -3.2 μm.

Single-particle analysis

Collected images were processed live using cryoSPARC v3 (*90*). Micrograph stacks were corrected with the local-motion correction software implemented in cryoSPARC and the CTF parameters were initially estimated using CTFFIND-4 (*72*). Automated particle-picking was applied in cryoSPARC, first using the blob picker and then using templates. Particle picks were 2x binned, extracted and classified in 2D to discard artefacts. 996,000 particles were selected and separated by 3D classification. The best 3D class, containing 254,000 particles was further processed and separated into clusters based on conformational heterogeneity using 3D variability analysis (*91*). Two resulting maps were further refined into a 5.2 Å resolution map of TOM core complex and a 7.4 Å map of the precursor-bound TOM holo complex. Out of the initial 996,000 particles, 350,000 were separated by 3D classification, leading to two refined 10 Å maps of MBP.

3.10.2 TOM holo + pALDH

TOM holo complex was incubated with pALDH at a 1:8 ratio for 1 h at room temperature. The mixture was loaded onto a Superdex 200 Increase size-exclusion column (GE), equilibrated with 50 mM KPO4 pH 8.0, 50 mM KCl, 1 mM EDTA, 1 mM TCEP and 0.02% GDN (Figure 3.5). 3 μl of the peak fraction were applied to a glow-discharged C-Flat 1.2/1.3 Cu grid. Using a Vitrobot Mark IV (FEI), the grid was blotted for 3 s at 100% humidity at 4°C, and flash-frozen in liquid ethane.

Micrographs were collected using a Titan Krios G2 electron microscope (ThermoFisher Scientific) at 300 keV, equipped with a Gatan K3 camera in counting mode and a Gatan BioQuantum energy filter. Dose-fractionated movies were acquired at 105,000x magnification and 0.83 Å pixel size. Total accumulated dose was 55 e/A^2 during a 3 s exposure picture. Images were collected with defocus range between -1.2 and -3.0 μ m.

Single-particle analysis

22,000 collected images were processed using Relion 4.0 (*89*). Micrograph stacks were motion-corrected with MotionCor2 (*71*) and CTF parameters were initially estimated using CTFFIND-4 (*72*), both as implemented in Relion. Particle-picking was performed in crYOLO: a small data set was used to manually train a model, which was subsequently applied to all micrographs (*92*). Particle-picks were imported back into Relion, extracted with 4x binning to 3.3 Å pixel size, and classified in 2D to discard artefacts. 1.5 million particles were selected, re-extracted with 1.5x binning (1.3 Å pixel size) and separated by 3D classification. The best 3D class contained 620,000 particles that were re-extracted without binning. The corresponding map was 3D refined and postprocessed in Relion to 4.6 Å resolution using a global mask. The resulting map was used for polishing, reaching a 4.2 Å resolution, as assessed by the gold-standard FSC = 0.143 criterion.

TOM core complex map

Polished particles and maps were imported into cryoSPARC v3 (*90*). Non-uniform refinement resulted in a 3.6 Å resolution structure of the TOM core complex. After a round of 3D classification without alignment, 304,000 particles were subjected to another non-uniform refinement with imposed C2 symmetry, which produced a 3.37 Å resolution map. Local refinement in cryoSPARC, with a mask around the entire core complex but excluding the micelle, further improved resolution to 3.32 Å, as calculated by the gold-standard FSC = 0.143 criterion.

TOM core complex + pALHD complex map

Those 304,000 particles and map were non-uniform refined using 10 Å as maximum alignment resolution limit. This resulted in a 4 Å resolution map, according to the gold-standard FSC, with imposed C2 symmetry, and was subsequently used to study the preprotein bound TOM core complex.

TOM holo complex map

The initial 620,000 polished particles were further processed in Relion to reaveal the presence of Tom20. Briefly, using UCSF ChimeraX (*93*), a mask was designed to cover the cytoplasmic side of the complex. Particles were classified into five groups in 3D without alignment using the mask, resulting in two classes with distinct Tom20 conformations. These classes, with 140,000 and 120,000 particles each, were further 3D refined and postprocessed to a resolution of 6.6 Å and 6.7 Å respectively according to the gold-standard FSC.

3.11 Model building

3.11.1 Sequence analysis and structure prediction

Mitochondrial targeting sequences of precursor proteins of interest were compared with TargetP-2.0 (*94*). Secondary structure analysis of precursor proteins and of TOM subunits was performed with the Bioinformatics Toolkit of the Max Planck Institute of Biology (*95*), which offers and overview of secondary structure features like coiled coils, transmembrane helices and disordered regions. Quick2D (*96*) is a software within the Bioinformatics Toolkit that compares the outputs of multiple prediction algorithms in one, including PSIPRED3 (*97*), SPIDER2 (*98*), PSSpred (*99*), DeepCNF-SS (*100*), PCOILS (*101*), MARCOIL (*102*), TMHMM (*103*), Phobius (*104*), PolyPhobius (*105*), DISOPRED3 (*106*), SPOT-Disorder (*107*) and IUPred (*108*).

Tertiary structures of the precursor proteins, Tom70 and Tom20 were obtained from AlphaFold (*75*) using the AlphaFold platform set up in-house by Dr. Juan Castillo at the Max Planck Institute of Biohysics. Similarly, quaternary structures of TOM subcomplexes in association with precursors proteins were generated with AlphaFold-Multimer (*109*).

3.11.2 Modelling

A statistical prediction model of TOM core complex dimer was generated with AlphaFold-Multimer, and fitted into the refined core map using the recently developed model building software ISOLDE (*110*), within UCSF ChimeraX (*93*), and Coot (*76*). Additional real-space refinement of the model was performed in Phenix (*77*). Phosphatidylcholine and diosgenin were fitted into the structure using Coot. Preprotein-bound TOM core complex was rigid-body-fitted in UCSF ChimeraX. TOM core with Tom20 models were based on the AlphaFold prediction of Tom20 and Tom22. Tom20 and Tom22 were rigid-body-fitted into each conformation map, relaxed into the structure using Coot, and then merged with the TOM core model.

3.12 Data management

Images and diagrams in this thesis were created with InkScape, BioRender and Microsoft Power Point. Plots were made using Python3. Protein molecular weights were measured using ExPASy compute pI/MW tool (*87*). Gels and blots were scanned using a ChemiDoc Imaging System (Bio-Rad). Electrostatic potential of Tom40 was calculated using APBS-PDB2PQR software suite at pH 7 (*111*). 3D visualisation of maps and models was carried out in UCSF ChimeraX.

The cryoEM map and atomic model of TOM core complex have been deposited in the Protein Data Bank and in the Electron Microscopy Data Bank with accession codes PDB 8B4I and EMD-15849. The cryoEM maps of TOM complex with Tom20 in two different conformations have been deposited in the Electron Microscopy Data Bank with accession codes EMD-15850 and EMD-15856.

Chapter 4

Results

4.1 Protein preparation

4.1.1 TOM core complex

Neurospora crassa hyphae were grown at the laboratory of Prof. Nussberger at the University of Stuttgart. TOM core complex was purified from isolated mitohcondria via hexahistadine tag at the C-terminus of Tom22. The first purification step consisted of binding TOM core complex in DDM to a NiNTA column and eluting with high imidazole concentration.

A second purification step consisted on running the sample through an anion exchange column and eluting the complex using a salt gradient. TOM eluted at 300 mM KCl. We analysed the elution fractions by SDS-PAGE (Figure 4.1). Our purified TOM core complex contained all subunits belonging to the basic import pore, Tom40, Tom22, and a single band at ∼10 kDa corresponding to small TOM subunits, in agreement with published literature (*23*).

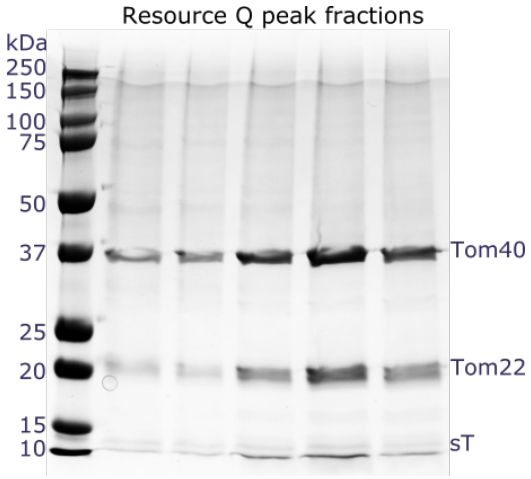

Figure 4.1: **Purification of TOM core complex.** Coomassie-stained SDS-PAGE gel after anion exchange chromatography of TOM core sample in DDM. Sample provided by Prof. Nussberger.

4.1.2 TOM holo complex

One hundred liters of *Neurospora crassa* hyphae resulted in small mitochondrial pellets, which were then swollen in order to separate their membranes. The outer membrane was separated through gradient centrifugation, producing a bright orange OMV fraction with 1 mg/ml concentration (Figure 4.2).

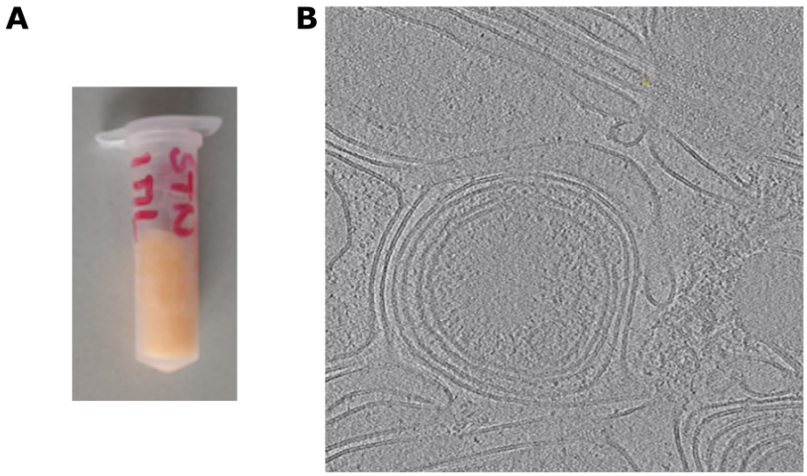

Figure 4.2: **Outer membrane vesicles.** A) 1 ml of outer-membrane vesicles prepared by Prof. Nussberger. B) Tomographic slice of the OMVs sample at 42,000x magnification obtained by Lea Dietrich.

We solubilized OMVs in glyco-diosgenin (GDN) and isolated TOM holo complex using a NiNTA column. Afterwards, we purified the complex by size exclusion chromatography, and analysed its subunit composition by SDS-PAGE (Figure 4.3). We identified the TOM core subunits, Tom40, Tom22 and the sTs, but also Tom70 and Tom20. These results were consistent with the literature (*18*). However, we also detected a ∼30 kDa protein, which we identified as VDAC, an abundant porin in the mitochondrial outer membrane and a common contaminant in TOM preparations (*83*).

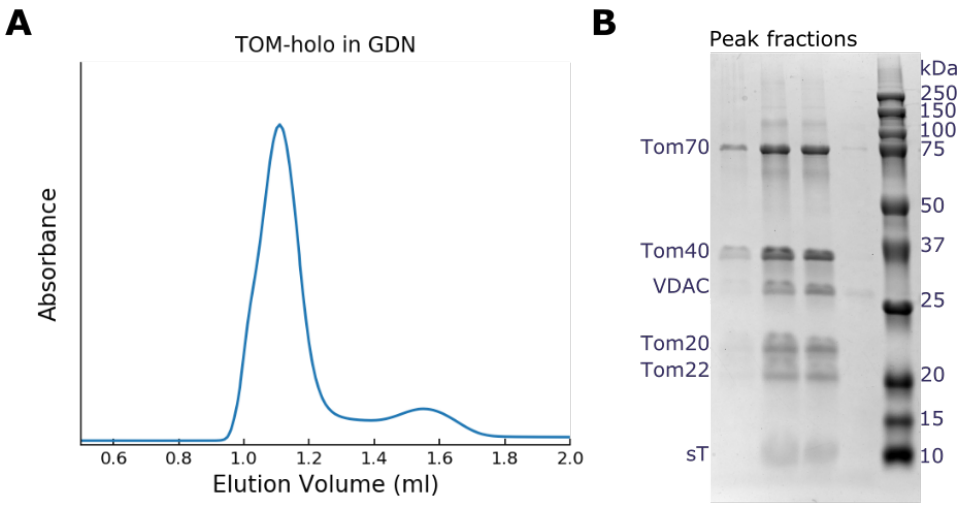

Figure 4.3: **Purification of TOM holo complex.** A) Size exclusion chromatography of TOM holo complex in GDN, using a Superdex 200 inc column (Cytiva). B) Coomassie stained SDS-PAGE of fractions corresponding to the size exclusion peak.

In collaboration with Dr. Thomas Bausewein, we initially attempted to purify the holo complex from whole mitochondria, but Tom70 and Tom20 seemed to dissociate from the general import port. It has been previously suggested that Tom70 is more abundant in sites of contact between the outer and inner membrane (*18*), so we might have lost it by harsh breaking of mitochondria. Upon separation of membranes and isolation of the complex from OMVs, we saw traces of Tom70 and Tom20 in our samples. Dr. Thomas Bausewein performed detergent screening to optimize solubilization. We saw that dodecylmaltoside (DDM) led to a total loss of the subunits of interest. Lauryl maltose neopentyl glycol (LMNG) and digitonin proved to conserve all subunits,

with partial loss of Tom20 and Tom70. However, GDN finally provided consistent results in which the seven subunits would elute in a single peak after size-exclusion chromatography.

4.1.3 Precursor proteins

We compiled a list of mitochondrial targeting precursor proteins and used TargetP-2.0 to predict their signal sequence and potential cleavage site (*94*). We then used the Bioinformatics Toolkit (Quick2D) to predict their secondary structure (*96*), comparing multiple algorithms that predict potential helices, strands, transmembrane domains or disordered regions. For details on these algorithms, see section 3.11.1. Lastly, we compared the presequences based on their amino acid composition and length, as shown in Table 4.1.

Name	Location	Organism	Sequence	Length
pHsp60	Matrix	Yeast	MLRSSVVRSRATLRPLLRRAY	21
pSu9	Inner Membrane	N. crassa	MASTRVLASRLASQMAASAKVARPAVRVAQVSKRTIQ TGSPLQTLKRTQMTSIVNATTRQAFQKRAY	67
pAtp1	Inner Membrane	N. crassa	MFRNALRQSTRAVGAFSATGRVAARNAAPVVSAVQAR TY	39
pDld2	Matrix	Yeast	MLRNILVRSTGSNFKFAGRYMKSSALLGYYRRVNYY	36
pCit1	Matrix	N. crassa	MAPVMRLGSAALRSSIHLTSRQTAFTAARCY	31
pIdh1	Matrix	N. crassa	MLSRNSMRLARAVAQPTPRLARGF	24
pCox4	Inner Membrane	Yeast	MLSLRQSIRFFKPATRTLCSSRYLL	25
pAldh	Matrix	Rat	MLRAALSTARRGPRLSRLLS	20
pCyB2	IMS	Yeast	MLKYKPLLKISKNCEAAILRASKTRLNTIRAY	32
pAco1	Matrix	Yeast	MLSARSAIKRPIVRGL	16

Table 4.1: **Presequence comparison.** Collection of different presequences considered for experiments in this thesis. The amino acid composition of the presequences was colored based on their charge (blue, positive) and hydrophobicity (green).

Six presequences were preliminarily selected based on their size (pAco1, pALDH, pHsp60, pSu9, pAtp1 and pCyB2), but only one short and one long presequence were chosen for further experimentation. The rest of this thesis is limited to the use of pSu9 and pALDH, as they showed consistent binding to TOM complex.

pSu9-MBP

pSu9 from *N. crassa* has been long used in mitochondrial translocation experiments (*112*). It targets the inner membrane, where it forms part of the C-ring of the F_0 domain of ATP synthase. We analysed the full sequence of this protein and were able to identify its cleavage site using TargetP-2.0 (Figure 4.4A). We used Quick2D to get a consensus of secondary structure prediction (Figure 4.4B). This way, we found that pSu9 spans residues 1-67 of the full protein sequence and that it is likely to either form helices or be disordered.

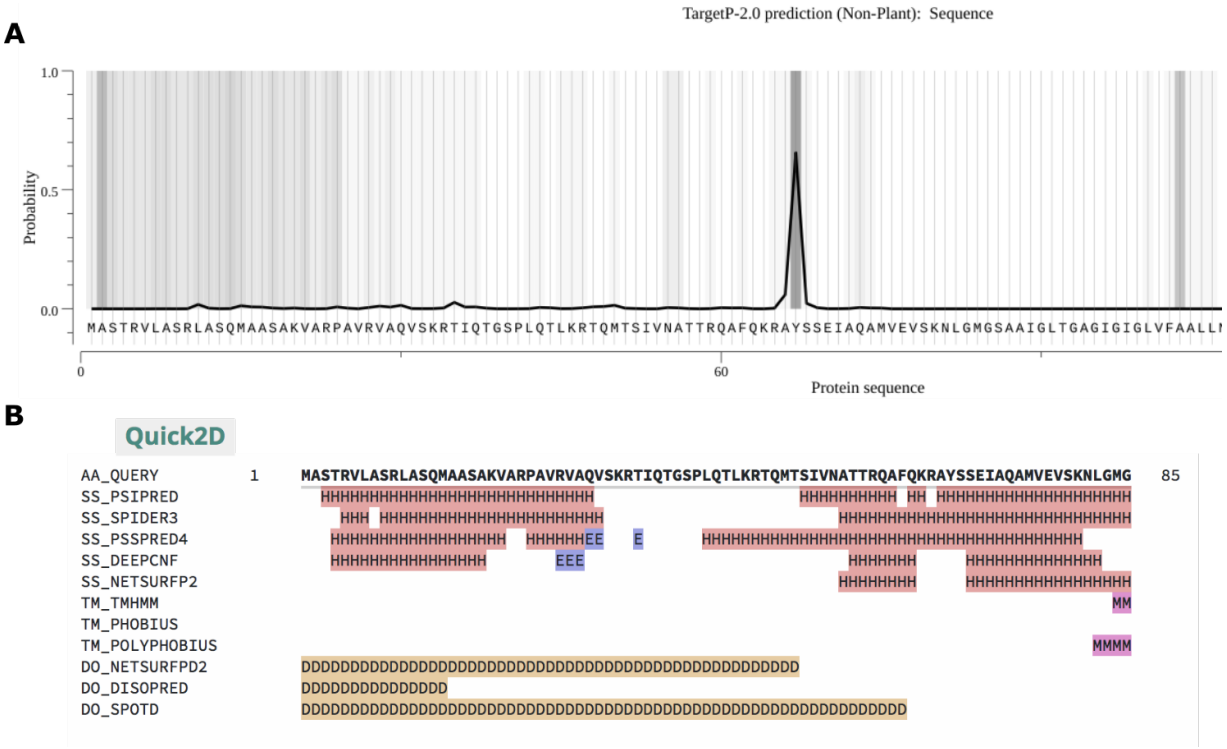

Figure 4.4: **Su9 presequence analysis.** Analysis of the truncated sequence of ATP synthase subunit 9. A) TargetP-2.0 prediction of its presequence cleavage site. B) Quick2D consensus analysis of secondary structure prediction. The algorithms used for prediction are listed on the left. Red denotes an α-helical region, violet a β-strand region, gold a disordered region, and magenta a transmembrane domain.

This presequence gene was cloned into a vector, together with the sequence of maltose binding protein (MBP). Recombinant expression in *E. coli* resulted in pSu9-MBP: a precursor protein construct formed by the presequence and a large folded domain. The construct contains a StrepII-tag at its C-terminus for purification and has a total of 448 residues (Supplementary material 2). Its predicted weight is of 50 kDa and runs at ~45 kDa in SDS-PAGE.

pALDH

The presequence of aldehyde dehydrogenase (pALDH) of *Rattus norvegicus* has been used in the past to study Tom20 as a mitochondrial receptor (*48*). Aldehyde dehydrogenase is an enzyme located in the mitochondrial matrix involved in alcohol metabolism. Similar to pSu9, we identified the cleavage site of pALDH and analysed its predicted secondary structure (Figure 4.5) using TargetP-2.0 and Quick2D. We determined that pALDH spans the first 19 residues of aldehyde dehydrogenase and that it could either form a helix or be disordered.

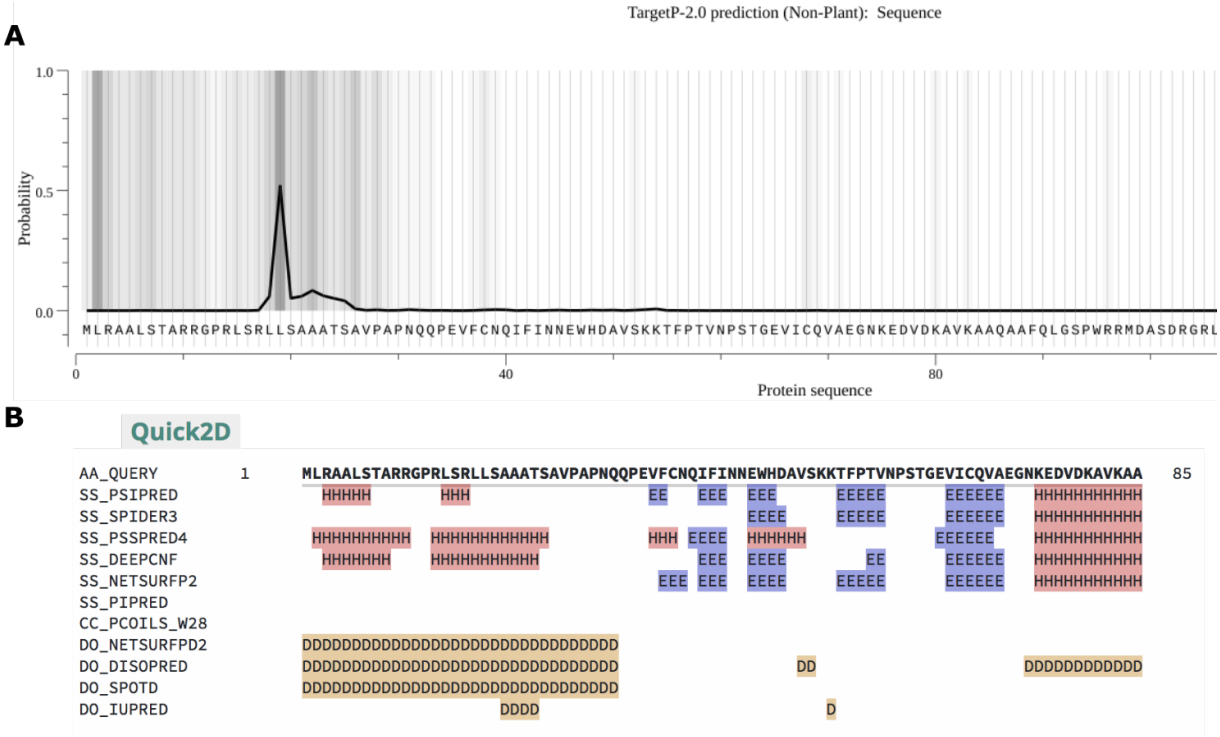

Figure 4.5: **ALDH presequence analysis.** Analysis of the truncated sequence of aldehyde dehydrogenase. A) TargetP-2.0 prediction of its presequence cleavage site. B) Quick2D consensus analysis of secondary structure prediction. The algorithms used for prediction are listed on the left. Red denotes an α-helical region, violet a β-strand region, gold a disordered region, and magenta a transmembrane domain.

We added a StrepII-tag to the C-terminus of pALDH. Because of its small size, we outsourced the synthesis and expression of the construct to GenScript. The full construct has 32 residues and has a predicted molecular weight of 3.5 kDa (Supplementary material 3). Upon delivery, lyophilized pALDH was resuspended in ddH_2O and aliquoted.

4.2 TOM core and holo complexes

4.2.1 Nano-differential scanning fluorimetry

We performed nanoDSF with both TOM core and holo samples in order to find their melting temperature as an indicator of protein stability. This indicates the transition point when half of the protein population in each sample are present in an unfolded state (*113*). Our measurements show thermal denaturation of proteins by measuring tryptophan fluorescence at 350 nm and 330 nm, at 70 temperature points, ranging from 20°C to 90°C (Figure 4.6). We compared the behaviour of the complexes under different buffer conditions, by increasing their pH by half a point between 5 and 9. We found that both complexes are stable up to high temperatures, with little difference in varying pH conditions. We determined an average melting temperature of 55°C for TOM core, and a higher melting temperature of 69°C for TOM holo. This highlights that the additional subunits stabilize the TOM holo complex, in comparison to the core complex.

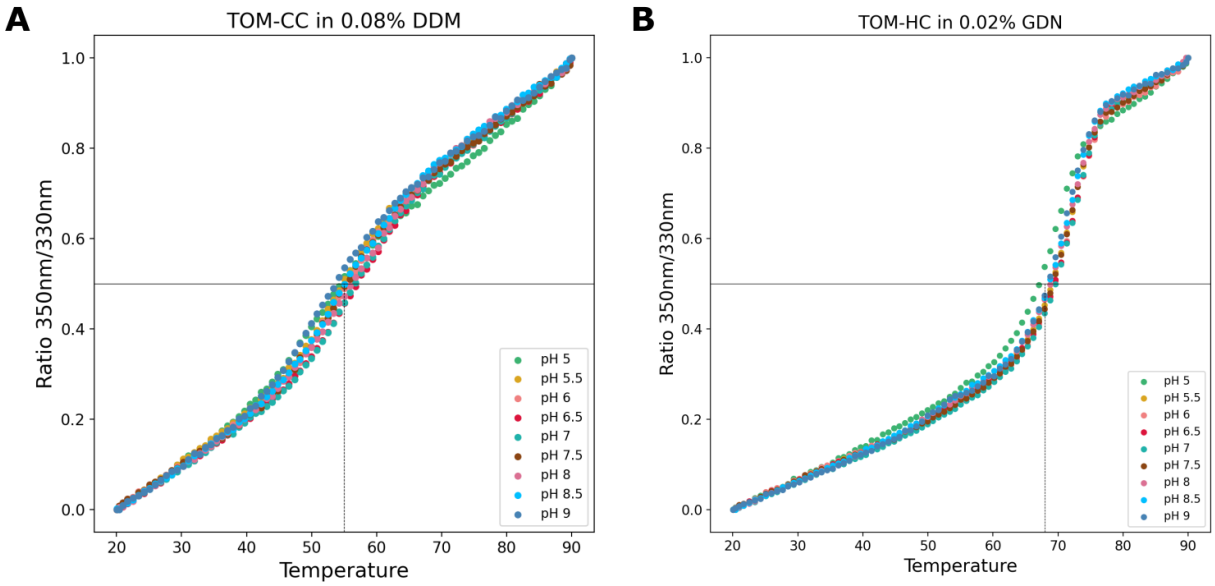

Figure 4.6: **Differential scanning fluorimetry of TOM complex.** We tested the stability of both TOM A) core and B) holo complexes by nanoDSF in pH from 5 to 9, at a temperature range from 20°C to 90°C.

4.2.2 Negative-stain electron microscopy

Prior to cryoEM, we analysed all TOM samples by negative-stain electron microscopy. This enabled us to screen the samples quality and concentration, as well as to attain an initial low-resolution protein structure.

TOM core complex

Figure 4.7 shows 2D class averages of the TOM core sample in DDM, obtained from imaging 47 negatively stained samples. In Figure 4.7, particle class averages show the protein complex in different orientations. Classes 1 and 2 are clear examples of TOM top views, corresponding to the view from either the cytoplasm or IMS side. This view is characterized by an eight-shape appearance, marked by the two translocation pores.

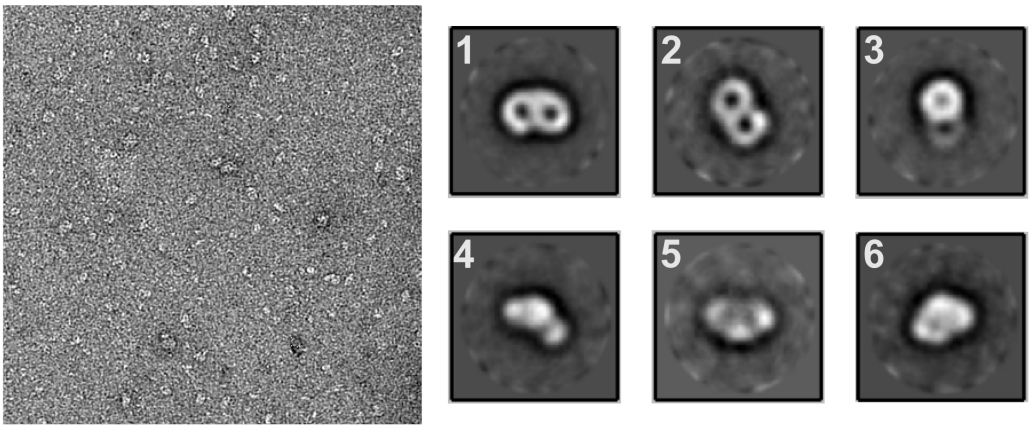

Figure 4.7: **Negative stained TOM core complex.** Left: Micrograph of negative stained TOM core complex in DDM at a 49,000x magnification. Right: Selected 2D averages ordered by decreasing particle number.

Class 4 shows TOM core complex from its side, as seen from the membrane. It has a characteristic elongated shape, seemingly comprised of three round sections. Classes 3, 5 and 6 show the complex from other angles. The size of the particles, as measured in Relion on their longest axis, ranges from 124 Å to 130 Å, in agreement with the expected TOM dimensions.

TOM holo complex

We classified negatively stained TOM holo particles from 40 micrographs, shown in Figure 4.8. It is easy to identify the elongated side view of the complex in class 1, as well as a two pore-containing top view in class 4. These, as well as class 3, have larger dimensions compared to the core complex class averages, increased from 137 Å to 150 Å. Class 5, however, is a rounder top view that seems to contain an additional feature on one side. This may be related to a TOM trimer, seen in the past by Model et al. (*40*), or caused by the presence of Tom20 and Tom70 subunits.

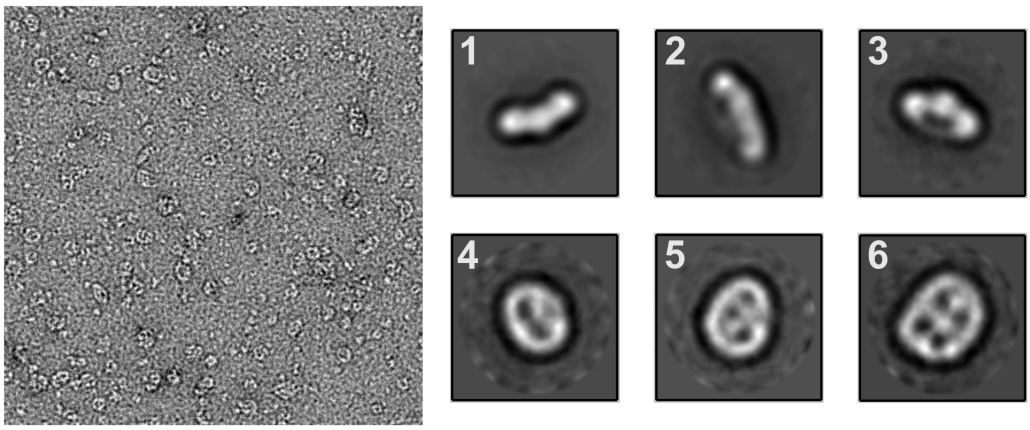

Figure 4.8: **Negative stained TOM holo complex.** Left: Micrograph of negative stained TOM holo complex in GDN at a 49,000x magnification. Right: Selected 2D averages ordered by decreasing particle number.

Additionally, classes 2 and 6, with dimensions of 178 Å by 135 Å, show a larger version of the complex. We consider that class 6 corresponds to the top view of tetrameric TOM complex reported by Tucker et al. (*25*), and class 2, to its side view (Figure 1.8).

4.2.3 Native mass spectrometry

We studied TOM complex through LILBID mass sprectrometry in collaboration with Prof. Mogner at the Goethe University Frankfurt. Using this technique, we compared subcomplex formation of the purified TOM core and holo complexes. Figure 4.9 shows dissociation of monomeric subunits at high laser intensity, up to 45 m/z. Peaks corresponding to TOM core subunits match those seen by Mager et al. (*24*). In

TOM holo spectrum we observe three additional peaks. The first one at 20,10 m/z, matching the 20.23 kDa predicted mass of Tom20, and a second one at 10,18 m/z, most likely corresponding to doubly-charged Tom20. Lastly, we see a small peak at 29,83 m/z that matches the predicted VDAC molecular weight of 30.50 kDa.

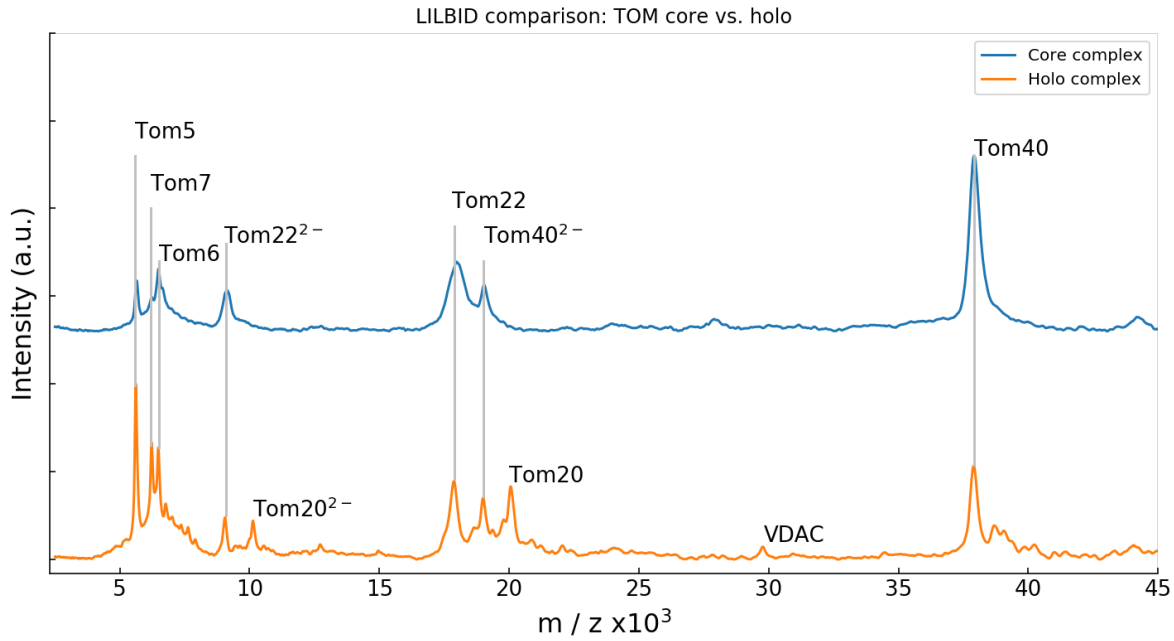

Figure 4.9: **LILBID of TOM core and holo complexes at high-laser intensity.** Spectra of both complexes up to 45,000 m/z at high laser intensity with assignment based on the predicted molecular weights of each subunit. Additional to TOM core subunits, peaks corresponding to Tom20 and VDAC are visible in the spectra.

Figure 4.10 shows spectra obtained from low laser intensity up to 125,000 m/z. We observed two monomeric components. Tom40, with a predicted weight of 38.15 kDa, appears at 37,93 m/z in both samples. However, only in the TOM holo spectrum we observe a shoulder at 69,36 m/z that corresponds to monomeric Tom70 with a 69.34 kDa predicted molecular weight. In addition, at this laser intensity we noticed larger assemblies formed by TOM subunits. TOM core spectrum shows interactions of Tom40 with itself, forming the general import pore dimer ($Tom40_2$), and its interaction with Tom22 and small TOM subunits ($Tom40_1sT_1$, $Tom22_1Tom40_1$, $Tom22_1Tom40_1sT_1$). In the TOM holo spectrum we found subcomplexes formed by both core and holo subunits. Interestingly, we can see multiple interactions between Tom70 and other subunits,

namely Tom70 with small TOM subunits ($Tom70_1sT_2$), Tom70 and the translocation pore ($Tom70_1Tom40_1$) and Tom70 with other receptor subunits ($Tom20_1Tom70_1$, $Tom22_1Tom70_1$ and $Tom22_2Tom70_1$).

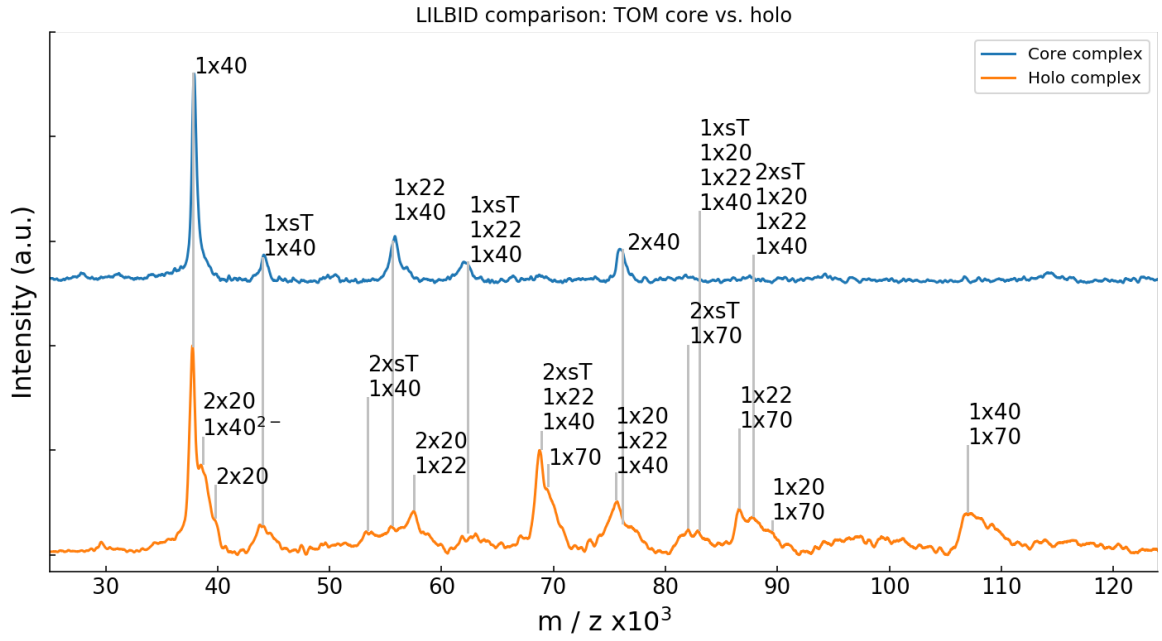

Figure 4.10: **LILBID of TOM core and holo complexes at low-laser intensity.** Spectra of both complexes up to 125,000 m/z at lower laser intensity. Tom40 and Tom70 appear as single molecules. The rest of the peaks reveal different subcomplexes formed by TOM subunits.

Our spectra also shows a shoulder at 39,79 m/z that roughly matches the predicted size of $Tom20_2$, 40.46 kDa, which would indicate possible dimerization of Tom20. Additionally, we can see the Tom20 dimer interacting with other subunits ($Tom20_2Tom40^{2-}{}_1$, $Tom20_2Tom22_1$). In the past, a crystal structure of pALDH bound Tom20 was reported to be a dimer. However, based on the residues involved in the interaction, dimer formation was deemed as biologically irrelevant. Lastly, we can see a subcomplex formed by Tom20, Tom22 and Tom40 ($Tom20_1Tom22_1Tom40_1$), which we also see in company of one and two small TOMs. This could be a hint of the conformation of a TOM holo protomer.

Table 4.2 summarizes the peaks observed in TOM holo's spectra, assigned based on predicted experimental masses (ExPASy) of all subunits and their possible subcomplex interactions. In the subcomplexes, sT represents one of the small TOM subunits with a value of 5.975 kD, calculated from the average mass of Tom5, Tom6 and Tom7.

Protein	Predicted Mass (kDa)	LILBID Mass (kDa)
Tom5	5.402	5.657
Tom7	6.061	6.231
Tom6	6.463	6.550
Tom22	16.816	17.901
Tom20	20.228	20.099
Tom40	38.150	37.925
Tom70	69.340	69.363
$Tom20_2$	40.456	39.785
sT_1Tom40_1	44.125	44.020
sT_2Tom40_1	50.101	53.483
$Tom22_1Tom40_1$	54.966	55.553
$Tom20_2Tom22_1$	57.272	57.623
$sT_1Tom22_1Tom40_1$	60.941	62.650
$sT_2Tom22_1Tom40_1$	66.917	68.859
$Tom20_1Tom22_1Tom40_1$	75.194	75.692
sT_2Tom70_1	81.291	81.980
$sT_1Tom20_1Tom22_1Tom40_1$	81.169	82.857
$Tom22_1Tom70_1$	86.156	86.453
$sT_2Tom20_1Tom22_1Tom40_1$	87.145	87.762
$Tom20_1Tom70_1$	89.568	89.496
$Tom40_1Tom70_1$	107.490	107.005

Table 4.2: **TOM holo complex, subcomplexes and subunits.** Predicted and experimental masses of subunits identified in our TOM holo sample, and their possible subcomplexes. Predicted masses were calculated using with ExPASy. sT represents one small TOM subunit: Tom5, Tom6 or Tom7.

4.2.4 Structure prediction

We investigated the structure of both complexes using structure prediction tools. These tools gave us an insight into the secondary and tertiary structure of the individual TOM subunits. Thanks to recent software developments, we were able to predict quaternary structure of the complexes, providing us with hints of interaction between subunits and their location. Nevertheless, it is important to remember that these

are only statistical predictions based on amino acid sequences of polypeptide chains, compared to other published protein structures. Interactions between the complex and its environment are not taken into consideration, disregarding the influence of lipid and water molecules in protein structure.

TOM core complex

We used AlphaFold Multimer to predict the structure of *N. crassa* TOM core complex. By default, we generated 25 prediction models of the dimer by providing two copies of each subunit. Unsurprisingly, the predicted structure was similar to the published structures of the *S. cerevisiae* and *H. sapiens* TOM core complex. These resulting models were analysed using UCSF ChimeraX and compared based on the estimated per-residue confidence provided by AlphaFold's algorithm. Figure 4.11 shows a representative model.

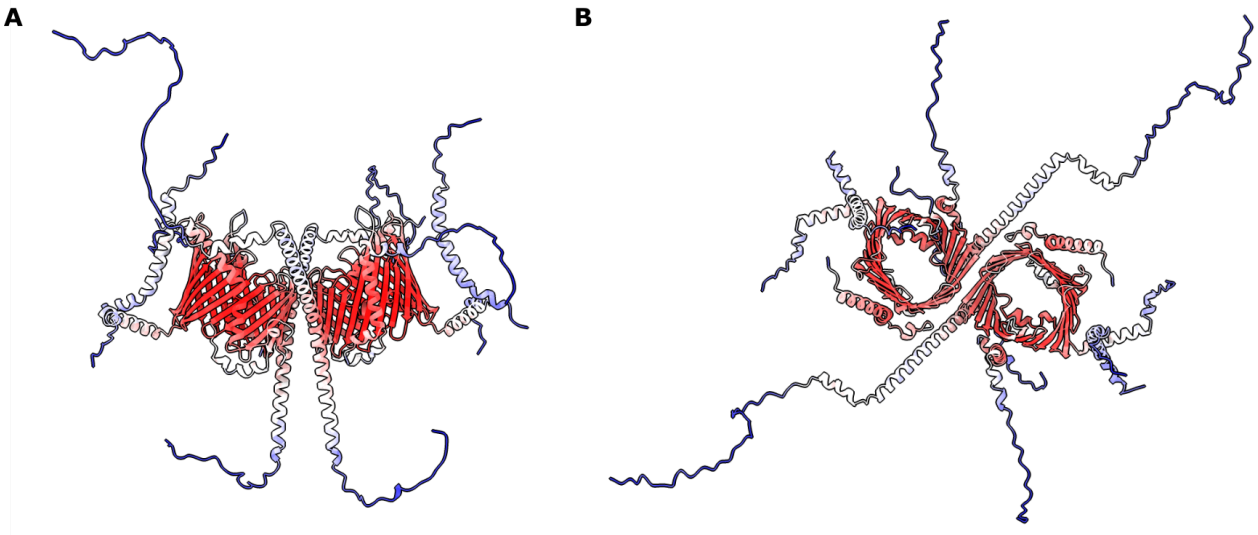

Figure 4.11: **Structure prediction of *N. crassa* TOM core.** AlphaFold prediction model of *N. crassa* TOM core dimer composed by two copies of Tom40, Tom22, Tom7, Tom6 and Tom5 as seen A) from its membrane side and B) from its cytoplasmic side. The model is colored according to reported per-residue confidence estimated rate. Red indicates high confidence, white, medium and blue, low.

High prediction confidence level was colored red. This was the case of both Tom40 β-barrels, their internal helices, and the transmembrane regions of Tom7 and Tom6. This was consistent in all 25 predicted models, suggesting their high stability in the

complex. Regions of mid-confidence were colored white, including parts of Tom22 and Tom5, as well as the IMS helices of Tom40. Low-confidence level was blue, which was common in disordered N and C-termini of all helical subunits. Low-confidence regions are likely flexible, which explains why those regions are not visible in the available cryoEM TOM core structures obtained by single-particle analysis (Figures 1.6 and 1.7).

We also analysed a possible structure of TOM core trimer, previously found through crossliking experiments and a low-resolution cryoEM map (*40*, *41*). Figure 4.12 shows a representative prediction model, generated with three copies of each TOM core subunit. The prediction models differed in the lower-confidence disordered domains, except for one model where the output was a core dimer next to an independent TOM protomer. The predicted trimer conformation matches the interactions suggested by a crosslinking model proposed in the literature (*41*, *42*). Compared to the TOM dimer model, in the timer prediction a single Tom22 links two Tom40 barrels, probably stabilized by the tail of Tom7. However, all small TOM subunits seem conserve to their positions around Tom40.

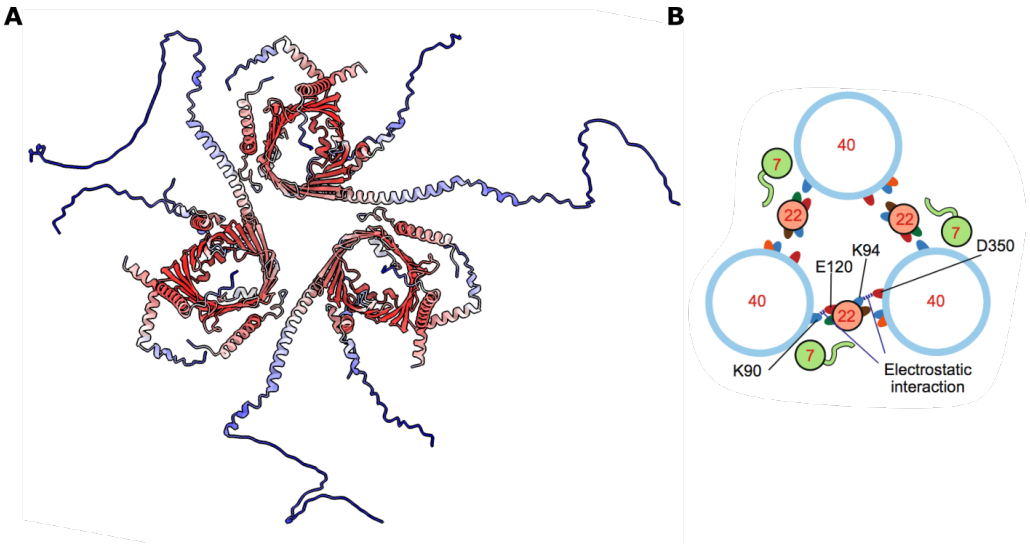

Figure 4.12: **Prediction model of *N. crassa* TOM core trimer.** A) AlphaFold prediction model of TOM core trimer composed by three copies of Tom40, Tom22, Tom7, Tom6 and Tom5, seen from its cytoplasmic side. The model is colored according to estimated per-residue confidence rate. Red indicates high confidence, white, medium and blue, low. B) Crosslink-based structure of TOM trimer according to Araiso et al., 2019.

TOM holo complex

The structures of Tom70 and Tom20 are considerably less understood than those of the TOM core subunits. As in the previous section, we used the Quick2D Bioinformatics toolkit of the Max Planck Institute of Biology (*96*) for secondary structure analysis of Tom70 and Tom20 though multiple algorithms. Subsequently, we built prediction models of these subunits and TOM holo complex using AlphaFold.

Tom70

We predicted formation of helices or sheets formed by the 624-residue sequence of Tom70. We identified 28 α-helices, which appear numbered in Figure 4.13 (red). Two small helical regions between α2 and α3 were not labeled because of their short length.

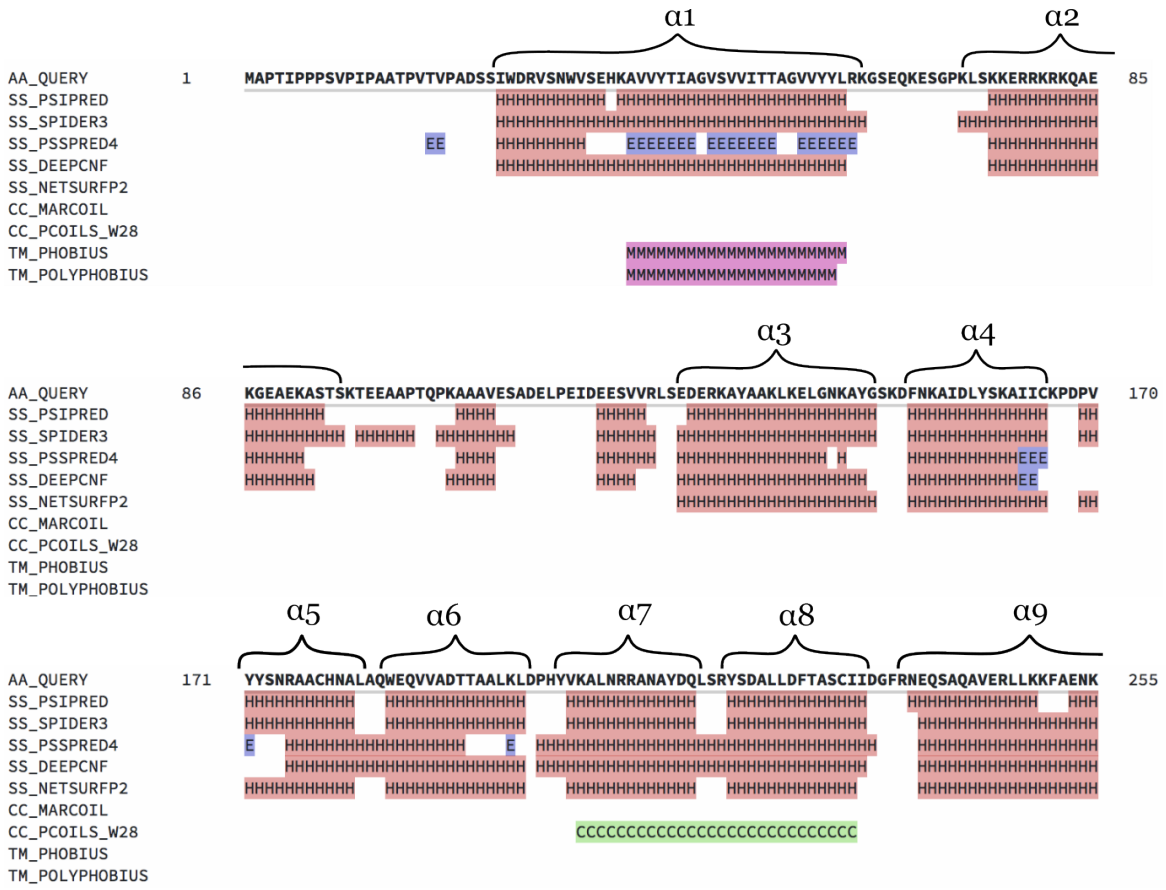

Figure 4.13: **Tom70 sequence analysis.** Analysis of the sequence of *N. crassa* Tom70 using Quick2D to predict its secondary structure and transmembrane domain. The algorithms used for prediction are listed on the left. Red denotes an α-helical region, violet a β-strand region, green a coiled coil, and magenta a transmembrane region. The sequence continues on the next page.

Figure 4.13: **Tom70 sequence analysis (continued).** Analysis of the sequence of *N. crassa* Tom70 using Quick2D to predict its secondary structure and transmembrane domain. The algorithms used for prediction are listed on the left. Red denotes an α-helical region, violet a β-strand region, green a coiled coil, and magenta a transmembrane region.

Additionally, we predicted a transmembrane domain within α1, at the N-terminal of Tom70 (magenta), and identified possible formation of coiled coils (green) in four regions, suggesting that helices in the soluble domain of Tom70 are packed close together.

Using AlphaFold we generated prediction models of Tom70, which differed from each other in their disordered regions, showed in blue in Figure 4.14A. This model predicted 26 α-helices, as expected, closely matching the 26 helices that form the published cryoEM human Tom70 structure (*27*), shown in Figure 4.14B. In the prediction model we can observe two characteristic sections of the soluble domain of human Tom70, a chaperone-docking area formed by α3 to α9 and a precursor-binding pocket spanning from α10 to α28 (*45*, *47*).

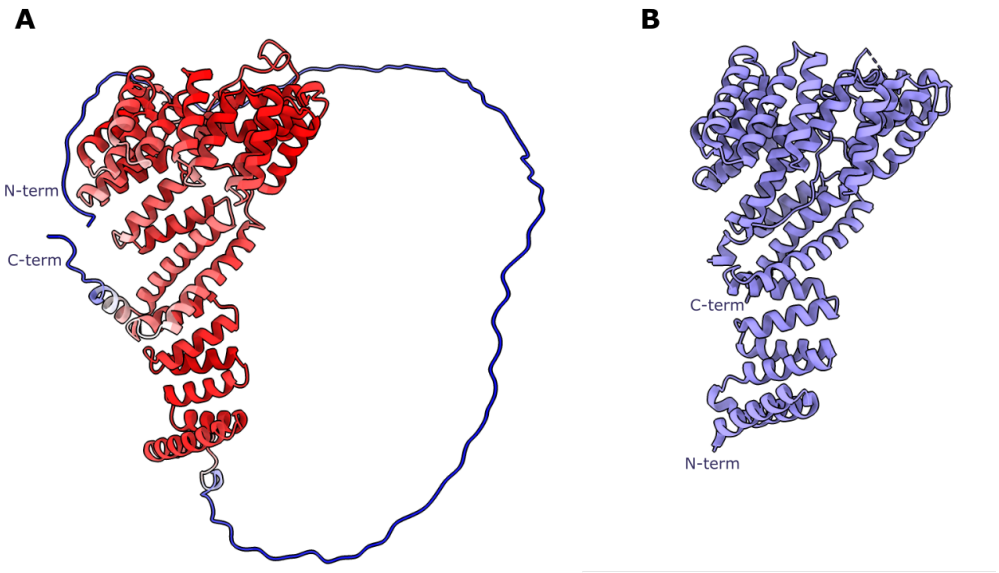

Figure 4.14: **Structure of Tom70.** A) AlphaFold prediction model of the full *N. crassa* Tom70 sequence, colored according to reported per-residue confidence estimated rate. Red indicates high confidence, white, medium and blue, low. B) CryoEM model of the soluble domain of human Tom70 (Gordon et al., 2020, PDB 7KDT).

The N-terminus of the Tom70 AlphaFold model appears completely disordered compared to the predicted Quick2D model. The first helix in the prediction corresponds to helix α3 in Figure 4.13. Even its predicted transmembrane domain appears completely disordered, and seems to bind the precursor binding pocket. As this interaction is unlikely, we assume that it was predicted based only on hydrophobic interactions.

Careful analysis of the Tom70 sequence helped us gain insight of its entire structure. Both prediction software indicated that its soluble, helical receptor domain (red in Figure 4.14A) is stable, but that, because of its long disordered region, the subunit as a whole is highly flexible.

Tom20

Quick2D predicted formation of seven α-helices along 181 residues in Tom20 (Figure 4.15). At its N-terminal, α1 was predicted to form a transmembrane domain (magenta). Next, α2 is a long helix that connects the known peptide binding domain with the transmembrane helix. Helices α3 to α6 match Tom20's crystal structure, shown in Figure 1.11B. At its C-terminus, Quick2D predicts a short helix, α7, surrounded by unstructured domains.

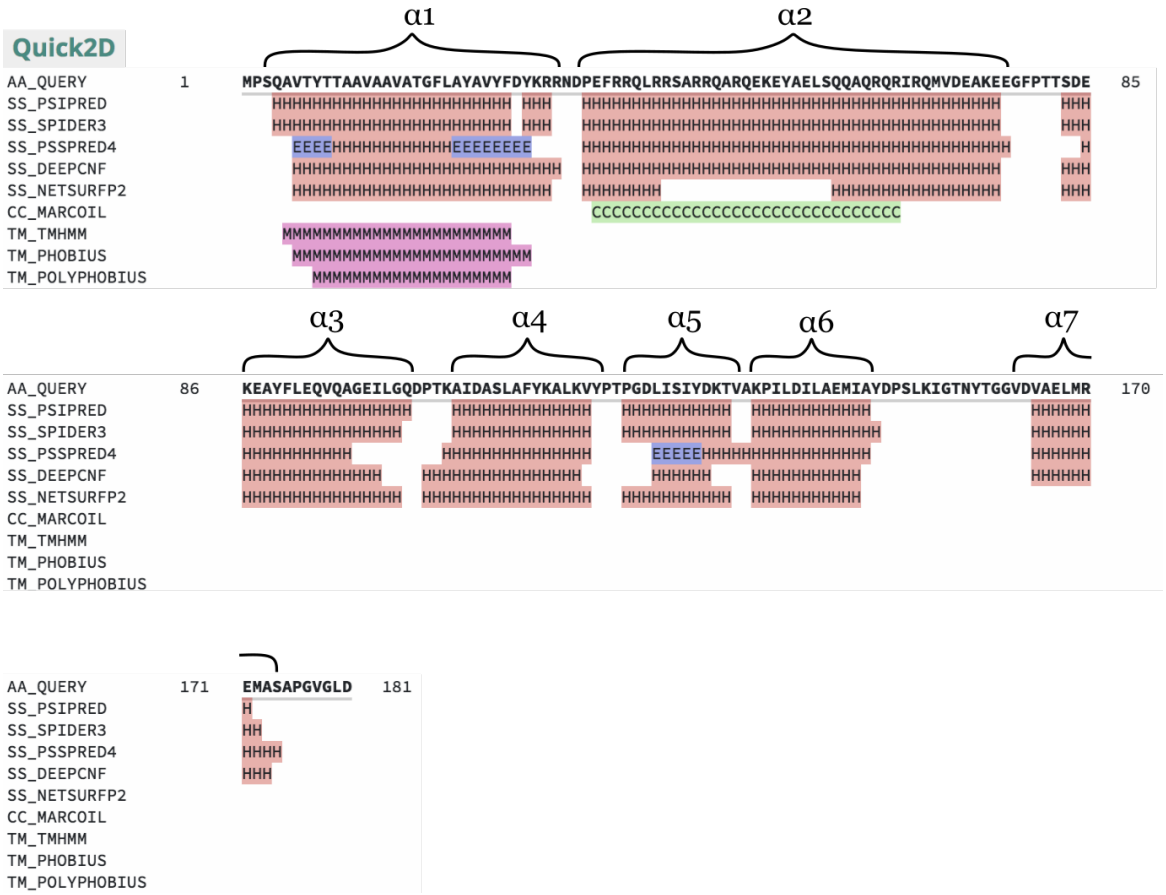

Figure 4.15: **Tom20 sequence analysis.** Analysis of the sequence of Tom20 using Quick2D to predict its secondary structure and transmembrane domain. The algorithms used for prediction are listed on the left. Red denotes α-helical regions, violet β-strand regions, green a coiled coil, and magenta a transmembrane region.

The AlphaFold prediction model of Tom20 suggest a stable, mostly helical structure (Figure 4.16A). Similar to the Quick2D prediction, α1 and α2 appear as two long helices with high confidence level. Helices α3 to α6 form the cytoplasmic peptide binding domain, matching the previously published crystal of the Tom20 TPR (Figure 4.16B) (*49*). Last, α7, which is missing from the crystal structure, appears in the prediction with low-confidence level, indicating possible flexibility.

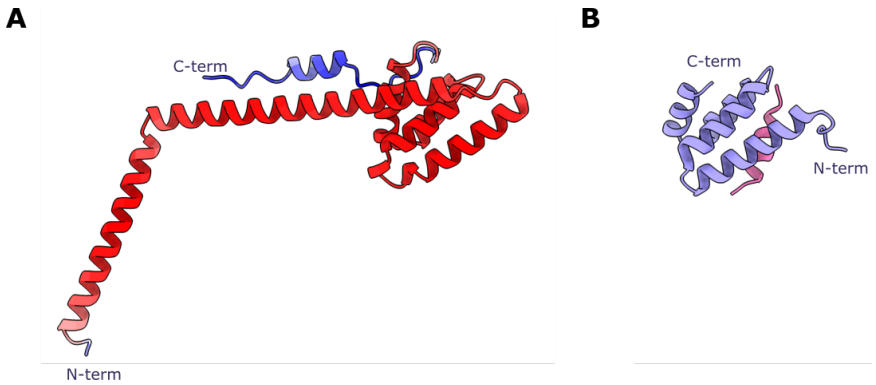

Figure 4.16: **Structure of Tom20.** A) AlphaFold prediction model of the full *N. crassa* Tom20 sequence, colored according to the reported per-residue confidence estimated rate. Red indicates high confidence, white, medium and blue, low. B) Crystal structure (Obita et al., 2007, PDB 2V1T) of the receptor domain of Tom20 (violet) bound to pALDH (pink).

Multimers

Based on subcomplexes observed through LILBID-MS, we predicted structures of different subunit combinations using AlphaFold Multimer. Initially, we focused on generating prediction models of the monomeric subcomplex composed of $Tom20_1Tom22_1Tom40_1$. Figure 4.17A shows four models that illustrate the remarcable similarity observed within a set of 25 predicted outcomes. In general, only disordered domains of Tom22, and Tom20 α7, appeared in different positions. Tom40 and Tom22 display the same interaction in do the core dimer. Tom20 interacts mainly with the helical N-terminal of Tom22. This region appears at mid and high-confidence level, meaning that it could be a stable interaction. Furthermore, the disordered N-terminal of Tom22 seems to wrap around Tom20 α2, in line with studies that suggest that Tom22 stabilises the receptor function of Tom20 (*114*). Lastly, the Tom20 peptide binding domain hovers above Tom40, on the cytoplasmic side, in close proximity to the longest loop of Tom40.

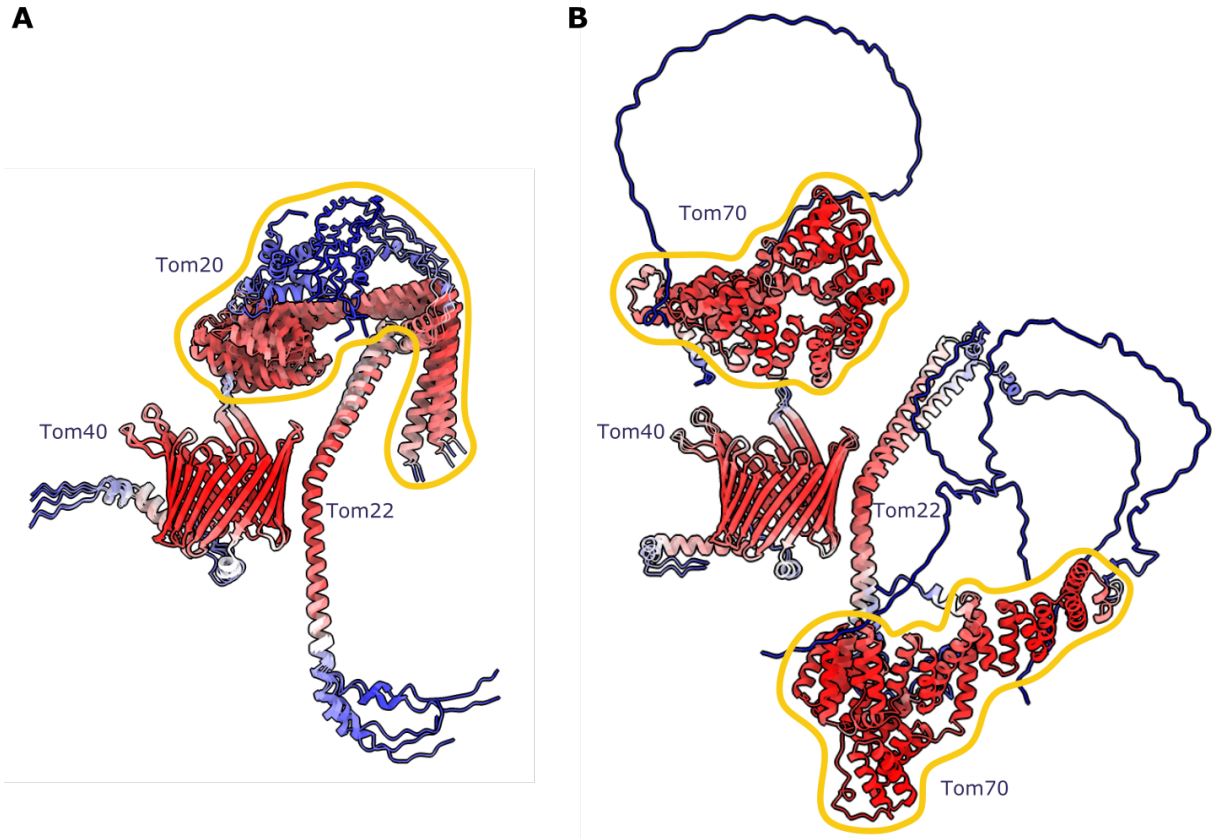

Figure 4.17: **Multimer predictions with TOM holo subunits.** AlphaFold Multimer prediction models of *N. crassa* Tom40 and Tom22 with A) one copy of Tom20 (highlighted in yellow) and B) one copy of Tom70 (highlighted in yellow). Multiple prediction models are shown, colored according to reported per-residue confidence estimated rate. Red indicates high-confidence, white, medium and blue, low.

We generated additional prediction models of the $Tom70_1Tom22_1Tom40_1$ subcomplex. For simplicity, Figure 4.17B shows only two of 25 prediction models obtained with AlphaFold Multimer. In this case, Tom40 and Tom22 conserved their positions throughout the predictions, differing only in the disordered regions of Tom22. However, Tom70 did not appear at a defined location in the subcomplex. In some predictions, the Tom70 receptor domain was on the cytoplasmic side of Tom40, at a similar position as Tom20 in Figure 4.17A. In other prediction models the Tom70 soluble domain sat along the membrane axis, resulting in biologically unlikely arrangements. Furthermore, the N-terminal of Tom70 was completely disordered, with no clear interaction site with Tom22 and Tom40. This prediction model reinforces the notion that Tom70 is

only lightly bound to the TOM core complex and that it probably has a very flexible interaction with Tom40.

We then predicted the entire TOM core dimer with one and two copies of Tom20. The models of the TOM core dimer with one Tom20 did not provide any new insight, as they resembled the interactions in $Tom20_1Tom22_1Tom40_1$ (Figure 4.17A). The presence of a second copy of Tom20 resulted in formation of a $Tom20_1Tom22_1Tom40_1$ dimer. As shown in Figure 4.18, each Tom20 binds the N-terminal of Tom22, with their receptor domains interacting with Tom40, hovering close to the translocation pores.

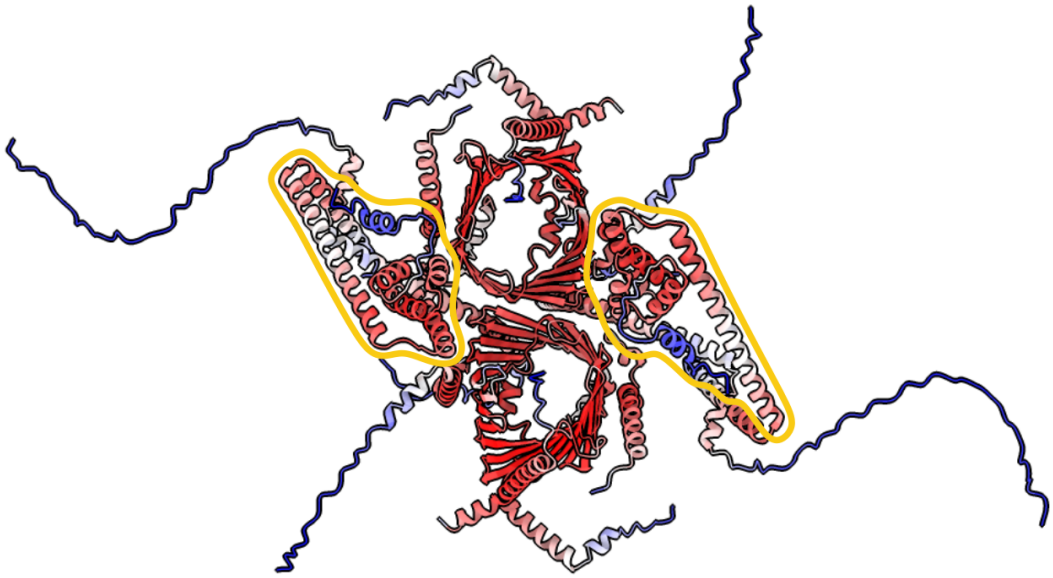

Figure 4.18: **Prediction of TOM core complex with Tom20 (top view).** AlphaFold Multimer prediction model of *N. crassa* TOM core complex with two copies of Tom20 (yellow outline). Colored according to reported per-residue confidence estimated rate. Red indicates high-confidence, white, medium and blue, low.

In the resulting models, however, the receptor domain of Tom20 seems to hover at a greater angle than in the $Tom20_1Tom22_1Tom40_1$ protomer (Figure 4.19). It is difficult to interpret such prediction data in terms of stoichiometry, but it is possible that two copies of Tom20 repel each other while waiting for an incoming presequence,

or that a TOM dimer containing two copies of Tom20 is not biologically viable due to potential steric clashes.

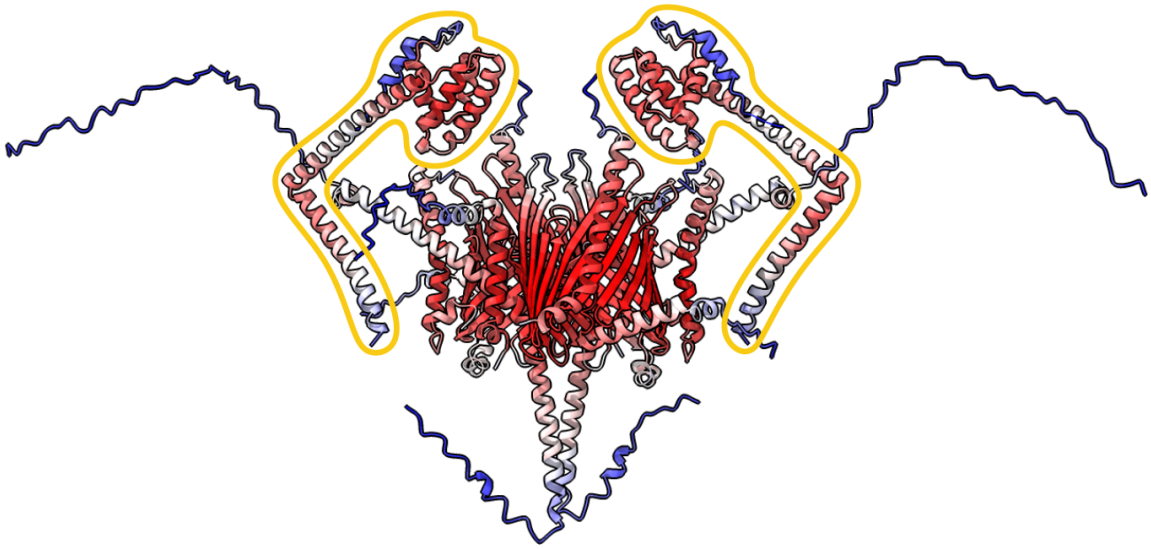

Figure 4.19: **Prediction of TOM core complex with Tom20 (side view).** AlphaFold Multimer prediction model of *N. crassa* TOM core complex with two copies of Tom20 (yellow outline) as seen from the membrane side. Colored according to reported per-residue confidence estimated rate. Red indicates high-confidence, white, medium and blue, low.

Finally, we ran a similar test with one and two copies of Tom70. However, these results matched those determined with the $Tom70_1Tom22_1Tom40_1$ monomer. Both single and double copies of Tom70 appeared at all directions around TOM, without a strongly defined position. In some cases, one copy of Tom70 appeared to hover on the cytoplasmic side of Tom40, as in Figure 4.17B. However, when predicting a TOM core + Tom20 + Tom70 model with AlphaFold, all cytoplasmic Tom70s in close proximity to the Tom40 barrel were displaced by Tom20.

4.3 Interaction of TOM holo complex with precursor proteins

In order to study the binding mechanism of precursor proteins to TOM, we incubated our TOM holo sample with two precursor proteins: pSu9-MBP and pALDH.

4.3.1 TOM holo and pSu9-MBP

We incubated TOM holo complex in GDN with our pSu9-MBP precursor, and analysed their interaction through clear Native PAGE, without addition of charged dyes or detergents, in duplicate. After transferring the gel to a PVDF membrane and staining it with Ponceau S, we found that TOM holo complex runs above 720 kDa, but that the preprotein-bound complex runs slightly lower (Figure 4.20). We then cut the membrane sections, and incubated one half with α-Tom22 and the other with α-StrepII antibodies to identify their bands. α-Tom22 was kindly provided by Prof. Nussberger.

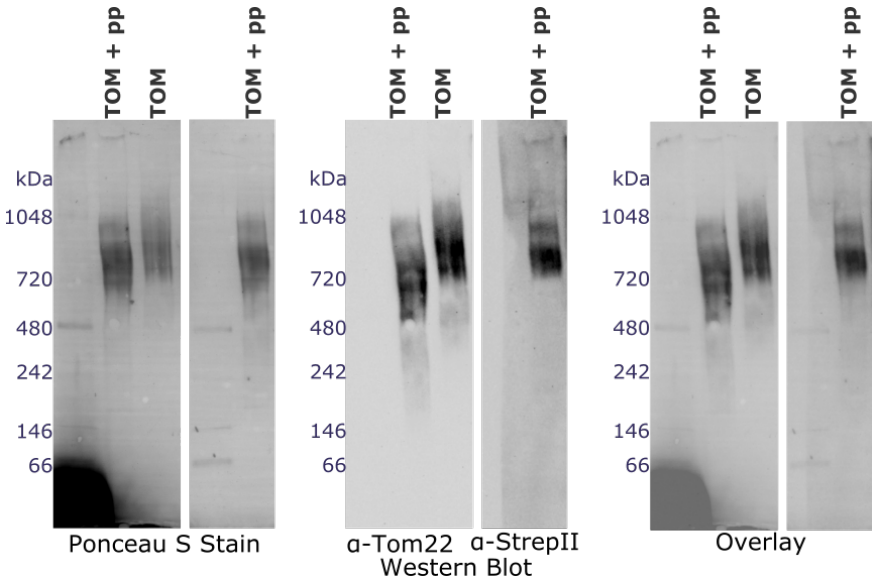

Figure 4.20: **Native PAGE of TOM holo with pSu9-MBP.** TOM holo with and without precursor protein pSu9-MBP (pp) were resolved through Native PAGE and Ponceau S staining. TOM bands were identified by Western blot after incubation with α-Tom22 and α-StrepII antibodies.

Western blot against α-Tom22 confirmed that the bands visualized with Ponceau S stain correspond to TOM. The results include multiple bands corresponding to the complex, perhaps due to different stoichiometric conformations of the complex within our sample, or due to different oligomerization states. Although size exclusion chromatography of the complex showed only one peak containing the TOM complex (Figure 4.3),

charge applied in Native-PAGE may have separated different oligomers. Furthermore, blotting against α-StrepII showed that the presequence was fully bound to TOM.

We next resolved the precursor-bound complex through a size exclusion column. We obtained two large elution peaks and analyzed their contents through SDS-PAGE (Figure 4.21). The fist peak corresponded to pSu9-MBP bound TOM holo complex, and the second, to unbound excess pSu9-MBP.

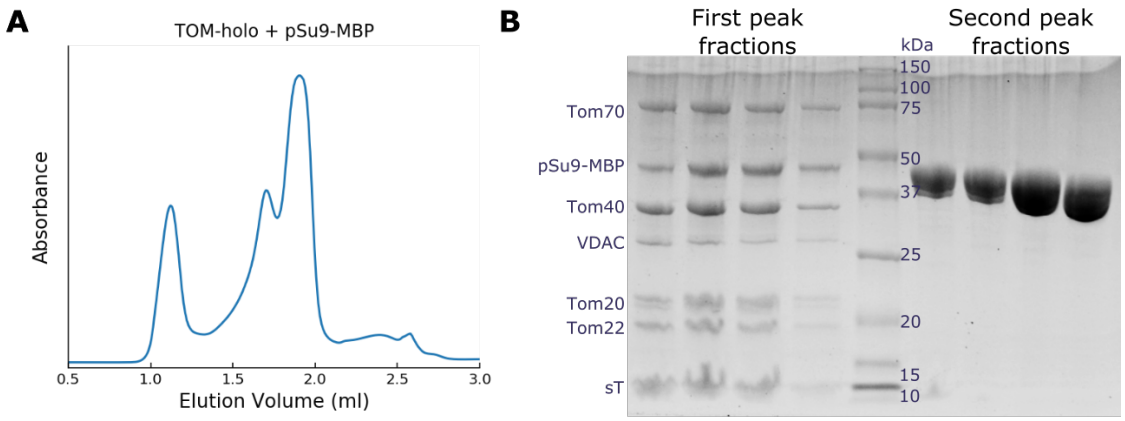

Figure 4.21: **Gel filtration of pSu9-MBP bound TOM holo.** A) Size exclusion chromatography of pSu9-MBP bound TOM holo complex in a Superdex 200inc column (Cytiva). B) Coomassie-stained SDS-PAGE of resulting elution fractions.

This confirmed that the precursor protein had strongly attached to the complex. Following gel filtration, we stained the sample for electron microscopy, collected 100 micrographs, and analysed class averages. Compared to negatively stained TOM holo complex, precursor-bound classes do not show a clear top view with visible pores, but a round particle with a fuzzy center (class 1). In agreement with our gel-filtration results, this fuzziness is a clear indication of a precursor binding and filling the pores. Classes 2 and 6 show side view-like particles, with a cloudy element on one side, which may reflect flexible binding of the precursor protein or the presence of flexible subunits. The opposite seems to be the case for classes 5 and 7, where we can see a small protein in focus with a larger blurred domain. The small feature measures around 60 Å in its longest axis, and is likely a MBP dimer.

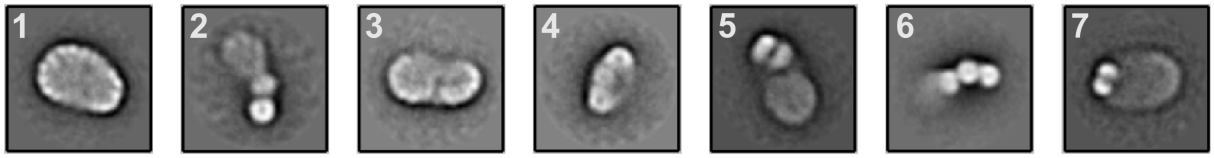

Figure 4.22: **Negatively stained TOM holo complex + pSu9-MBP.** Selected averages from 2D classification of the TOM core complex in DDM after incubation with pSu9-MBP. The averages are ordered by decreasing particle number.

Crosslinking mass spectromety

In collaboration with Dr. Julian Langer from the Proteomics Laboratory at the MPI of Biophysics and MPI for Brain Research, and Dr. Thomas Bausewein in our group, we investigated the interactions of the complex and pSu9-MBP through XL-MS. For this purpose, we crosslinked the complex with two different chemicals: 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) and disuccinimidyl sulfoxide (DSSO). EDC is a zero-length crosslinker that reacts with a carboxyl group and an amino group to form an amide bond. DSSO forms a 10 Å link between two amine groups and contains two cleavable bonds around a central sulfoxide. We tested both crosslinkers at different concentrations to find ideal working conditions, and resolved the samples on SDS-PAGE. Resulting gel bands were excised out and sent to Dr. Langer's laboratory for analysis. Digested peptide fragments were compared against the sequences of Tom70, Tom40, Tom22, Tom20, Tom7, Tom6, Tom5, pSu9 and VDAC from *N. crassa*.

DSSO crosslinking resulted in 48 links summarized in Figure 4.23A. Here we saw abundant croslinking between K33 in pSu9 to Tom40, Tom20 and Tom7, possibly reflecting different stages of translocation. We also saw interactions between K53 in Tom20 with Tom40, Tom22 and Tom7, shedding light on the location of Tom20 within the complex. Additionally, we obtained multiple hits of Tom70 with itself, meaning that we crosslinked its TPR-rich receptor domain.

Crosslinking with EDC, on the other hand, provided a cleaner insight with 244 hits. First, we found several self-links within pSu9-MBP and VDAC, but no interactions with each other nor with other subunits. Next, we saw a great number of links between Tom20 and Tom70, highlighted in gray in Figure 4.23B. This indicates a strong relation between both subunits. Tom70 also seems to interact with Tom40, Tom7 and Tom6. Tom70 produced many self-links within its soluble region. Similar to DSSO results, we saw many links between Tom20 (K53 and D181), Tom40 , Tom22 and Tom7, plus some interactions with Tom6.

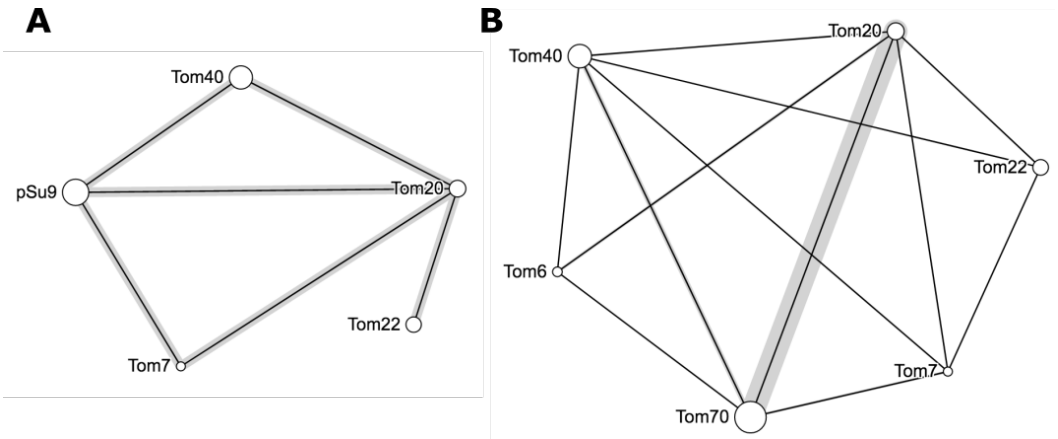

Figure 4.23: **Crosslinking of TOM holo with pSu9-MBP.** Crosslinks obtained from crosslinking TOM holo complex with pSu9-MBP using A) DSSO and B) EDC. The size of the circles relate to protein size, the gray shadows reflect the relative number of crosslinks between two proteins. Generated using xiNET.

Neither crosslinker provided hits involving Tom5, probably because it is mostly embedded in the membrane, and lacks lysines in its short soluble domain. It is worth keeping in mind that in both cases, the samples were crosslinked in solution, meaning that we cannot rule out link formation between two independent complexes.

Structure prediction

Given that we are interested in translocation, we also attempted to predict the interaction of pSu9-MBP with Tom40. Figure 4.24A shows a predicted model of pSu9-MBP. At a lower confidence rate, the presequence is a partial helix, shown in blue, with its C-terminal connected to a highly stable, fully folded MBP, in red. AlphaFold Multimer predicted pSu9-MBP inside the translocation pore, crossing at different

angles, delimited by the inner Tom40 helices. However, it unexpectedly appears to be inserted from the IMS side of Tom40, rather than from the expected cytoplasmic side (Figure 4.24B). Regardless, the contact points between both proteins may still reveal hints on interacting partners and can be used for comparison with cryoEM data. Although most of the prediction models showed a similar conformation, in a few cases pSu9-MBP appeared completely outside the pore.

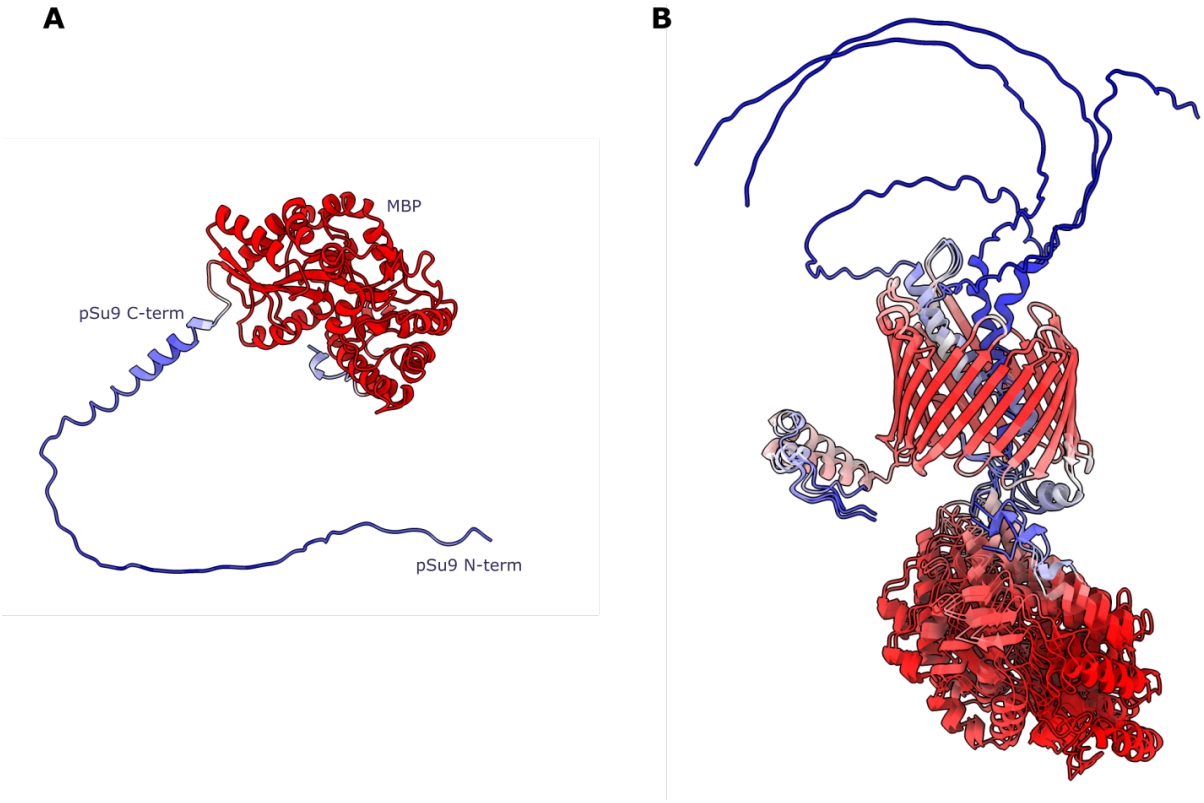

Figure 4.24: **Multimer prediction of Tom40 with bound pSu9-MBP.** A) AlphaFold prediction model of precursor construct pSu9-MBP. B) Four AlphaFold Multimer prediction models of the interaction between Tom40 and pSu9. Colored according to reported per-residue confidence estimated rate. Red indicates high-confidence, white, medium and blue, low.

4.3.2 TOM holo and pALDH

We tested the interaction of the complex with our shorter precursor construct, pALDH. Similar to our study with pSu9-MBP, we incubated TOM holo in GDN with pALDH and subjected the mixture to size exclusion chromatography. After separating unbound pALDH, we checked the eluted fractions by SDS-PAGE (Figure 4.25). Because pALDH is small, it runs together with the small TOM subunits (sT)

and it cannot be distinguished by SDS-PAGE alone. We confirmed elution of the precursor and the complex in the main peak by performing a dot-blot against α-StrepII antibody.

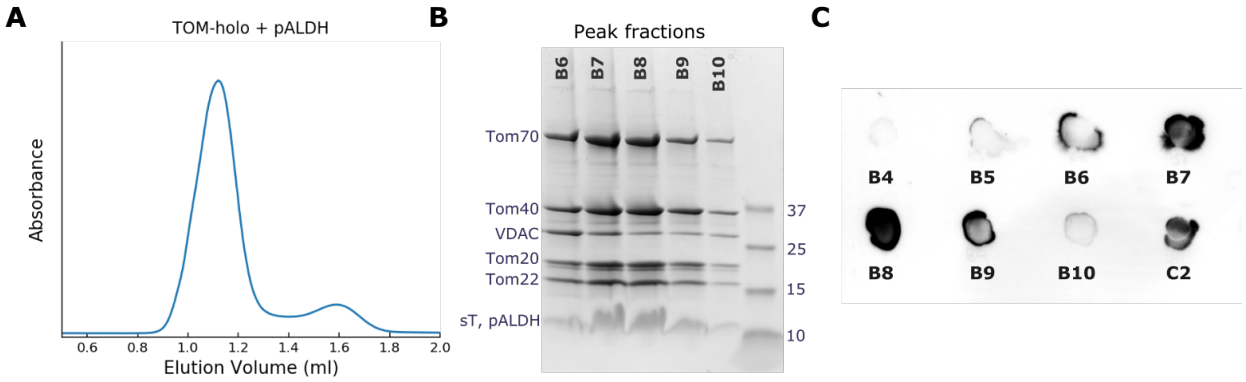

Figure 4.25: **Purification of TOM holo with pALDH.** A) Size exclusion chromatography of TOM holo complex with pALDH, using a Superdex 200inc column (Cytiva). B) Coomassie-stained SDS-PAGE of fractions corresponding to the size exclusion peak. C) Dot-blot incubated against α-StrepII antibodies.

After gel filtration, we analysed the sample through negative-stain electron microscopy. Obtained 2D averages showed a similar structure to that observed with the TOM holo complex alone (Figure 4.8). As the presequence only has a size of 3.5 kDa, and is small in comparison to the holo complex, it is difficult to visually confirm its binding only based on negative-staining microscopy.

Structure prediction

We generated a prediction model of pALDH (4.26A). The 32 residue construct resulted in a partially helical structure, with a surprisingly high confidence level. Likewise, we generated an AlphaFold Multimer prediction of Tom40 and pALDH. Figure 4.26B shows four models in which pALDH is bound to the pore. In this case, pALDH appears bundled closer to the cytoplasmic side of Tom40, apparently caught in mid-translocation towards the IMS side, delimited by the internal helices of Tom40. As before, some prediction models showed Tom40 completely independent of pALDH.

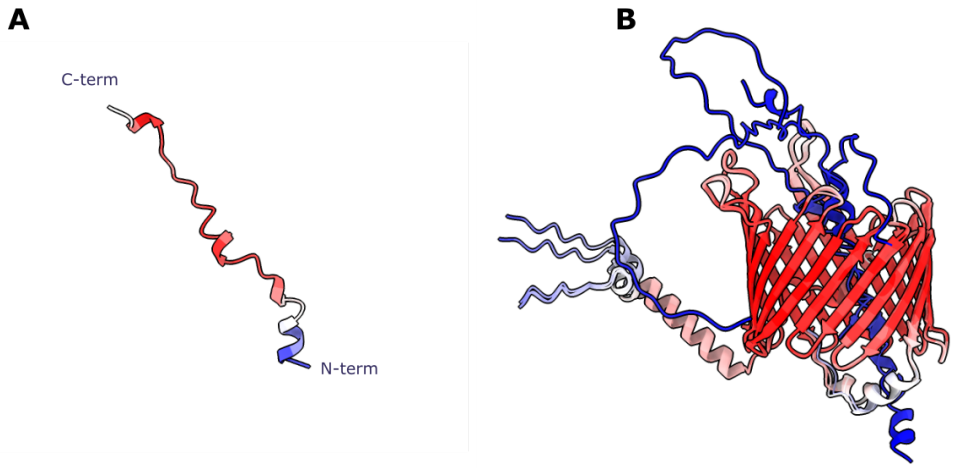

Figure 4.26: **Multimer prediction of Tom40 with pALDH.** A) AlphaFold prediction model of precursor construct pALDH. B) Four alphaFold Multimer models of the interaction between Tom40 and pALDH. Colored according to reported per-residue confidence estimated rate. Red indicates high-confidence, white, medium and blue, low.

In comparison to Figure 4.26A, all bound pALDH had low-confidence level and some lost all helical segments. Some prediction models show the presequence at a stage of apparent early translocation. Others show the N-terminal of pALDH peeking out of the pore, near the helical C-terminal helix of Tom40. In a few cases, the C-terminal of the presequence is inserted into the pore, rather than the N-terminal, depicting inverse translocation as seen with pSu9-MBP (Figure 4.24). Nevertheless, these results give us an insight into translocation pathways through the pore and will be useful in future model building.

4.4 CryoEM structure of TOM complex

4.4.1 CryoEM of TOM holo and pSu9-MBP

After purification of TOM holo complex with precursor pSu9-MBP (Figure 4.21), we froze the best elution fraction, previously determined by SDS-PAGE, with 1.8 mg/ml concentration, on a carbon-coated, copper mesh grid for cryoEM. Using a Krios microscope, we acquired 9,442 movies and picked close to 1,100,000 particles using a template picker in cryoSPARC. Figure 4.27 shows an example micrograph collected at 105,000x magnification, and some extracted particles. Interestingly, some

of these particles show one or two dots close to the cytoplasmic side of the complex, distinguished by the curvature of the micelle.

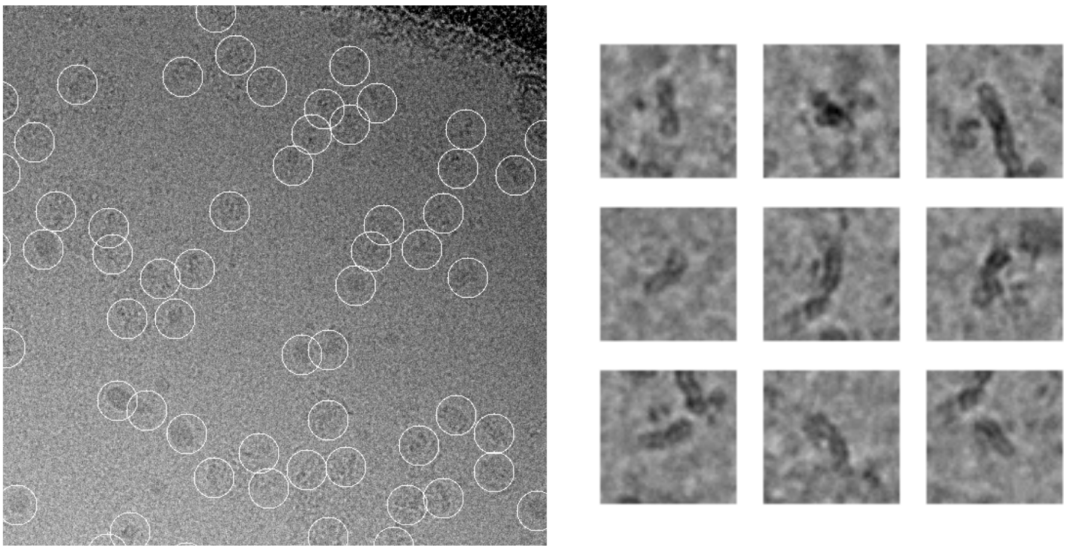

Figure 4.27: **CryoEM of TOM holo and pSu9-MBP.** Left: Micrograph of frozen TOM holo complex in GDN incubated with pSu9-MBP at 105,000x magnification. Right: Example particles of the complex.

After a single round of sorting, we separated particles corresponding to noise from the 996,000 protein-containing particles. Figure 4.28 shows selected averages obtained from 2D classification in cryoSPARC. The side views of the complex were abundant and easy to identify. Note that some of them display a diffuse cloud emanating from the cytoplasmic side of the complex, as seen in classes 1 and 4. Classes 5 and 10 show a more defined feature in this position.

Top views were less abundant. Class 12 shows a top view, with two clear pores centered in the micelle, and another dark feature on the side, resembling a third pore. This will be discussed in the next chapter. Furthermore, classes 7, 9 and 11 show a protein of smaller size than TOM complex. These probably correspond to MBP monomer and dimers, as we observed in negative-stain imaging(Figure 4.22). Lastly, classes 3, 6, and 8 portray the complex from alternative sides.

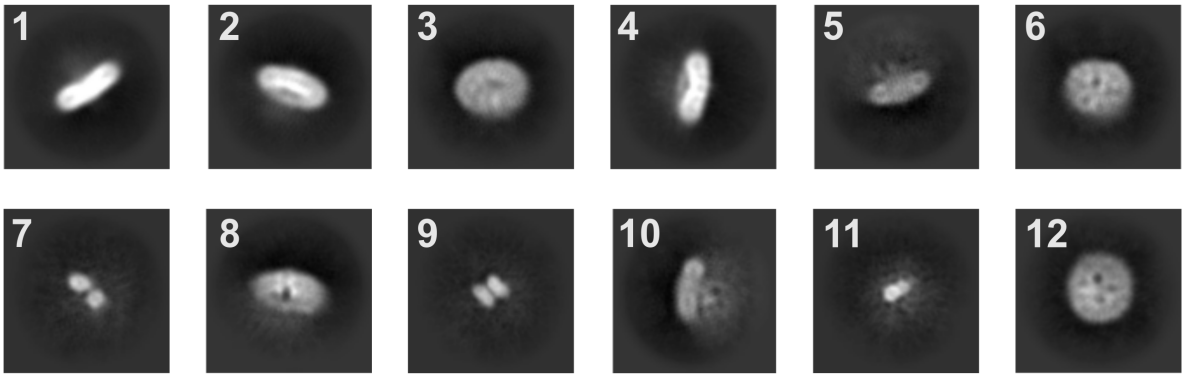

Figure 4.28: **2D averages of TOM holo with pSu9-MBP** Selected averages from the 2D classification of nearly one million particles containing TOM holo complex and pSu9-MBP. The averages are ordered according to their decreasing particle number.

After further 2D classification, we selected a set of 250,000 particles that best resembled the dimer complex and separated them in 3D through ab-initio reconstruction and heterogeneous refinement. We used the best class, containing half of the particles, for 3D variability analysis, which helped us separate particles into three clusters (Figure 4.29). The first cluster contained 35% of the particles and was refined to 5.2 Å resolution after non-uniform refinement in cryoSPARC (green). At a low density threshold, its micelle was round with the complex positioned at its center. At high-density threshold, the TOM core dimer, with its five subunits per protomer, was clearly visible(Figure 4.30).

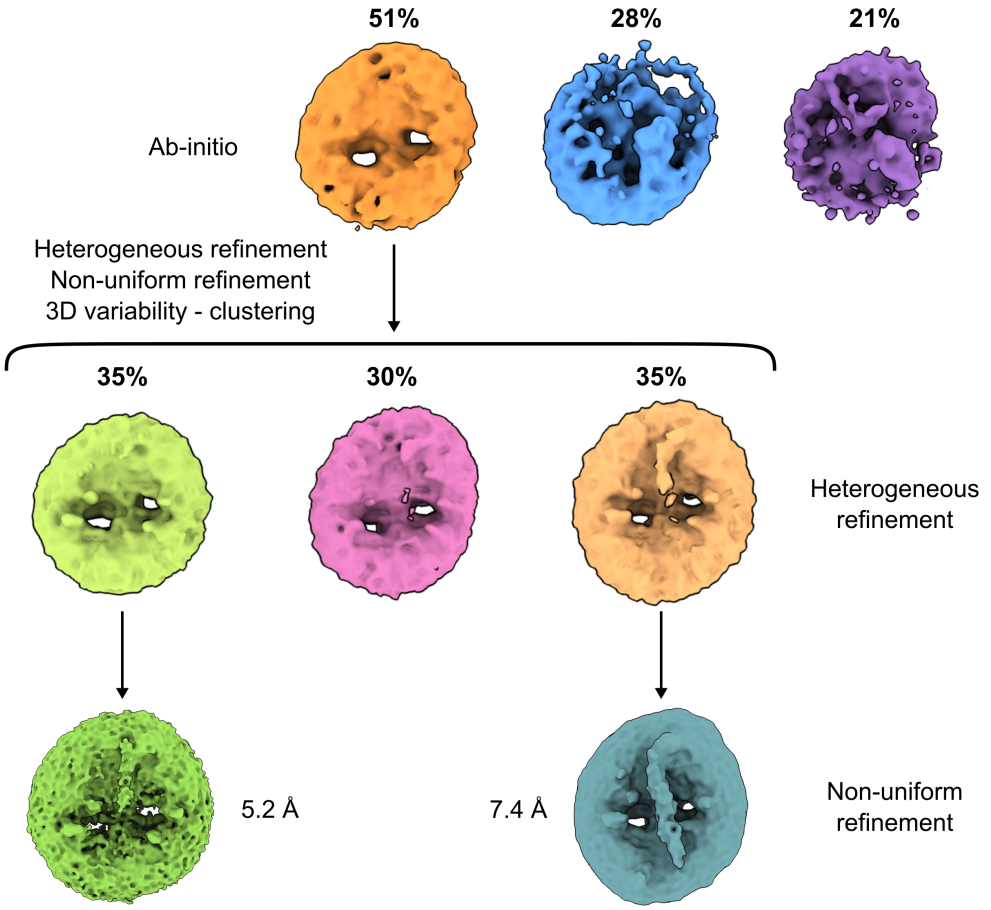

Figure 4.29: **3D classification of TOM holo with pSu9-MBP.** Ab-initio and 3D variability classification of preprotein bound TOM data set, leading to two final maps at 5.2 Å and 7.4 Å resolution.

Following the subunit assignment made by Bausewein et al., (2017), we identified both Tom40 translocation pores, as well as their three helices, although individual β-strands in the β-barrel are not clearly distinguishable. Two copies of Tom22 span the membrane and protruding out from the micelle into the IMS side of the complex. On the cytoplasmic side however, one copy of Tom22 is larger and curves towards one side. Likewise, we identified Tom5, Tom6 and Tom7, which display their characteristic shapes and are contained within the micelle, yet one Tom5 is better resolved than the other.

The second cluster in Figure 4.29, with 30% of the particles (pink), aligned poorly and as a result displayed little features resembling TOM complex at a high density threshold. The third cluster, however, with the remaining 35% particles (orange), had

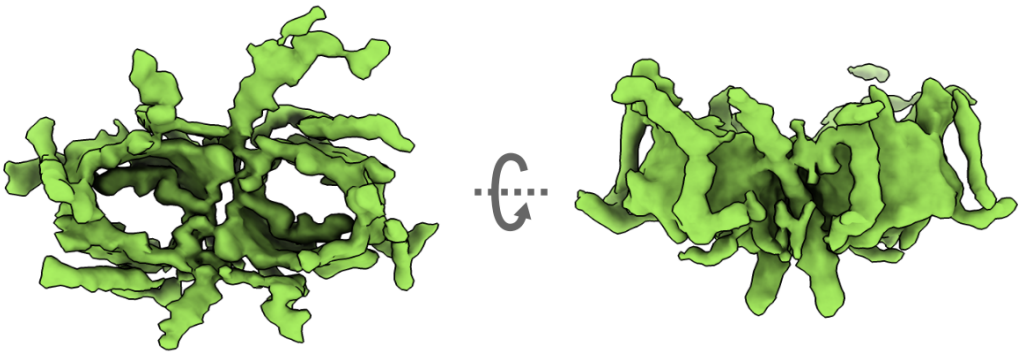

Figure 4.30: **Mid-resolution TOM core complex.** We resolved TOM core complex from our pSu9-MBP data set at a 5.2 Å resolution. Although individual residues are not distinguishable in the map, we found the five core subunits in the complex.

a larger micelle and included an elongated feature extruding towards the cytoplasmic side of the complex. After refinement, a 7.4 Å map showed a poorly resolved TOM core complex at a high density threshold. At a low density threshold, the cytoplasmic elongated feature became more evident. It rises from the edge of the micelle, almost aligned to one pore and crosses towards the other end. It is interesting to note that the map is not symmetric, meaning that this new density does not end at the exact opposite side from where it begins.

Compared to the core complex, there are three new entities in the data set, Tom20, Tom70 and pSu9-MBP. Our prediction model for Tom70 was far too large to fit into this elongated feature (Figure 4.14), and we discarded it as a candidate. However, the cytoplasmic domain of Tom20 has a similar size (Figure 4.16). The position of the emerging feature and its distance to the pores resembles that of our TOM core + Tom20 multimer prediction model, shown in Figure 4.18. The density is, however, too long for Tom20 alone. This suggests that Tom20 could be interacting with either a less defined Tom20 on the opposite side, or perhaps pSu9-MBP.

We further processed particles containing this feature. After a second round of refinement with limited alignment resolution, we obtained a 10 Å map that allowed us to see a second feature attached to the complex at lower density threshold (Figure

4.31). This new globular density is attached to the top of the elongated rod, and lies between the two translocation pores. This rounder density, located on the cytopalsmic side of the complex, may belong the bound pSu9-MBP.

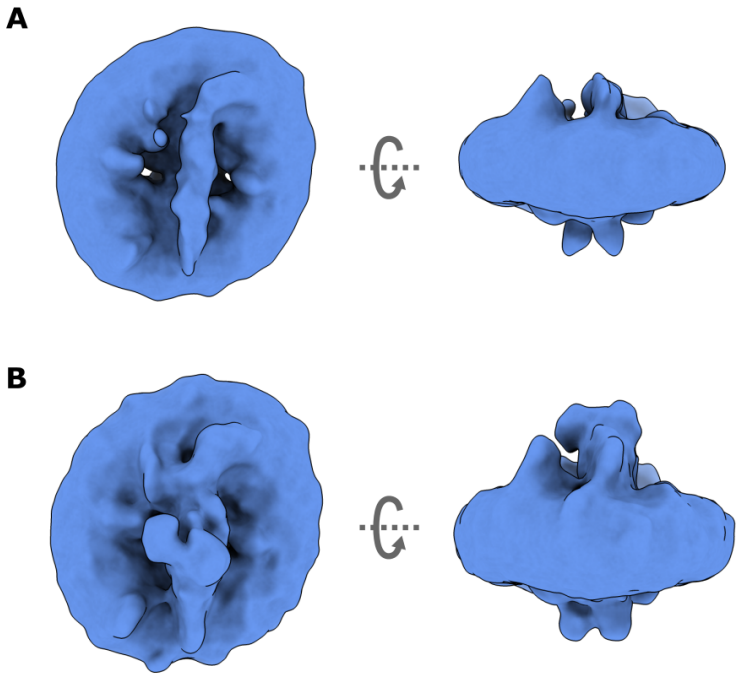

Figure 4.31: **Low-resolution structure of TOM with pSu9-MBP.** A) 10 Å resolution map of the TOM holo complex sample with an elongated density protruding from the micelle. B) The same map at lower density threshold shows an additional globular density bound to the complex.

We returned to the 2D classes containing smaller elements resembling the precursor protein by itself (Figure 4.28) and processed 350,000 particles in 3D. We obtained monomeric and dimeric maps with dimensions of 35 Å by 60 Å that comfortably fit our predicted MBP domain. Both maps, obtained from roughly 100,000 particles each, have 10 Å resolution and are shown in Figure 4.32.

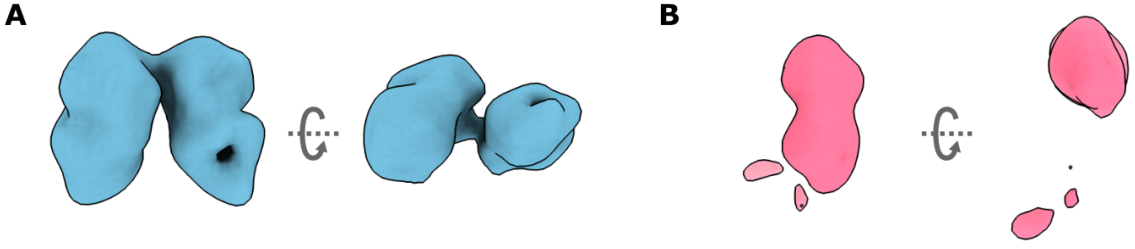

Figure 4.32: **Low-resolution structure of MBP** A) Dimeric and B) monomeric structure of MBP obtained from our TOM holo complex with pSu9-MBP data set.

We overlaid MBP monomer onto our TOM holo structure (Figure 4.33). We saw that the globular density in the TOM map closely resembled monomeric MBP, suggesting interaction of the presequence with the complex. More specifically, a presquence bound Tom20. At lower density threshold, it is possible to see thin rods connecting the cytoplasmic features to the core complex, near Tom5 and Tom7, which could mean that pSu9 is approaching the translocation pores. However, at this resolution, we can only speculate.

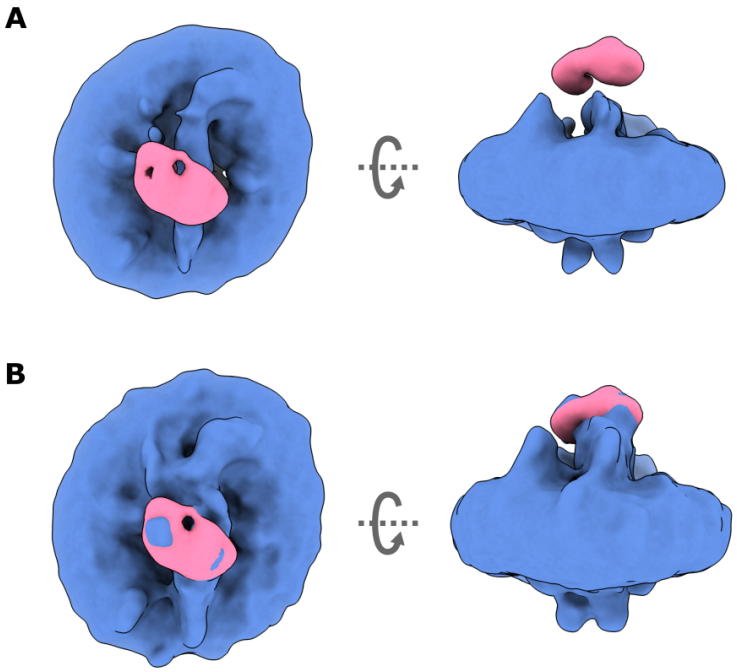

Figure 4.33: **MBP bound TOM complex.** Monomeric map of MBP overlaid to our 10 Å TOM holo complex map at A) low and B) lower density threshold.

4.4.2 CryoEM of TOM holo and pALDH

After gel filtration, we froze pALDH-bound TOM holo complex at 2 mg/ml protein concentration in copper mesh grids for cryoEM. We recorded 22,000 movies at 105,000x magnification in a Krios microscope. Figure 4.34 shows a representative micrograph and extracted particles, some of them showing small domains protruding from the cytoplasmic side of the complex. These domains appear in different positions. Since our pALDH precursor protein is a small peptide chain of 3 kDa, it is likely that these protruding densities correspond to Tom20 and Tom70.

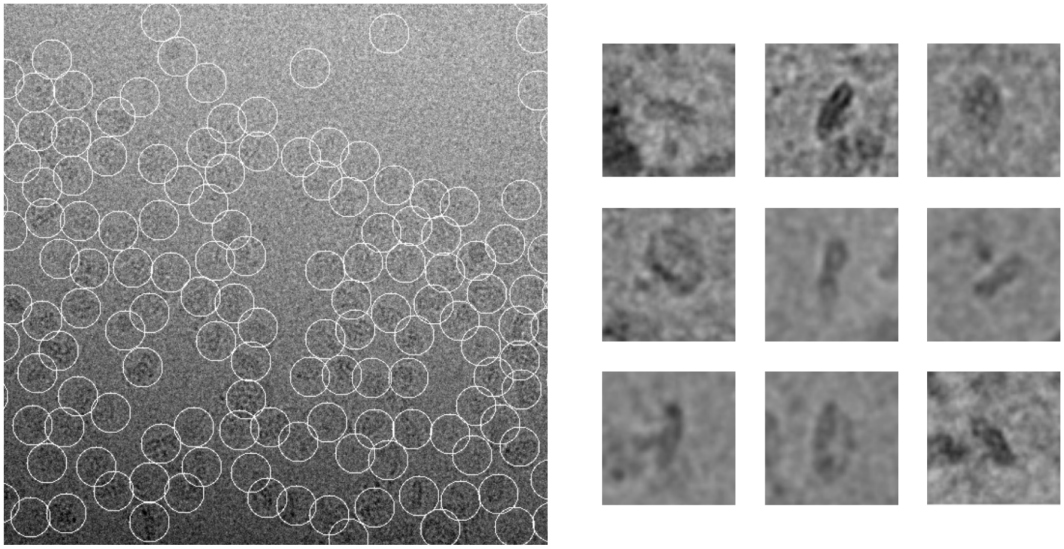

Figure 4.34: **CryoEM of TOM holo with pALDH.** Left: Micrograph of frozen TOM holo complex in GDN incubated with pALDH at 105,000x magnification. Right: Example extracted particles in different orientations.

We imported the movies into Relion and picked particles using a trained Cryolo model. We extracted nearly 2 million particles in a four times binned box and classified them in 2D to clear out bad non-protein particles and noise. Figure 4.35 shows 2D class averages of the remaining 1.5 million particles. Similar to the negative stain data set of TOM holo complex (Figure 4.8), we saw a mixture of particles of different sizes, indicating the presence of both dimer and tetramer complexes in the sample.

When analysing top views in Figure 4.35, we can see apparent heterogeneity. Class 3 shows two pores towards one side of the micelle, accompanied by a third feature at its center. Classes 4 and 12 show two clear pores at the center, but their micelles are very different in shape. As for the side views, classes 9 and 11 seem to have some fuzziness emanating from their cytoplasmic side. In contrast, classes 5 and 15 lack the fuzzy cloud. This could indicate that whatever the cause of this cloud is, the precursor, Tom20 or Tom70, is not present in all particles in the data set. The larger classes, in contrast, show less details of the protein densities inside micelles. Classes 8 and 14, for example, seem to have a large fuzzy cloud attached to them, making it difficult to get an idea of any internal structure of the particles.

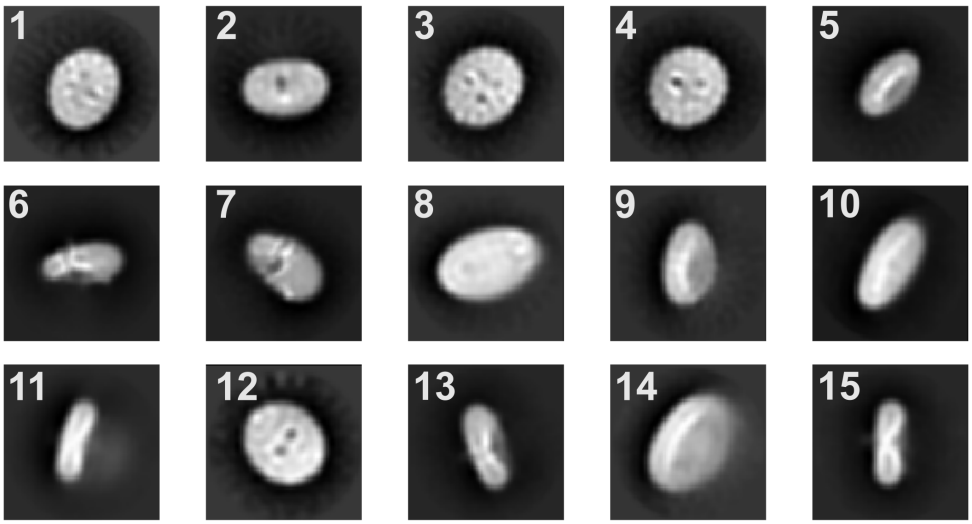

Figure 4.35: **2D averages of TOM holo with pALDH.** Selected averages from 2D classification of 1.5 million particles containing TOM holo complex with pALDH. The averages are ordered according to their decreasing particle number.

We then re-extracted particles at 1.5x bin and separated them into six classes in 3D. 40% of the particles fell into a well defined class (purple in Figure 4.36). In this class, two characteristic pores of TOM dimer are clearly visible, as well as the helical IMS termini of Tom22. At high-density threshold, the map discernibly shows the five subunits of the core complex. Additionally, at low-density threshold, this class showed a protrusion on its cytoplasmic domain, similar to one seen in the pSu9+MBP dataset. The second largest class in Figure 4.36 (soft pink) also showed two clear pores, Tom22 IMS termini, and the cytoplasmic feature. However, its micelle is considerably longer and its high-density features were less resolved.

We continued to work on the largest class (purple), containing 620,000 particles; processing of the rest of the particles will be further discussed at the end of this chapter. Following 3D refinement, Bayesian polishing and postprocessing, the quality of the map improved considerably, reaching a final resolution of 4.2 Å. This process helped better resolve the cytoplasmic density, which took the shape of two elongated rods ending in a small globular domain (Figure 4.37A). Both rods seem to emerge together from the edge of the micelle, and diverge as they approach the pores. Additionally, we observe two short rods parallel to the membrane, corresponding to the cytoplasmic

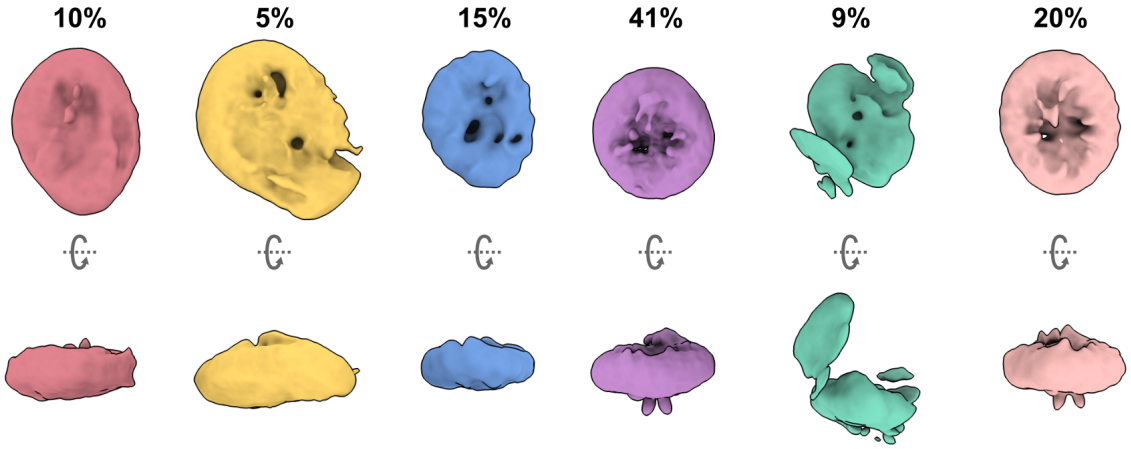

Figure 4.36: **3D classification of TOM holo with pALDH.** 1.5 million particles classified in 3D into 6 classes. A large portion of the particles fell into a class clearly depicting TOM dimer, with hints of a cytoplasmic domain.

ends of Tom5, as well as a small bump, corresponding to the largest loop in Tom40. In contrast, at high-density threshold (Figure 4.37B), the map conserves only features corresponding to the five components of TOM core complex.

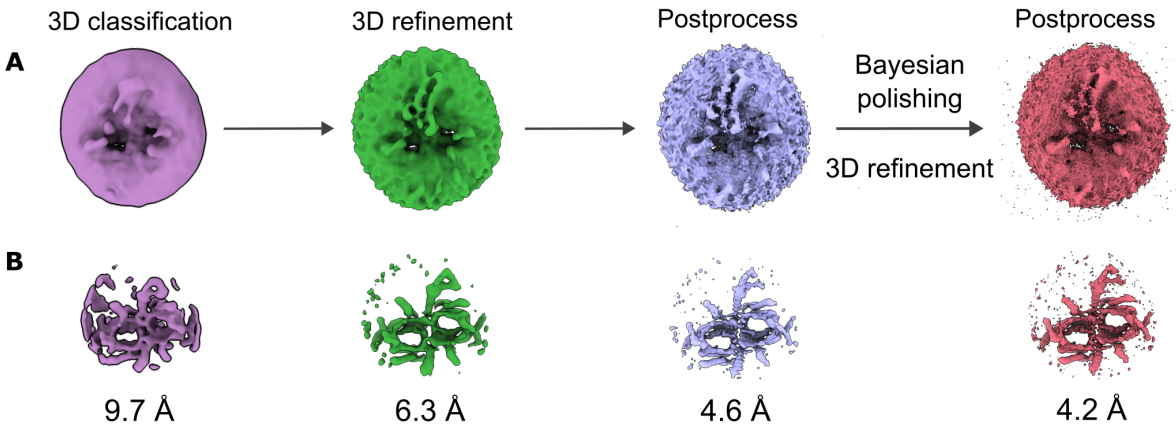

Figure 4.37: **Refinement of TOM holo with pALDH.** After refinement and bayesian polishing in Relion, we obtained a 4.2 Å resolution map showing TOM core complex with a better defined cytoplasmic domain.

Subsequently, we polished the particles in their original box size, i. e. without binning, and refined and postprocessed the complex with a local mask, excluding the detergent bell. This led to a small improvement in resolution up to 4.1 Å (Figure 4.38). The high-density map shows better defined helices, but no individual β-strands in Tom40. Interestingly, this map also shows one elongated copy of Tom22. A longer

cytoplasmic Tom22 is not surprising, as the termini of Tom22 are longer, but mostly unstructured. As we saw in AlphaFold, a short helix was expected at the membrane-cytoplasm interface of Tom22, perpendicular to its intermembrane helix (Figure 4.19). A more structured Tom22 was also observed in human TOM core complex (*27*).

Furthermore, in Figure 4.38 we see a short elongated density attached to the visible end of Tom22. This is the origin of the cytoplasmic domains we see at low-density threshold. As mentioned in the previous section, this cytoplasmic domain resembles Tom20 both in size and shape. However, as shown in Figure 4.37, at low-density threshold data suggests the presence of two copies of Tom20, one next to the other, completely breaking with the dimeric symmetry observed in the core complex.

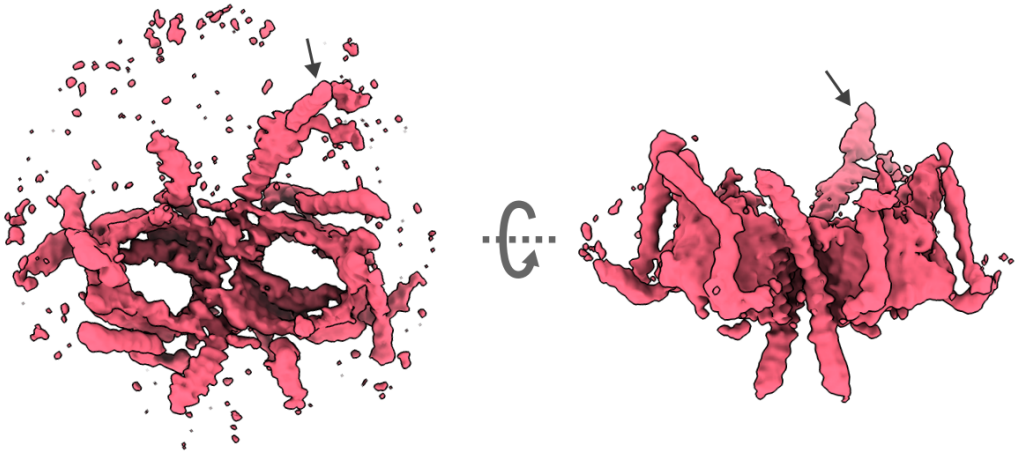

Figure 4.38: **Mid-resolution TOM core complex.** At 4.1 Å resolution, we resolved TOM core complex with a longer N-terminal Tom22 and a small density protruding towards the cytoplasm (gray arrows).

In the following sections, I will describe how these 620,000 particles (Figure 4.37 were used to obtain four TOM maps (Figure 4.39): one of TOM core complex at 3.3 Å resolution, one showing the interaction of precursor protein pALDH with Tom40 at 4 Å resolution, and two maps of Tom20 docked to TOM core at 6-7 Å. Finally, in the last section, I will discuss about particles that were excluded after 3D classification and the information gained from them.

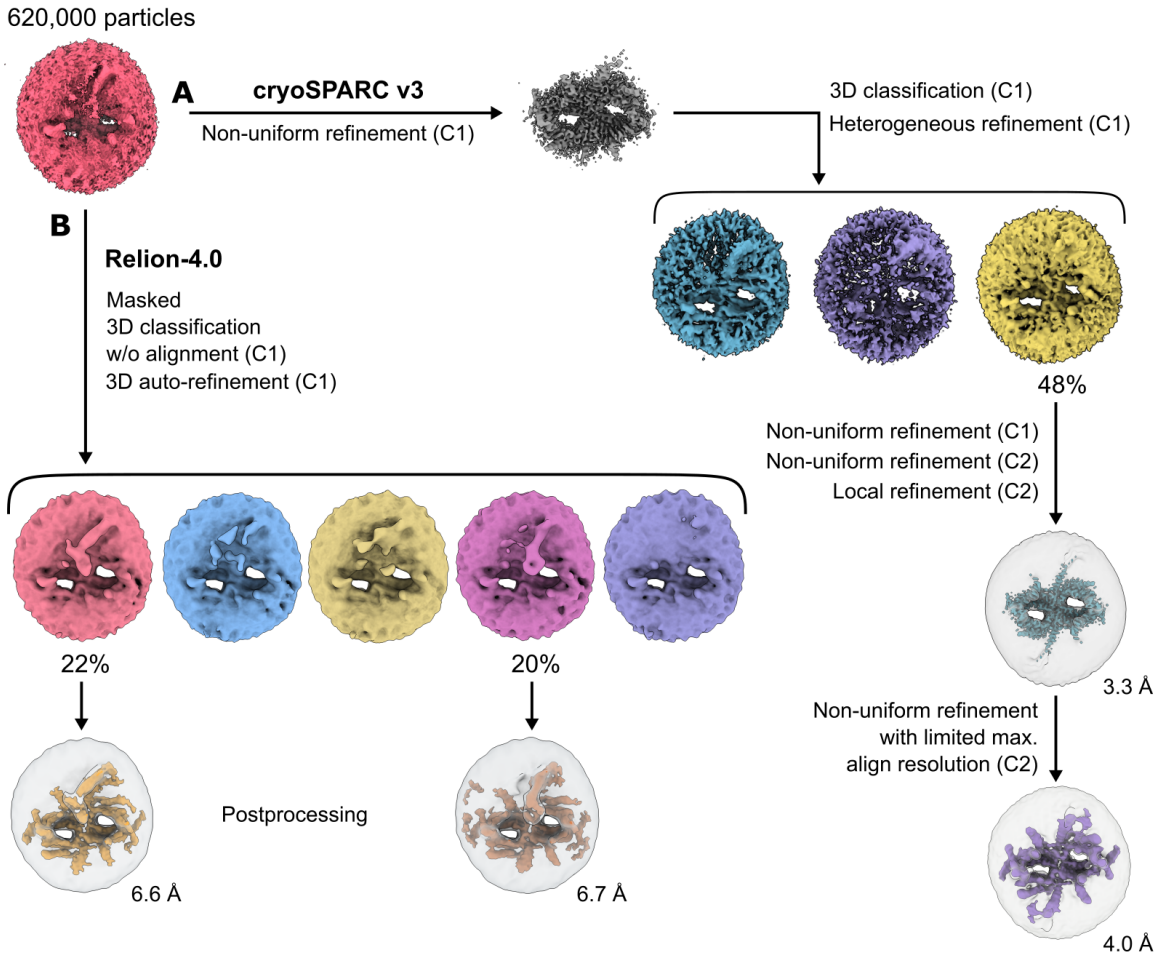

Figure 4.39: **Processing of TOM holo with pALDH.** Classification and refinement of TOM holo with pALDH data in A) cryoSPARC and B) Relion, resulting in four maps of TOM complex.

TOM core complex

Seeing no further improvement in resolution in Relion, we continued processing unbinned, polished particles in cryoSPARC. We refined the particles on their corresponding map, imported from Relion (Figure 4.38). This led to an improved map at 3.5 Å resolution, where individual β-sheets of Tom40 were visible, and all α-helices were clearly distinguishable. However, the increase in resolution came at the price of the complete absence of the Tom20 density and the elongated Tom22 cytoplasmic densities observed in Relion.

We performed a 3D classification without alignment of the particles with a global mask, followed by heterogeneous refinement of all classes (Figure 4.40). Two classes showed broken maps, but nearly half of the particles fell into the best class (yellow), showing a clear TOM core. This scheme corresponds to path A in Figure 4.39.

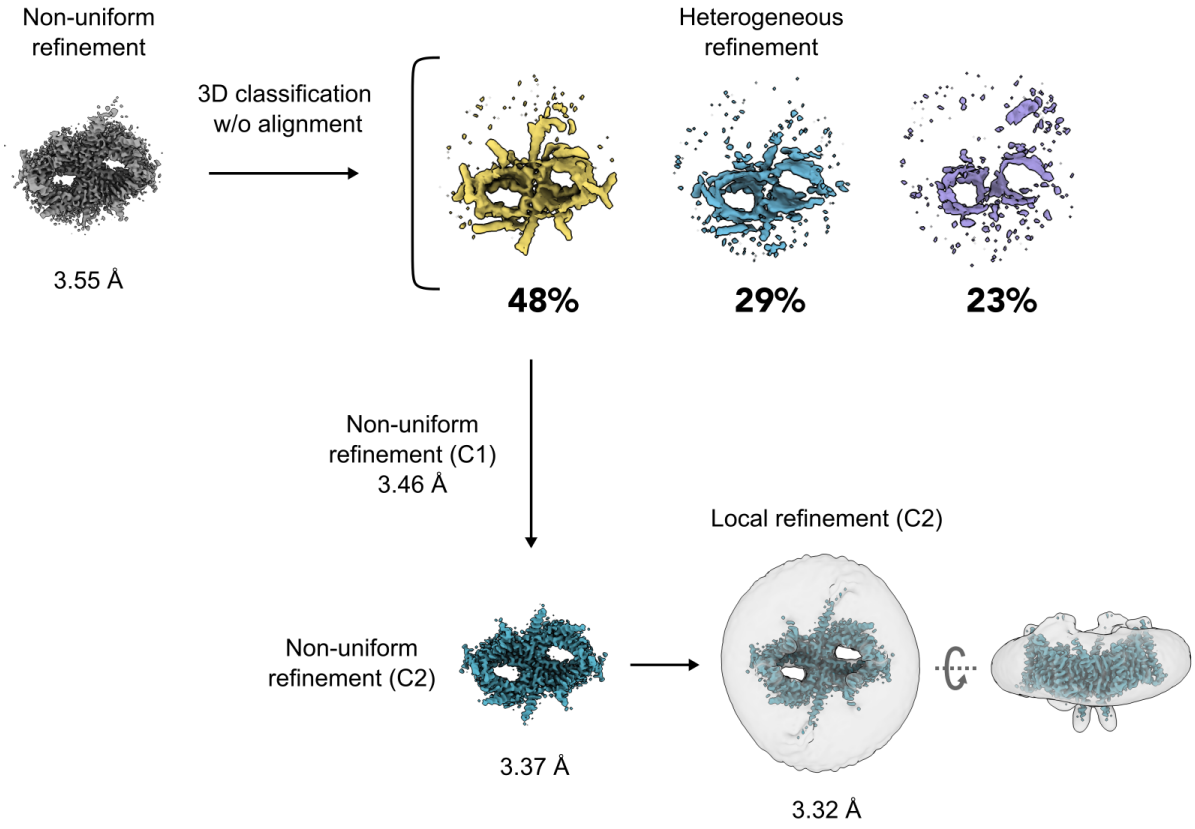

Figure 4.40: **Processing of TOM holo with pALDH in cryoSPARC.** Classification and refinement of TOM holo with pALDH particles in cryoSPARC, resulting in a high-resolution map of TOM core complex at 3.3 Å.

We then refined the 304,000 particles in cryoSPARC, applying C2 symmetry only in the last two refinement steps. Finally, we obtained a 3.3 Å structure of *N. crassa* TOM core complex. This structure lacks the cytoplasmic Tom20 features observed before. Local-resolution estimation shows the resolution of the map ranges mostly between 2.5 Å and 3.5 Å (Figure 4.41). The best resolved areas encompass the inside of both β-barrels, including the inner Tom40 α-helix, hinting at a very stable translocation

pore. The less-resolved areas are, as expected, the truncated termini helical subunits at both cytoplasmic and IMS sides of the complex.

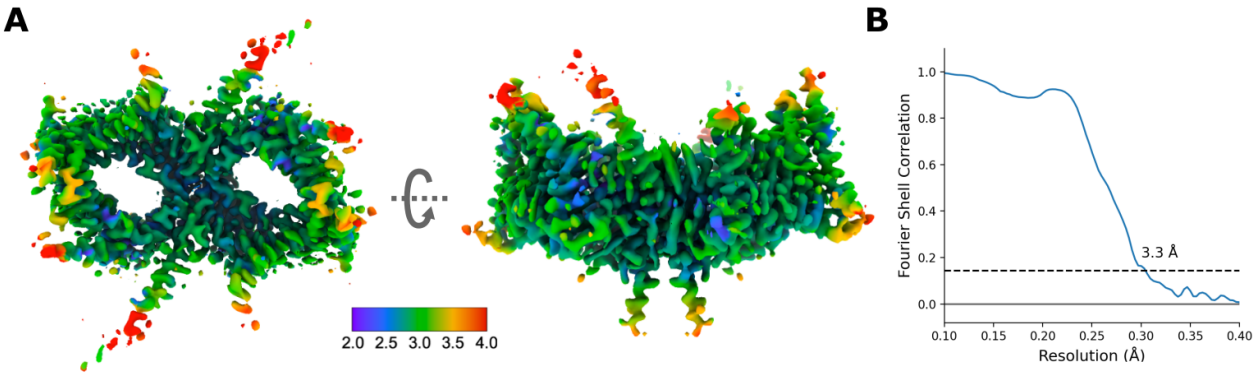

Figure 4.41: **High-resolution TOM core complex.** A) Local-resolution estimate of our final TOM core complex map ranges from 2 Å to 4 Å, with better definition inside of the pores, and poorer definition at the cytoplasmic and IMS termini of the helices. B) Fourier shell correlation of final local refinement.

At this resolution we were able to reliably model the five subunits of TOM core complex based on our previously published map (*23*). We used the AlphaFold prediction model of the core complex (Figure 4.11) as a base. We relaxed the model in our map using the recently developed model building software ISOLDE. We then fixed individual residue clashes in Coot and performed a final real-space refinement in Phenix. Figure 4.42 shows our final 3.3 Å resolution map and corresponding model, with individually colored subunits: Tom40 in yellow, Tom22 in blue, Tom7 in violet, Tom6 in green and Tom5 in pink.

As described by Bausewein et al. (*23*), two translocation pores are tilted at a ∼20° angle relative to the membrane normal. In between the pores, a V-shaped density at the symmetry axis runs parallel to both Tom22 helices. This density is also present in our maps refined without imposed symmetry, and has been observed in yeast TOM (*25*, *26*). We identified this density as a phospholipid and modelled it as phosphatidylcholine (PC). PC is kept in place by residue F309 in both Tom40 barrels. We observed two rotamer conformations for this reside.

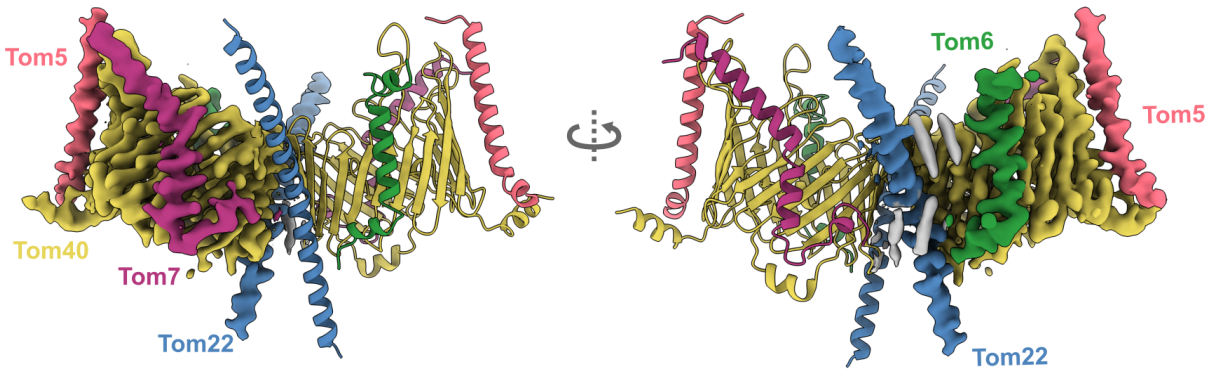

Figure 4.42: **TOM core complex model.** Composite structure of our TOM core complex map and model. Five Tom subunits are indicated as follows: Tom40 in yellow, Tom22 in blue, Tom7 in violet, Tom6 in green, and Tom5 in pink. Detergent and lipid molecules appear in gray.

On the outside of the complex we identified four large, rod-shaped densities that surround each Tom22. They span half of the membrane and were modelled as diosgenin, the backbone of the detergent used for solubilization, GDN (Figure 4.43B). We additionally found 10 small elongated densities on each side of the complex that we did not model. In its native environment, these densities may be replaced by PC, posphatidylethanolamine (PE) or cholesterol (*115*).

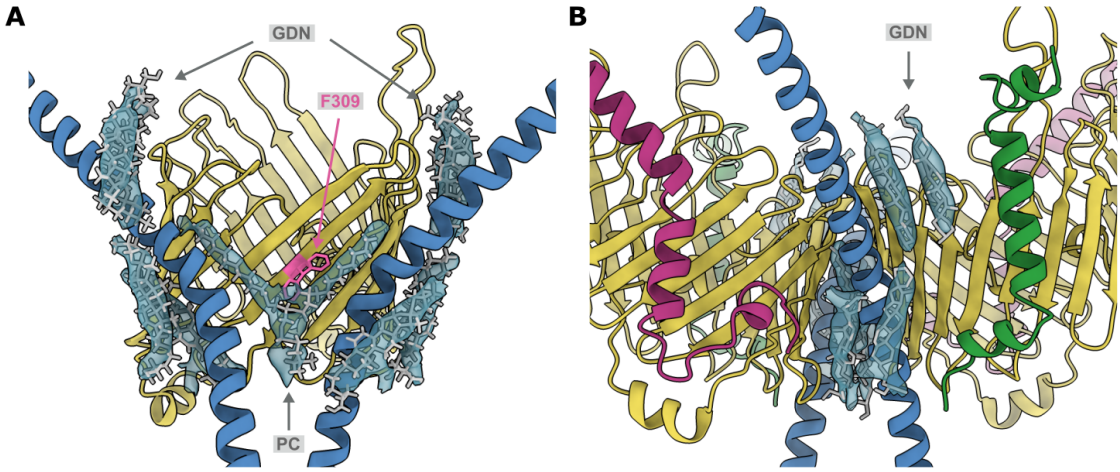

Figure 4.43: **Lipid and detergent surround the core complex.** Nine densities surround the subunits in the map. A) Phosphatidylcholine (PC) sits in the center of the complex and is held in place by F309 in Tom40 (pink). B) Four detergent (GDN) molecules surround Tom22 on each side of the complex.

Our map contains two copies of **Tom40**, each consisting of 19 β-strands and three α-helices. Figure 4.44 shows the N-terminal end of Tom40 in light orange, starting with two helices, α1 in the IMS, followed by α2 inside the pore. α2 crosses Tom40 from one side to the other, from IMS to cytosol, and limits its inner size to 35 Å by 20 Å, followed by β1, which marks the start of the barrel. In dark orange, Figure 4.44 shows β19 and α3, the C-terminal of Tom40. β19 interacts with β1 to close the barrel. α3 dips into the IMS and ends in a stable loop inside the pore, opposite to α2. At the cytoplasmic interface, loops between the Tom40 β-strands are long and interact with small TOM subunits. The longest loop, between β14 and β15, protrudes out of the membrane and is only visible in our map at lower threshold levels, indicating flexibility, in agreement with local resolution estimation.

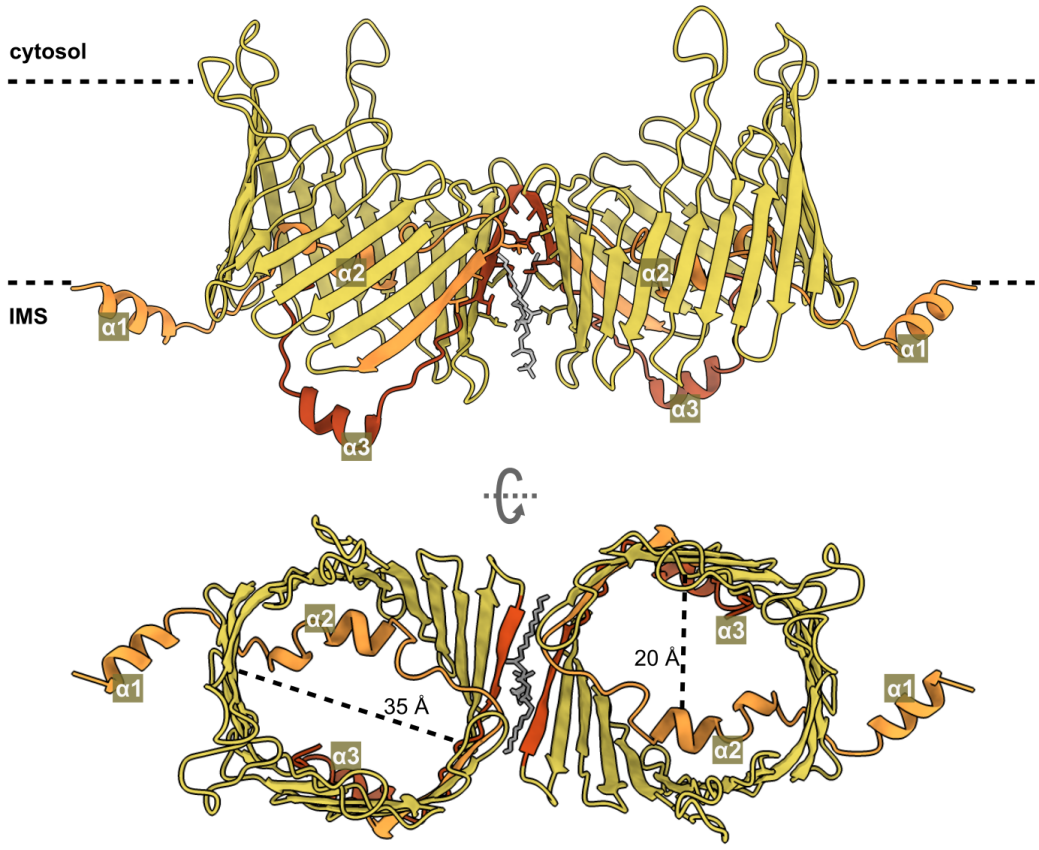

Figure 4.44: **Tom40 model.** Two pores of the TOM dimer are separated by a molecule of phosphatidylcholine, giving the barrels their tilted positions. The N-terminal of each barrel, including α1, α2 and β1 are colored in light orange. Their C-terminals, including β19 and α3 are colored in dark orange.

The symmetry axis of the dimer is at the point where both Tom40s meet. The barrels are separated by PC and only interact with each other on the cytoplasmic side: loops connected to strands β1 and β2 in one Tom40 interact with β19 of the other Tom40. This is also the area where β1 and β19 interact in each pore, marking the start and end of the barrels. PC spans strands β17, β18 and β19 and is confined by Tom22.

Two copies of **Tom22** extend from cytoplasm towards IMS at the dimer interface, stabilizing the tilted position of both Tom40 pores. Each copy of Tom22 is in contact with Tom40's β16-18 through hydrophobic interactions, and is surrounded by detergent and lipid molecules (Figure 4.45). The N-terminal domain of Tom22 is unstructured, flexible, and mostly invisible in our map. At the cytoplasmic inteface, Tom22 forms hydrogen bond interactions with two Tom40 loops. The Tom22 main helix transverses the membrane and kinks outwards at position P99. On the IMS, Tom22 protrudes out of the membrane with its disordered C-terminus unresolved.

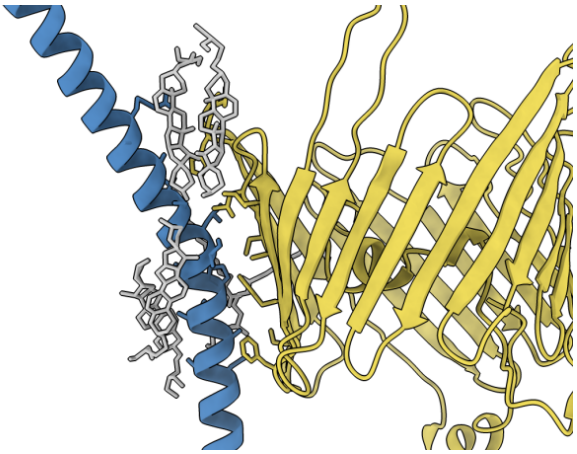

Figure 4.45: **Interactions between Tom22 and Tom40.** Tom22 stabilizes the tilt between pores, as it interacts hydrophobically with Tom40 and lipids molecules.

Opposite to Tom40 α2, we can see the unusual structure of **Tom7** in each monomer (Figure 4.42). Tom7 interacts with β1-6 in Tom40, and is kept in place by hydrophobic and hydrogen interactions between both subunits. The N-terminal of Tom7 is a helix that runs parallel to the membrane surface. A second helix transverses the membrane and reaches the IMS. Its C-terminal folds back into the membrane and is

mostly unstructured. This stretch of Tom7 contains three proline residues, giving it its characteristic shape, and is held in place by hydrogen bonds.

On the other side we find the transmembrane domains of **Tom6**. Starting at the cytoplasm, the N-terminal of Tom6 is likely flexible and not visible in our map. Tom6 continues with a helix that forms hydrogen interactions with the longest loop in Tom40, and then bends down into a second helix inside the membrane. Finally, at its C-terminal, Tom6 ends with a helix parallel to the membrane-IMS interface.

Two copies of **Tom5** are located at the extremes of the complex, attached to Tom40. We were able to partially model its N-terminal, which points towards the center of the complex. Tom5 is composed of a single helix perpendicular to the membrane that kinks outwards at the IMS interface due a proline residue. It interacts with Tom40 strands and α1 helix, mainly by hydrophobic interactions and hydrogen bonds.

pALDH bound TOM core complex

We continued to process TOM core particles and map, and refined them with limited alignment resolution and imposed symmetry. This led to a 4 Å resolution map showing hints of a precursor-bound Tom40. At low-density threshold (Figure 4.46, left), this map contains an elongated, thin rod, crossing through the inside of both Tom40 pores, from cytoplasm towards IMS. We suggest that this density corresponds to bound pALDH during translocation. However, this density is only visible at low-density threshold (Figure 4.46, right), indicating low-binding affinity of the precursor to the complex or that the precursor displays multiple binding modes. We nevertheless attempted to rigid-body-fit a prediction model of our pALDH construct into the map, superimposed to our TOM core model.

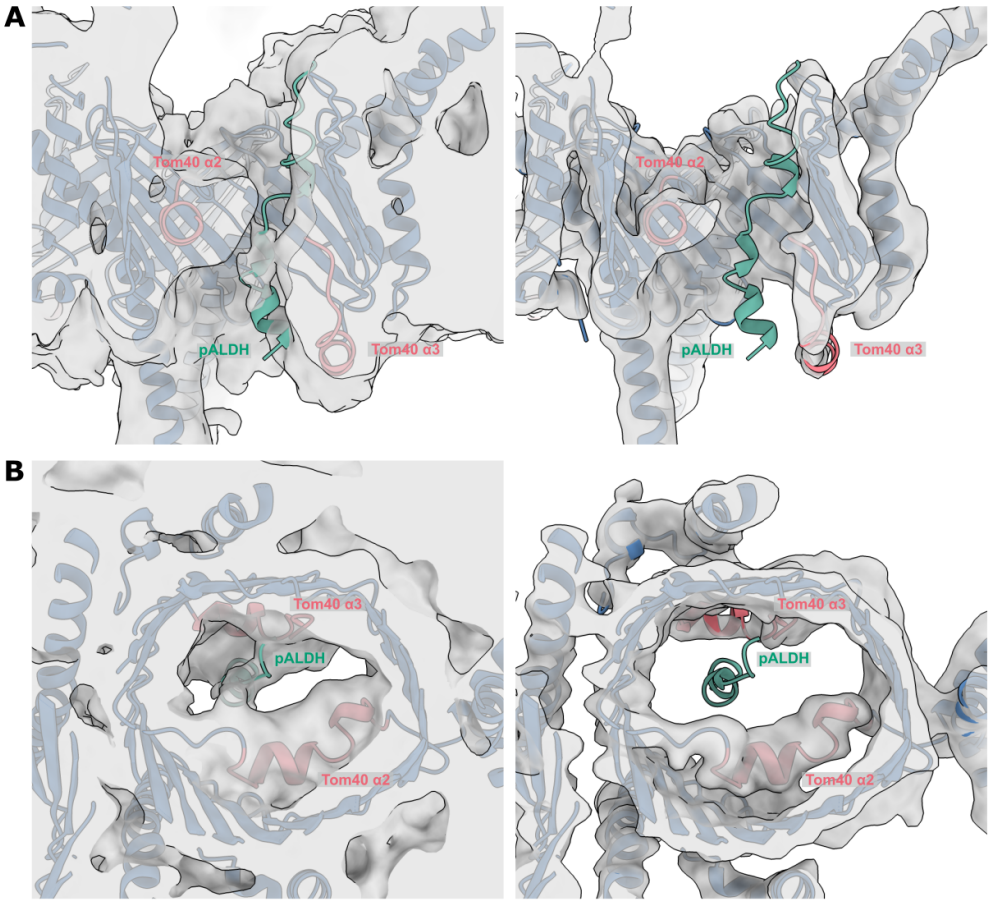

Figure 4.46: **Structure of pALDH bound TOM core complex.** Model-map superposition of precursor bound complex at lower (left) and higher (right) density thresholds. The rigid-body-fitted pALDH model generated with AlphaFold appears in teal, our TOM core model in blue, the 4 Å map of the TOM core complex in gray, and Tom40's α2 and α3 in pink. (A) Side view of the precursor inside Tom40 mid-translocation. (B) View of the interaction from the cytosol.

The map shows that pALDH extends from α2 to α3 inside Tom40, therefore, we were only able to fit the presequence within this region. Tom40 α2 and α3 are rich in hydrophobic residues. In the model-map superposition, we identified three points where the presequence density makes contact with Tom40. Figure 4.47 highlights three residues that come close to the pALDH density. In α2, Y60 seems to point inwards, likely guiding the precursor as it enters the pore. On the IMS side, L335 in α3 points upwards, marking an exit point for the precursor. Most interestingly, at the end of α3, F349 is an intermediate contact site. We can see how this phenylalanine, which is the last residue in Tom40, points towards the center of the pore. These residues may be important in the gating of precursor proteins.

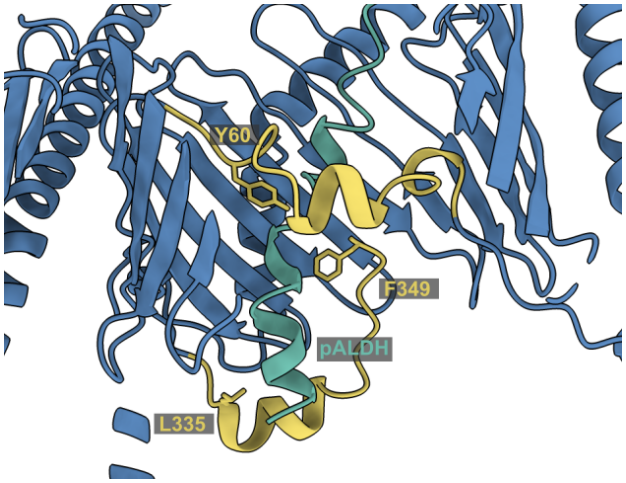

Figure 4.47: **pALDH bound TOM model.** Side view interaction between Tom40 and pALDH. Rigid-body-fitted model of pALDH (teal) inside TOM core complex (blue). Three residues in $\alpha 2$ and $\alpha 3$ (yellow) are at a close distance from pALDH: Y60, L335 and F349.

TOM holo complex

By further processing the original 620,000 particles in Relion, we were able to extract more information on the structure of TOM holo complex (path B in Figure 4.39). As shown in Figure 4.37, maps generated in Relion contained two rod-like features in the cytoplasmic side of the complex, similar to those observed in the TOM with pSu9-MBP map, related to Tom20 and the precursor protein (Figure 4.33). We designed a mask covering only the space above both Tom40 pores, excluding the micelle and any other TOM core features out. This mask enabled local 3D classification of cytoplasmic components without alignment, meaning that all particles remained aligned to the core domain. In this step, mask design was a key for successful separation of particles based only on the feature of interest. Figure 4.48 shows the mask on the refined map at low-density threshold and five classes resulting from classification, superimposed on the refined map at high density threshold.

After separating the particles, we refined each class with a global mask, covering the whole complex, including micelle. Figure 4.49 shows five maps at low and high-density threshold. The most abundant class (purple), containing over 200,000 particles, showed no trace of a cytoplasmic density. Two classes (red and pink), with close to 300,000 particles together, displayed a very well defined helical feature in different positions;

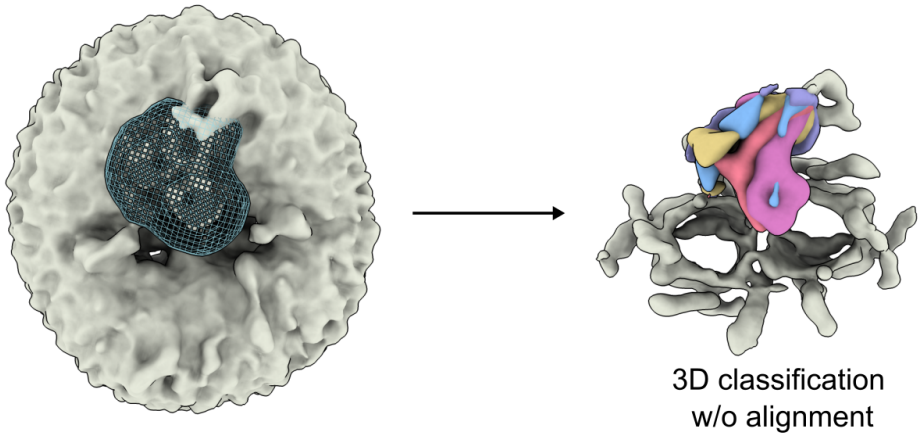

Figure 4.48: **Masked classification of TOM holo.** Masked classification without alignment of refined and polished data allowed us to separate the particles into five classes, based on their cytoplasmic features.

both with a globular element hovering over the translocation pores. These features are also visible at high-density threshold and are clearly bound to the rest of the complex (Figure 4.49B). The remaining classes (blue and yellow), with close to 150,000 particles together, contained undefined densities in the cytoplasm that mostly disappear at high-density threshold.

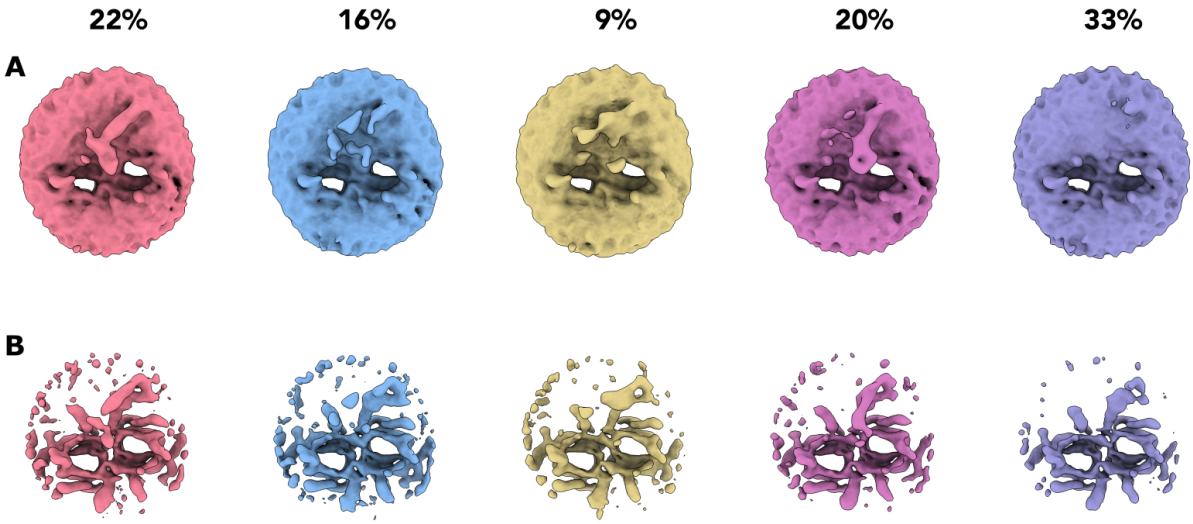

Figure 4.49: **Refinement of TOM holo 3D classes.** Global refinement of particles classified with a local mask, without alignment. The red and pink classes clearly show a defined feature in the cytoplasmic side of the complex. Maps are shown at A) low-density and B) high-density threshold.

The blue and yellow maps in Figure 4.49 lost all interesting cytoplasmic features after post-processing, but the red and pink maps improved in quality and provided a better insight on Tom20 location. Figure 4.50 shows a rod clearly present in both maps, extending from the edge of the micelle towards the translocation pores. In the maps, both rods share the same origin, but bifurcate as they approach the center of the complex. They stem from the same position as the density observed in TOM with pSu9-MBP maps (Figure 4.31). The position of both helices also matches the two elongated densities present in early low-resolution maps of this data set (Figure 4.37), suggesting that those maps contained the superposition separate states. At high-density threshold, the helices emerge from near the short N-terminal helix of Tom22, observed in Figure 4.38, although their globular domains seem to disappear at this threshold.

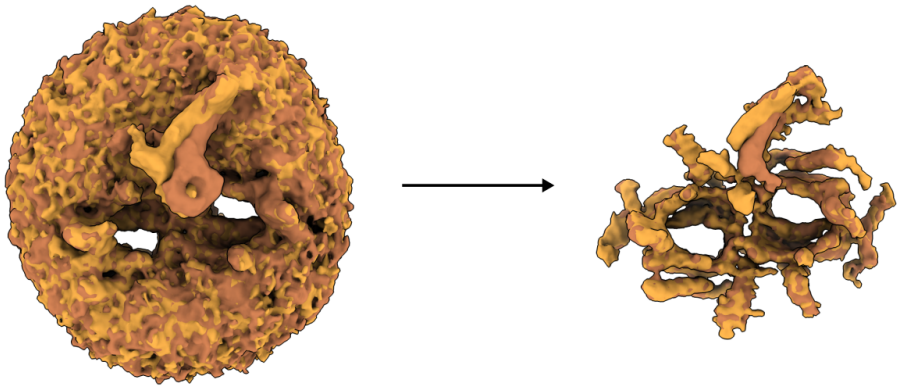

Figure 4.50: **Postprocessed TOM holo maps.** Comparison of two TOM holo classes after classification and refinement. Cytoplasmic helical features are visible in both low-density (left) and high-density (right) threshold.

Our LILBID and crosslinking experiments demonstrated that Tom22 and Tom20 interact with each other and form subcomplexes with other subunits (Figures 4.10 and 4.23). Our AlphaFold prediction models show that Tom20 docks at and is stabilized by the N-terminal of Tom22 (Figure 4.17 and Figure 4.18), closely matching the interaction in our two new maps. Based on these observations, we have assigned the cytoplasmic densities as Tom20, present in two distinct conformations.

Conformation 1 (C_1), which reached 6.7 Å resolution, reveals one copy of Tom20 hovering above the pores, roughly positioned at the same angle as Tom22. In conformation 2 (C_2), with 6.6 Å resolution, Tom20 is tilted towards one pore, in direction of Tom6. Figure 4.51 shows representative 2D averages of the particles contained in each data set, ordered in decreasing particle number. In both cases, the side view shows Tom20 protruding from the cytoplasmic side of the micelle in the 2D class average. Interestingly, in class 4 of C_1, two copies of Tom40 are clearly seen, but a third hole is visible in the larger side of the mi- celle. This is possible marked by the gap between Tom20 and Tom22 transmembrane helices.

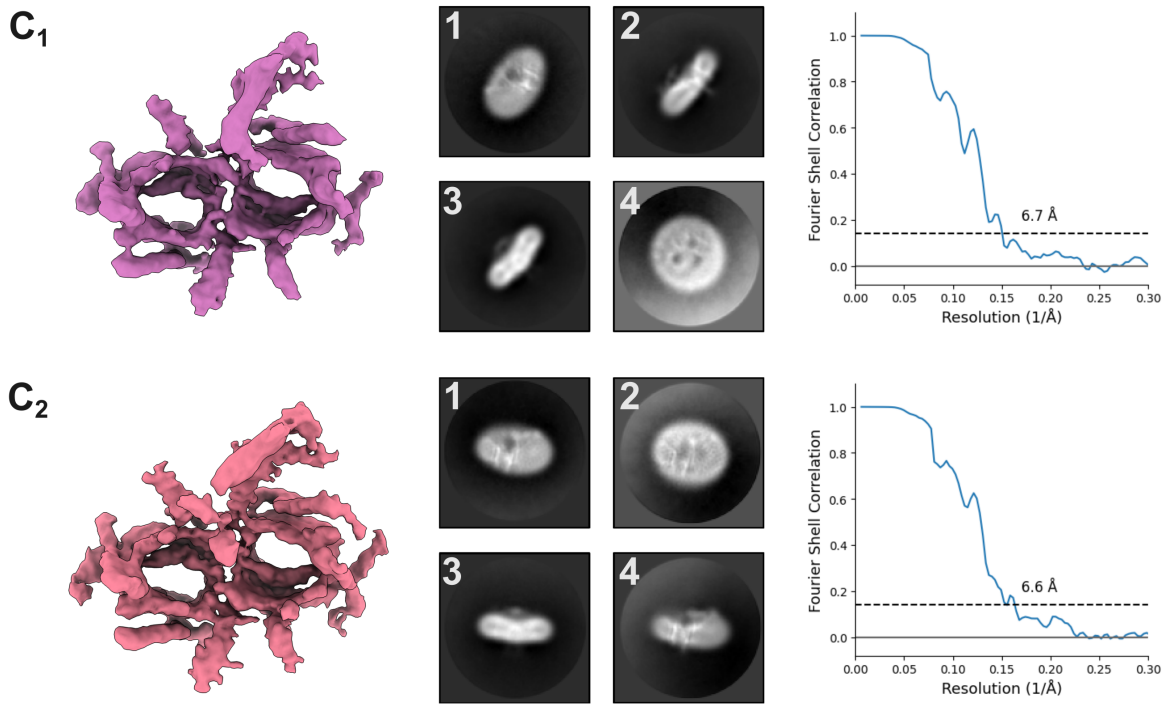

Figure 4.51: **Two conformations of Tom20.** Tom20 is present in our data set in two conformations. C_1, with 120,000 particles, reached 6.7 Å resolution. C_2, with 140,000 particles, reached 6.6 Å resolution. For each conformation, selected 2D averages and Fourier shell corralation (FSC=0.143) are shown.

After fitting our TOM core complex model into the maps, based on AlphaFold's prediction of the $Tom20_1Tom22_1Tom40_1$ subcomplex (Figure 4.17), we rigid-body-fitted Tom20. This confirmed that AlphaFold's prediction of Tom20 (Figure 4.16) matches our Tom20 densities, both in size and shape. In our model, Tom20 is docked to the complex at a single point of the the cytoplasm-membrane interface of Tom22 (Figure

4.52). It is also clear now that in C_1, Tom20 takes a central position, close to both pores, but that in C_2, it approaches the longest loop in Tom40, in close proximity to Tom6.

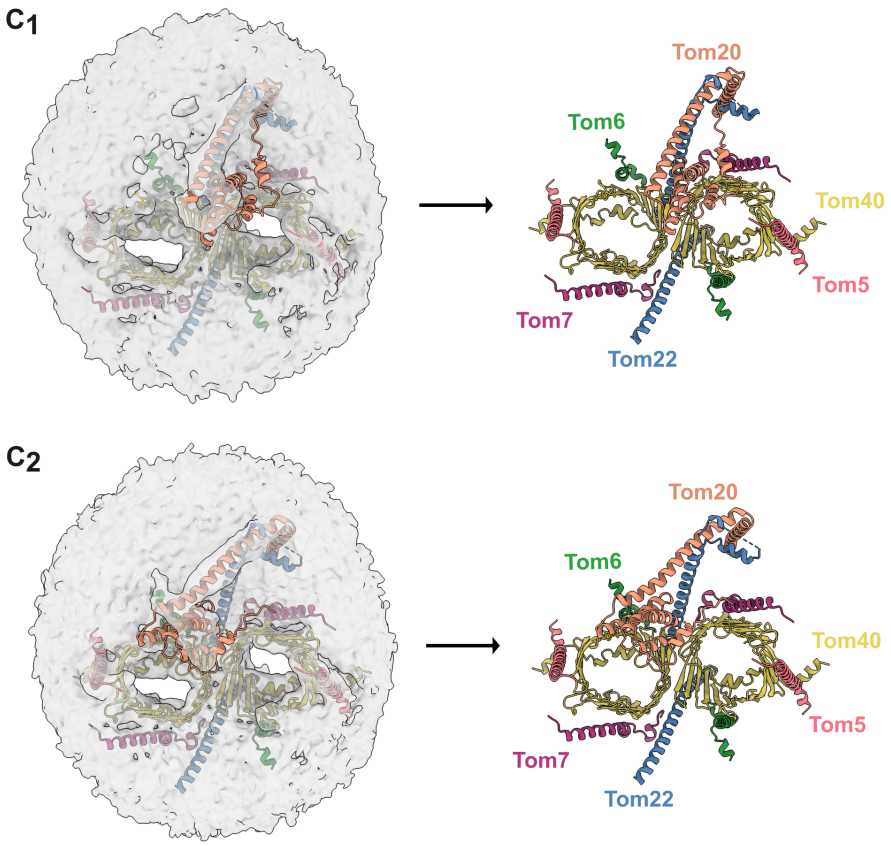

Figure 4.52: **Maps and models of TOM holo complex.** Rigid-body-fit of AlphaFold's Tom20 prediction in the cryoEM map of TOM holo (conformations C_1 and C_2) with our TOM core complex model.

Using the ISOLDE software, we relaxed the rigid-body-fitted Tom20 into both maps. However, it was challenging to precisely fit its receptor domain (Figure 4.16B) into the globular density at the end of the helices. Although we only approximately aligned them, the exact position of the individual helices remains elusive. Particularly, the C-terminal helix of Tom20 was difficult to fit and remains mostly outside of the density. In the AlphaFold prediction of Tom20, this C-terminal helix had a low confidence value, suggesting high mobility (Figure 4.16A).

It is interesting to note that both maps contain a small bulge protruding from the docking site, on top of Tom20. As shown in Figure 4.17, the AlphaFold prediction of $Tom20_1Tom22_1Tom40_1$ revealed that the disordered N-terminus of Tom22 can wrap around the longest cytoplasmic helix of Tom20. We therefore used this prediction model to fit a short part of cytoplasmic Tom22 as it wraps around Tom20. Figure 4.53 shows the docking site, as well as the distance at which the Tom20 receptor domain hovers above the translocation pores. This distance does not vary between both conformations, suggesting that Tom20 movement is only parallel to the membrane plane.

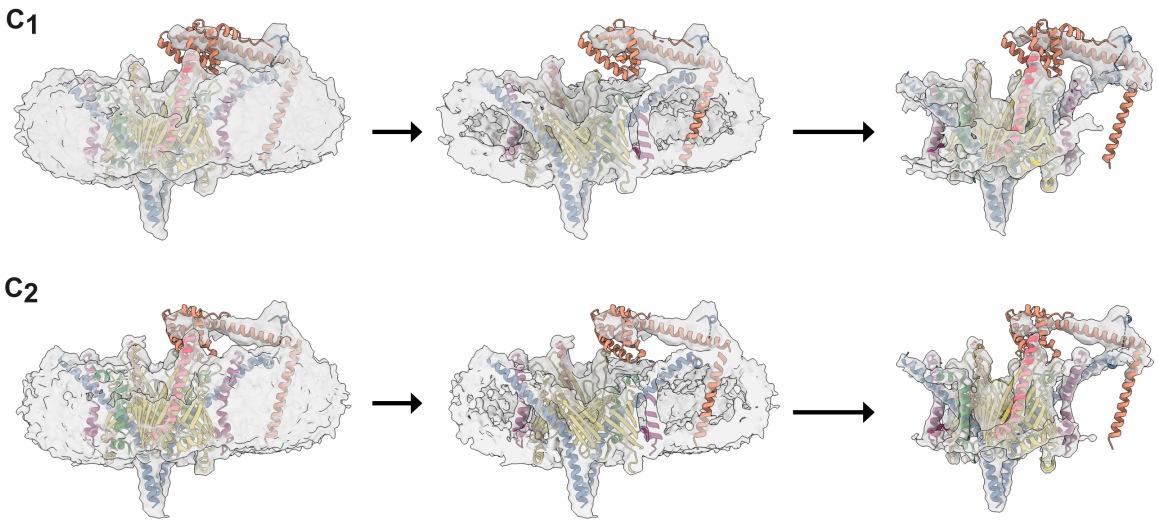

Figure 4.53: **Tom20 fit in TOM holo map.** Side view of TOM holo map-model superpositions, showing Tom20's helices in the Tom20 density. Both center maps depict a transversally cut complex, which highlights the intermembrane helix of Tom20.

At lower threshold levels, we see a small rod-density originating at the membrane-cytoplasm interface of both maps, spanning halfway through the micelle (Figure 4.53). In C_2, this rod is still partially visible at higher-threshold level. We used this to fit the N-terminal transmembrane helix of Tom20. This further indicates that Tom20 only binds to the rest of the complex at a single point, and that its N-terminal reaches the IMS without interacting with any other subunit.

Larger particles

After initial classification of TOM holo with pALDH, we discarded nearly 60% of the particles, due to their lack of higher-resolution features of the TOM complex (Figure 4.36). However, some of these remaining particles may belong to tetrameric TOM or other large complexes. We continued to process the three most promising classes and discarded the rest. As these particles were overall larger, we re-extracted the remaining 660,000 particles in a bigger box, but downsampled them 1.5 times. Initially, we separated the particles by 2D classification into 80 classes. Figure 4.54 shows selected averages depicting particles with various orientations.

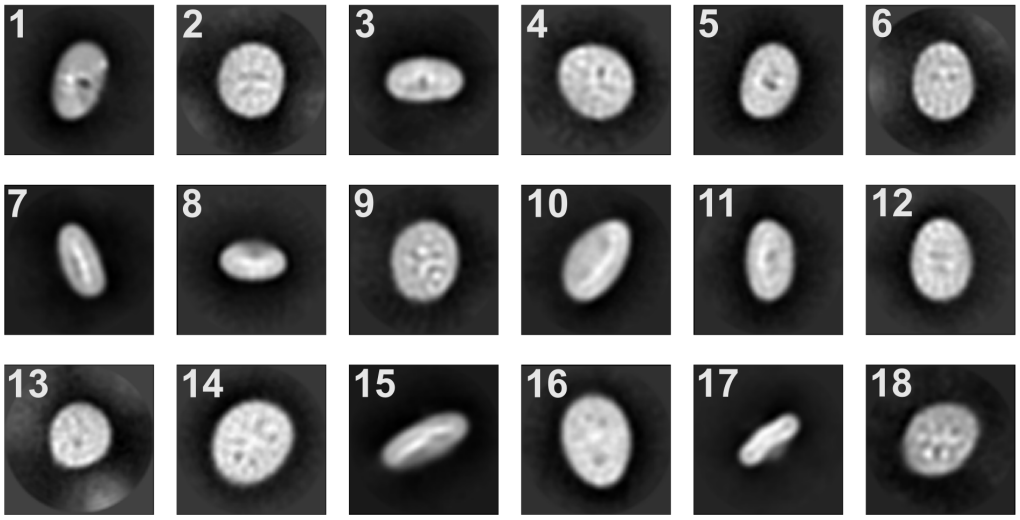

Figure 4.54: **2D averages of larger particles.** Selected averages from 2D classification of 660,000 particles depicting a larger version of TOM complex. The averages are ordered according to their decreasing particle number.

Once more, the classes have different shapes and sizes due to their orientation, but also hint at heterogeneity in the data set. Side views, such as in classes 7, 8 and 17, match size and aspect of the TOM dimer we resolved previously, but could also correspond to larger particles from a tilted view instead. Top views are also difficult to interpret, as some seem almost perfectly round, like class 2, or elongated, like class 14. Some apparent top views indicate two clear pores, like class 6, but others hint at more pores. Classes 9 and 18, for example, show complex features, making them difficult to interpret. As manual selection of the averages was complex, we continued

separating the data set through several rounds of 3D classification. The classification seemed to classify the particles by the size of the micelle (Figure 4.55). Most of the resulting maps lacked any internal structure and became very noisy at low density threshold levels. Two classes (yellow and orange, center), however, had two clear pores at their centers and structured cytoplasmic densities. Although they resemble TOM dimers, they have a larger micelle compared to the dimer structure resolved earlier.

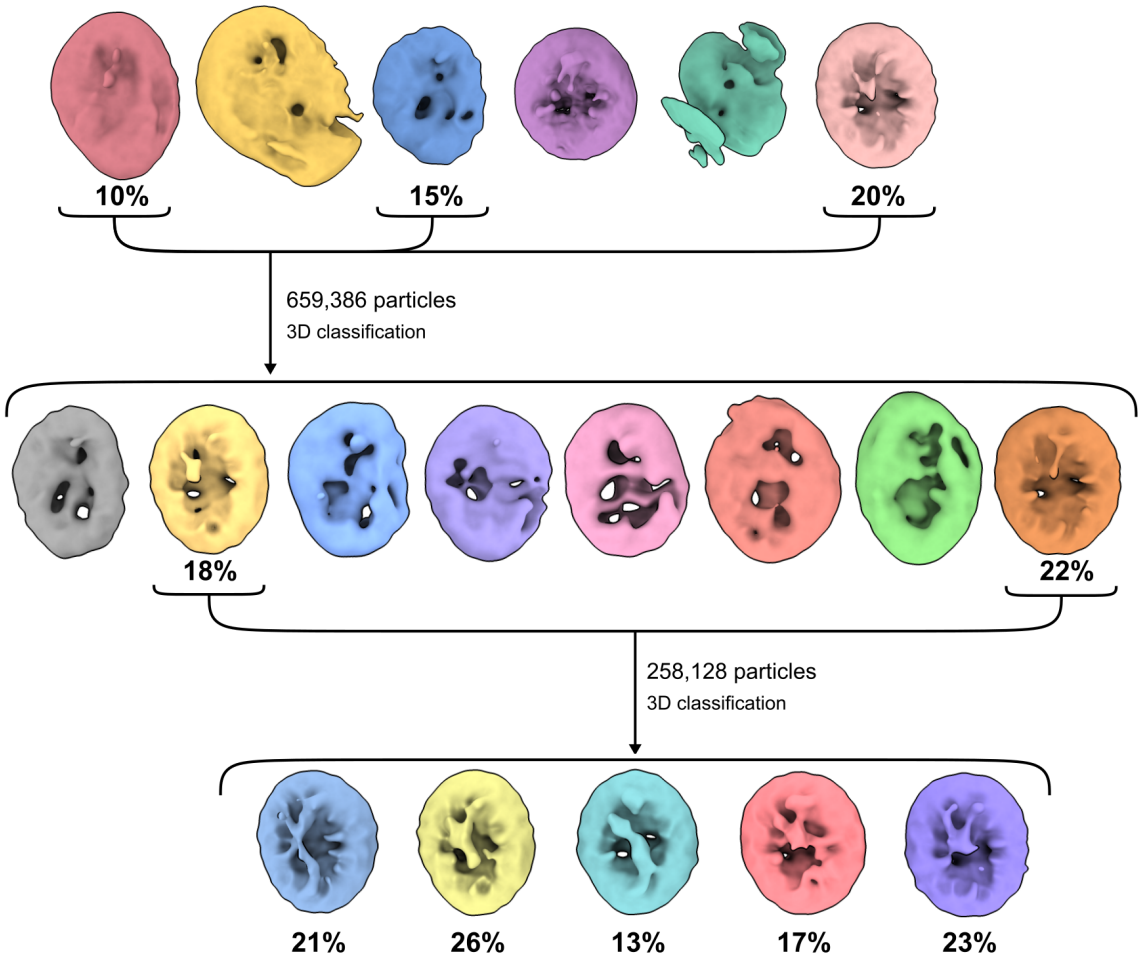

Figure 4.55: **Classification of larger particles.** 3D classification of TOM holo, focusing on the larger 660,000 particles in the data set. The final five classes seem to contain dimeric TOM in a large micelle, with an attached cytoplasmic feature in different positions.

Together, the two best classes contained 260,000 particles, which were classified into five classes in 3D. The resulting classes had similar shapes, but differed in their cytoplasmic features (Figure 4.55, bottom). These features remind us of Tom20 densities seen before, but appear in many different positions, in both sides of the pores. It is possible

that these particles contain one or more copies of Tom20 in different conformations from the two we have resolved, but that they are less stable. We continued to process these particles, but despite our attempts to further classify, align or refine them, we were not able to reconstruct any map beyond 16 Å resolution, nor to obtain any better insight of the features of interest.

Finally, we tried to reconstruct the tetrameric complex, by exclusively selecting the largest particles after 2D and 3D classification. However, despite these attempts, we were unable to produce a meaningful reconstruction. 3D reconstructions showed elongated pancake-structures, completely lacking any discernible internal features. As seen in the 2D class average, these large particles seem to have a lot of blurred information on their cytopalsmic side, making them difficult to properly align and classify.

Chapter 5

Discussion

In this project, we used cryoEM, as well as mass-spectrometry, differential scanning fluorimetry and computational prediction models to study structure and translocation mechanism of TOM holo complex. Our biochemical analysis confirms the successful purification of both, TOM core and TOM holo complexes, with all their corresponding subunits (Figures 4.1 and 4.3).

5.1 Biochemical analysis

Isolating TOM complex with Tom20 and Tom70 was a first challenge in this project. Thanks to the persistence of Dr. Stephan Nussberger and Dr. Thomas Bausewein, based on the work of Künkele et al., 1998, we were able to find the key: separating the membranes of the mitochondria, and working only with outer membranes. Producing OMVs required more rigorous work, more so than the already laborious task of isolating mitochondria from *Neurospora crassa*. However, this additional step allowed us to prepare intact and stable TOM holo complex. It has been reported that Tom20 and Tom70 have a low-affinity interaction with TOM core complex (*116*), however, our OMV isolated complex remained stable in GDN throughout our experiments. It is possible that Tom20 and Tom70 interact with other proteins in the IMS or inner

membrane, causing these subunits dissociate from the TOM core, when isolated from whole mitochondria (*18*).

Other examples in the literature show recent attempts to purify TOM holo complex for structural studies. In yeast for example, Tucker et al., (2019) attempted to overexpress all TOM core subunits, plus Tom20, but found no trace of Tom20 associating with the complex after purification (*25*). In contrast, Guan et al., (2021) reported bands corresponding to Tom20 and Tom70 after expression and purification of TOM in human cells, but saw no dejsity in their cryoEM maps correspondig to those two sununits (*28*). More recently, Su et al., (2022) successfully expressed, purified and crosslinked TOM holo (*43*). Using cryoEM they demonstrated that Tom20 is positioned close to the Tom40 pore. However, Tom70 was absent from their map. This reinforces the notion that these subunits are highly dynamic and bind transiently to the complex, and aremtherefore challenging targets for structural determination.

Remarkably, using nanoDSF, we were able to demonstrated the stability of the holo complex. Figure 4.6 illustrates how much more stable TOM holo complex is compared to TOM core, indicated by its consistently higher average melting temperature under various buffer conditions. It is possible that the presence of Tom20 and Tom70 allowed more lipids to remain attached to the complex after purification, making the TOM holo complex more stable. In addition, there is no indication of early dissociation of subunits prior to melting temperature. We were able, however, to separate the complex into its individual components after irradiating it with a high-intensity laser beam in our LILBID experiments (Figure 4.9). Once more, a comparison with the TOM core complex allowed us learn more about the holo complex. Contrasting both spectra confirmed the presence of Tom20 in the holo sample: two peaks corresponding to singly (20,099 m/z) and doubly (10,185 m/z) charged Tom20. At lower laser intensity we identified larger structures (Figure 4.10), including a peak corresponding to a singly charged Tom70 (69,363 m/z). This served as further affirmation that we had succesfully purified the TOM holo complex with all its subunits.

Our LILBID experiments with Dr. Nina Morgner showed different subcomplexes that can be formed by TOM subunits. We saw Tom20 and Tom70 interacting with core subunits, and shed a light into their location within the holo complex. In Figure 4.10, we identified three peaks corresponding to a subcomplex formed by Tom20, Tom22 and Tom40 ($Tom20_1Tom22_1Tom40_1$, $Tom20_1Tom22_1Tom40_1sT_1$, $Tom20_1Tom22_1Tom40_1sT_2$) and the smaller TOM subunits. This is great evidence of interaction of Tom20, Tom22 and Tom40, and the possible conformation of a protomer, formed by one copy of Tom20 per pore. We identified two peaks assigned to subcomplexes with two copies of Tom20, revealing interaction of Tom20 with itself. Four peaks corresponding to Tom70 subcomplexes revealed its interaction with the small TOMs, Tom20, Tom22 and Tom40 ($Tom70_1sT_2$, $Tom22_1Tom70_1$, $Tom20_1Tom70_1$, $Tom40_1Tom70_1$). This indicates a direct interaction of Tom70 with the translocation pore, and with other receptor subunits, but makes it difficult to pinpoint its exact location within the complex. It is possible, nonetheless, that its transmembrane helix is located close to the interaction site of Tom22 and Tom20, or that its cytoplasmic domain interacts with both Tom40 and Tom22.

Our XL-MS experiments, in collaboration with Dr. Julian Langer, provided additional insight into cooperation between subunits and new information on precursor protein binding (Figure 4.23). Both samples showed links between Tom20 and Tom22, as well as between Tom20 and Tom40, reinforcing the existence of $Tom20_1Tom22_1Tom40_1$ subcomplex, and provided evidence of interaction between Tom20 and Tom7. The sample incubated in DSSO showed interaction between the pSu9-MBP construct and Tom20, but not with the translocation pore. EDC crosslinks were more abundant and proved interaction of Tom20 C-terminal receptor domain and Tom6. Even more interestingly, we saw multiple links between Tom20 and Tom70, mainly in their cytoplasmic domains, suggesting close contacts between both subunits. Multiple links between Tom70 and Tom40, and Tom70 and the small TOMs reinforce our LILBID results. These connections are consistent with other published results (*41*, *51*). Site-specific crosslinking experiments present multiple interactions in yeast between Tom40

and Tom20, as well as interactions between the cytoplasmic domains of Tom20 and Tom22. Likewise, pSu9 has been successfully crosslinked to the inside of Tom40 and to cytoplasmic domains of Tom20 and Tom22.

5.2 Structural analysis

We analysed two presequence bound TOM holo samples by cryoEM. The first sample, incubated with pSu9-MBP, resulted in two maps of the complex at intermediate resolution. At 5 Å resolution, one map allowed us to identify five characteristic subunits of TOM core complex, with an elongated copy of Tom22 (Figure 4.30). This map, however, did not reach sufficient resolution for modelling of side chains nor showed individual β-strands in Tom40. More interestingly, we obtained a 10 Å resolution map of pSu9-MBP bound TOM (Figure 4.31). This map clearly displayed an elongated feature sticking out of the micelle, on the cytoplasmic side of TOM, crossing the structure from one extreme to another. At lower density threshold we identified MBP attached to the elongated density in the map.

These new features, in comparison to other published TOM core structures (*23*, *25*, *27*, *35*), highlight the presence of TOM holo subunits and presequence in the particles. The elongated feature was too thin to fit the AlphaFold prediction model of Tom70. However, it roughly matched Tom20 in size (Figure 5.1). At this point it was still unclear whether the new density corresponded to Tom20, and in that case, how many Tom20 copies interact with the core complex.

From this sample we reconstructed two maps, the first one of to the core complex, and the second the holo complex (Figure 4.29), even though the peak was monodisperse during purification. It is thus possible that Tom20 and Tom70 dissociate from the complex at a later stage. Our nanoDSF and MS experiments demonstrated stability of the complex and interaction of Tom20 and Tom70 with other subunits. Another possibility is that the classification was mainly driven by strong features in the core complex, as opposed to weaker additional densities present in the cytoplasmic domain.

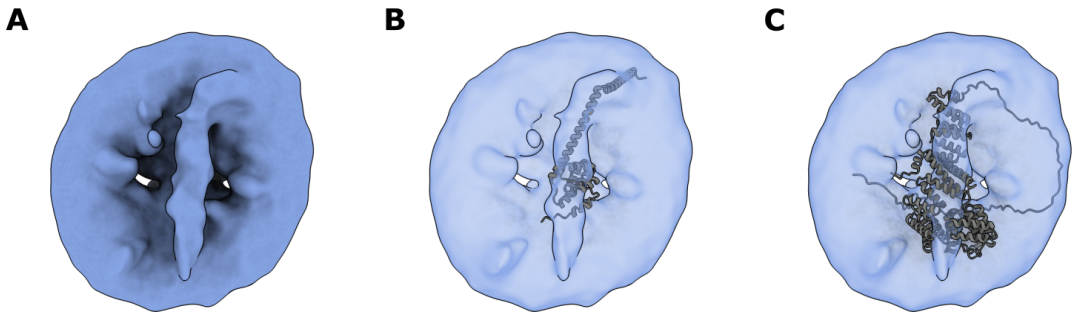

Figure 5.1: **Fitting in TOM + pSu9-MBP map.** A) Low-resolution map of pSu9-MBP bound TOM holo complex, showing an elongated density protruding from its micelle. Transparent overlay of the same map at and AlphaFold prediction models of B) Tom20 and C) Tom70 fitted into the cytoplasmic density.

Furthermore, as the isolated complex cannot complete translocation of precursors on its own because of the absence of membrane potential, pSu9-MBP may bind to the complex at different locations, making the particles too noisy for successful alignment. Unfortunately, neither further processing nor increasing the size of the data set led us to a higher resolution.

We then studied TOM holo complex with the shorter pALDH precursor construct, hoping to reduce noise in the sample. Similar to the pSu9-MBP data set, some individually extracted particles in the pALDH data set had a small, round domain attached to the cytoplasmic side of the complex (Figures 4.27 and 4.34). pALDH does not contain any large, foldable domains, and Tom20 could be too small to provide reasonable contrast at this magnification. It is then likely that the round domain pertains to Tom70 instead. We noticed that in both samples, this round Tom70 domain appears at a distance of the complex, in different positions, indicating a flexible link. This might suggest the presence of Tom70 attached to TOM in the collected data set. We processed this data set in two ways: one focused on high-resolution information of the TOM core complex using cryoSPARC, and one focused on lower resolution information of TOM holo complex using Relion (Figure 4.39). In the end, we generated four maps of TOM complex, which allowed us to model the core complex, rigid-body-fit

Tom20 in the holo complex, and visualize the translocation of pALDH. In comparison to the pSu9-MBP sample, pALDH helped stabilize the complex and reduce noise during processing, leading to higher-resolution maps.

5.2.1 TOM core complex

Processing in cryoSPARC, we arrived to a 3.3 Å resolution map of TOM core complex (Figure 4.41). Based on the AlphaFold prediction (Figure 4.11), we were able to model the first high-resolution structure of *N. crassa* TOM core complex dimer, with its five characteristic subunits. Figure 5.2 shows the symmetric TOM map and model, as seen from the cytoplasm. The original AlphaFold prediction shows soluble ends of transmembrane helical subunits with a low-confidence score. We were not able to model those regions in our map, as they lack secondary structure and are too flexible to be seen by cryoEM. However, we successfully modelled all loops between the Tom40 strands, and Tom40's IMS α3 helix.

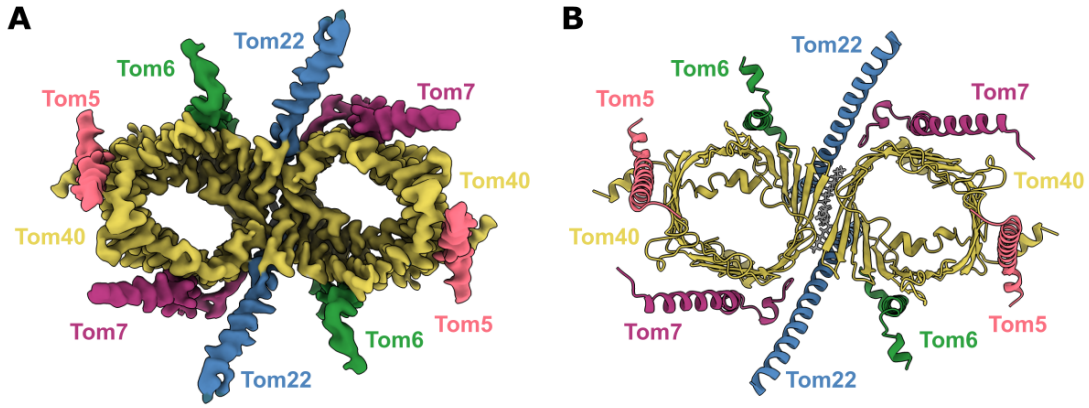

Figure 5.2: **Structure of TOM core complex.** A) Map and B) model of our TOM core complex structure seen from the cytoplasmic side. Adapted from Ornelas et al., 2023.

We also modelled four detergent molecules surrounding each copy of Tom22 (Figure 4.43). In nature, these densities are probably occupied by lipids. Between the translocation pores, we identified a phospholipid molecule, stabilizing the previously reported 20° tilt of Tom40 with respect to the symmetry axis (*23*). This lipid interacts with both copies of Tom22, and is held in place by F309 in both Tom40s. The tilt of the pores, plus the elongated transmembrane helices around them, create a funnel-like

shape at the membrane-cytoplasm interface. Figure 5.3 illustrates the local curvature this might induce on the membrane.

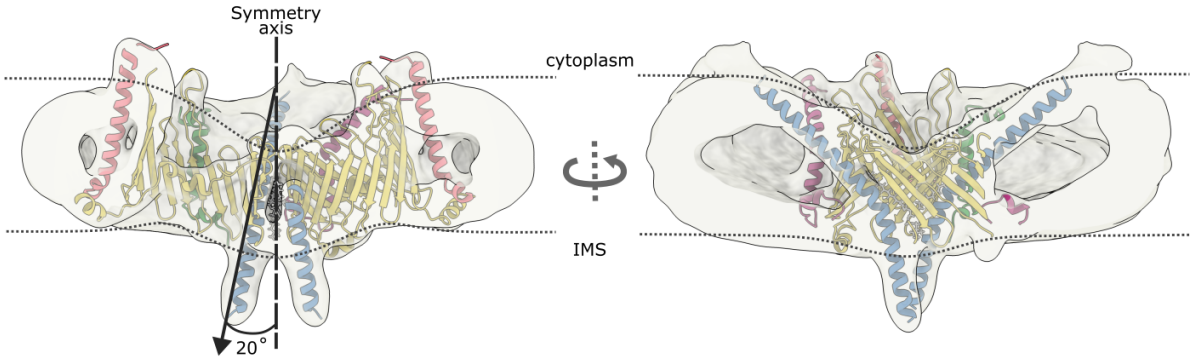

Figure 5.3: **TOM in the membrane.** Transversal cuts of TOM core model overlaid on a low-pass-filtered map, showing a 20° angle tilt between the pores. A dotted line highlights the funnel-like curvature of around the pores, and represents the outer mitochondrial membrane.

We compared our TOM core model with published models of TOM from other organisms (Figure 5.4A). The 19-strand translocation pore is highly conserved among *N. crassa*, yeast and human TOM, together with its inner helix. However, we found differences in α1 and α3. The N-terminal of human Tom40 is predicted to be long and disordered, it has not been observed in any published cryoEM structure, but it is present in the *N. crassa* and yeast models as α1. Similarly, α3 is missing in human Tom40, but differs between *N. crassa* and yeast Tom40. In yeast, it is longer than in *N. crassa* TOM, points towards Tom22 and is only partially visible in the deposited structures. Interestingly, as shown in Figure 5.4C, human Tom7 has an elongated C-terminal that reaches down towards the IMS and an elongated Tom40 loop that likely replaced α3 in the translocation exit pathway.

Furthermore, we see that in all three organisms Tom22 bends at a different angle in the IMS. Human Tom22 is tilted in the opposite direction than *N. crassa* Tom22 in the IMS, with close to a 25° separation between them respect to the symmetry axis (Figure 5.4D). Lastly, the models differ in the longest Tom40 loop, next to Tom6 (Figure 5.4B). In yeast, this loop is longer than in *N. crassa* Tom40, which makes it more flexible, explaining why it has not been modelled in any cryoEM map. On the

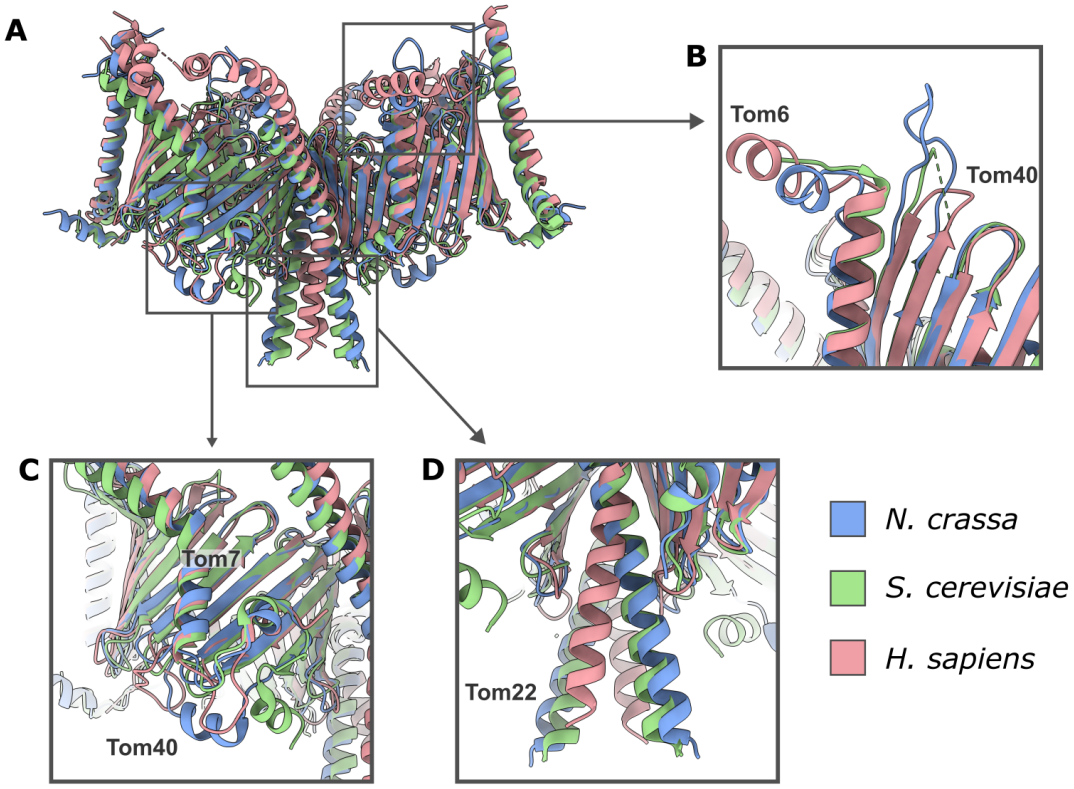

Figure 5.4: **TOM core complex comparison.** Superposition of TOM core complex models from *N. crassa* (PDB 8B41), yeast (PDB 6UCU) and human (PDB 7CP9) mitochondria. Cutout squares highlight important differences between the models, mainly in Tom7, Tom22 and Tom40.

contrary, it is very short in human Tom40. This loop has been associated to Tom20 in yeast (*41*), therefore its presence is important for translocation. It is unknown if human TOM has an alternative to the loop.

5.2.2 Precursor protein translocation

By further processing our TOM core map, we arrived at a symmetric 4 Å resolution map of the pALDH-bound complex. Figure 4.46 shows rod-like densities crossing the translocation pores, perpendicular to the membrane. Inside each pore, these rods span Tom40 from α2 to α3, making contact with the inner wall, close to F349 (Figure 5.5). This phenylalanine is of special interest because it is the last residue in Tom40 and it points inwards facing the center of the pore, indicating possible involvement in translocation. Interestingly, right behind F349 in Tom40's β4, a second phenylalanine, F106, sticks out towards the center of the pore.

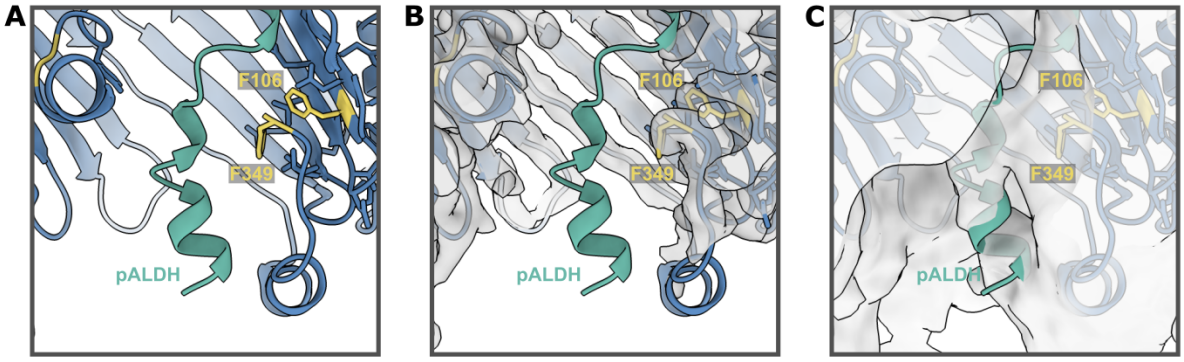

Figure 5.5: **Precursor protein translocation.** Interaction of precursor pALDH (teal) with hydrophobic residues inside Tom40 (yellow). A) Model of TOM core complex with rigid-body-fit of the prediction model of pALDH. B) Overlay with our high-resolution map of TOM core complex. C) Overlay with 4 Å resolution map showing a density corresponding to pALDH as it crosses the pore.

As previously discussed, yeast Tom40 α3 points outwards from the pore, yet it contains a phenylalanine residue in β4, at the same position as *N. crassa*'s F106. Human Tom40 lacks both α3 and a phenylalanine in β4, but has a valine residue in its place (Figure 5.6). This is remarkable as it may point to a hydrophobic interaction site within Tom40. The presence of a second hydrophobic residue at that position in *N. crassa* Tom40 might hint at a gating mechanism within the translocation pore; although its absence in other organisms is puzzling.

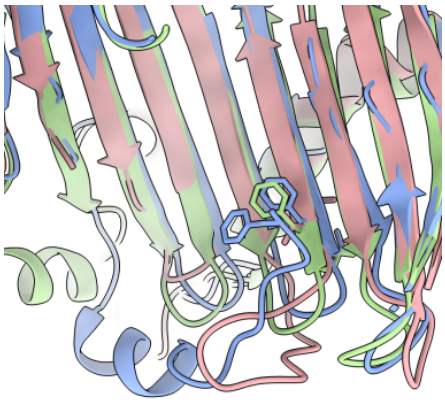

Figure 5.6: **Hydrophobic residues in Tom40.** Differences in certain hydrophobic residues between *N. crassa*, in blue (PDB 8B41), yeast, in green (PDB 6UCU), and human, in pink (PDB 7CP9) Tom40 provide an insight into a possible protein translocation mechanism.

Most precursor proteins contain a positively charged, hydrophobic, helical presequence that guides them into mitochondria (*10*, *35*). However, the outer mitochondrial membrane is porous and lacks a potential ($\Delta\psi$), opening the question of what drives translocation through Tom40 (*117*). An analysis of the inner surface of Tom40 sheds light into possible translocation pathways. Figure 5.7 shows the electrostatic potential surface of Tom40, containing large negative patches all around its inner wall (red), as it contains many acidic residues. Likewise, we can see clear hydrophobic patches (yellow) inside the pore, opposite to α2, towards the cytoplasmic entrance to the pore and close to α3. Combination of acidic and hydrophobic patches is likely to be fundamental for guiding the presequence helices through the pore, towards the IMS. Figure 5.7 also shows our superimposed pALDH bound TOM map, where we can see the pALDH density close to a large acidic patch and hydrophobic residues in the non-helix wall of Tom40.

It remains unknown whether the overall negative charge of the pore is enough to fully translocate presequences to the IMS. Our cryoEM structures suggest that pALDH has low occupancy in the TOM complex dimer and that it is bound to a relatively small number of TOM particles. It may bind to different regions of Tom40 through the translocation process, since a positive presequence can interact with various negatively charged regions of the pore and receptors. As we have not chemically restricted pALDH to the complex, by crosslinking it to Tom40 for example, presequence interaction may be transient. It is also possible that the isolated complex is not able to complete translocation without help of chaperones in the IMS or potential driven TIM23, as it would in mitochondria.

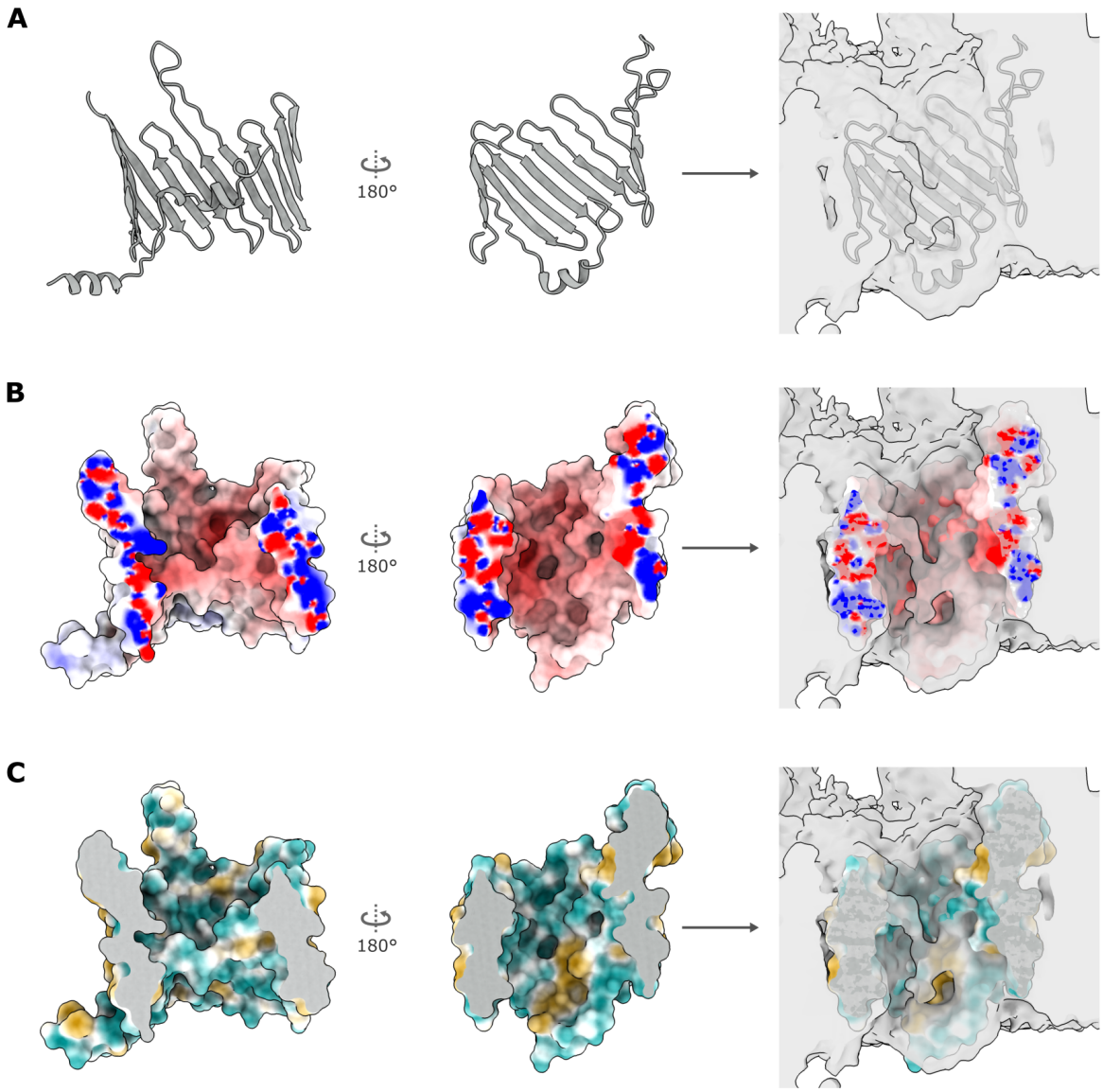

Figure 5.7: **Tom40 translocation channel.** Transversal cut of our Tom40 model, seen from its helix and non-helix side, plus an overlay with our presequence-bound 4 Å map. A) Ribbon representation. B) Surface representation of the electrostatic potential, with negative charges colored in red and positive in blue. Calculated using the APBS-PDB2PQR software suite. C) Hydrophobicity surface representation, with hydrophobic residues in yellow and polar residues in teal.

Our crosslinking results with the long crosslinker, DSSO, demonstrated binding between pSu9 and Tom40. This was not the case with zero-length crosslinker, EDC (Figure 4.23). In the literature, site-specific crosslinking experiments report binding of pSu9 to residues on the acidic patch of the non-helix side of yeast Tom40 (*41*). Although our structure does not show a direct interaction of pALDH with residues in that region, we see the presequence density stand within close range. Shiota et al., 2015 presented crosslinks of a presequence-less precursor, at various points in the cytoplasmic side of yeast Tom40, but our structure does not offer insight into binding at the cytoplasmic side. Lastly, by fitting the AlphaFold multimer prediction of Tom40 and pALDH (Figure 4.26) into our 4 Å TOM core map, we were able to compare possible binding sites of the presequence. The predicted position of pALDH in several models passes through the pore in a similar position to the one observed in our map, coming close to $\alpha 2$ and $\alpha 3$. Although we can also see more copies of the predicted pALDH presequence interacting with the cytoplasmic entrance of the pore.

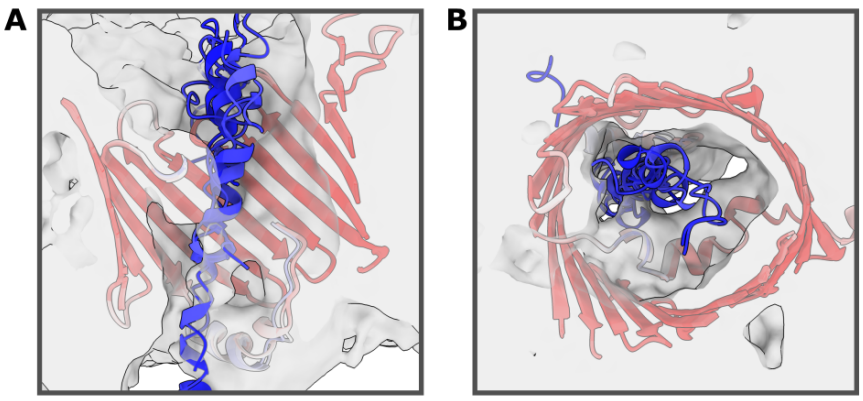

Figure 5.8: **Translocation prediction vs. map.** Superposition of our TOM core complex map with the AlphaFold multimer prediction models of pALDH bound Tom40, colored according to per-residue confidence rate. Transversal cut of Tom40 as seen from the A) membrane plane and B) cytoplasm.

5.2.3 TOM holo complex

The TOM holo complex has been a target of interest for structural biologists for decades. Using cryoEM we were able to determine the position of Tom20 in relation to the TOM core complex. Both our asymmetric TOM holo maps contain a copy of

Tom20 on one side of the core dimer, extending from the edge of the micelle, towards the center of the complex (Figure 4.50). Based on our MS and AlphaFold results, we were able confirm the identity of the Tom20 density in our map, providing us with new structural insights into the TOM holo complex.

Our maps show an extension of the transmembrane helix of Tom22, as it bends to one side, forming a short helix on the membrane-cytoplasm interface. Our 6.6 Å and 6.7 Å TOM holo maps suggest that Tom20 is attached to the core complex uniquely at this cytosol-membrane interface, and stabilized in its position by the unstructured N-terminal of Tom22 (Figure 4.53). Multiple studies suggest that beyond presequence recognition, Tom22 serves as an stabilizer of Tom20, while optimizing its receptor functions and increasing translocation levels (*34*, *41*, *114*). Shiota et al., 2015 used *in-situ* photocrossliking to demonstrate that Tom20 and Tom22 are close to one another in the complex, demonstrating that the N-terminal of Tom22 interacts with the soluble domain of of Tom20 in yeast. Furthermore, they saw that after incubation with a precursor protein, the links between Tom20 and Tom22 decrease, suggesting that the two subunits team up by simultaneously attaching to the incoming presequence instead. These interacting regions match our AlphaFold predictions of the $Tom20_1Tom22_1Tom40_1$ subcomplex (Figure 4.17A). Upon close examination, we found that precisely the most negatively charged region in the Tom22 N-terminal domain wraps around a very positive region in Tom20. Figure 5.9 highlights the region of charge complementarity in Tom20 and Tom22, which we suggest holds both subunits together through electrostatic interactions.

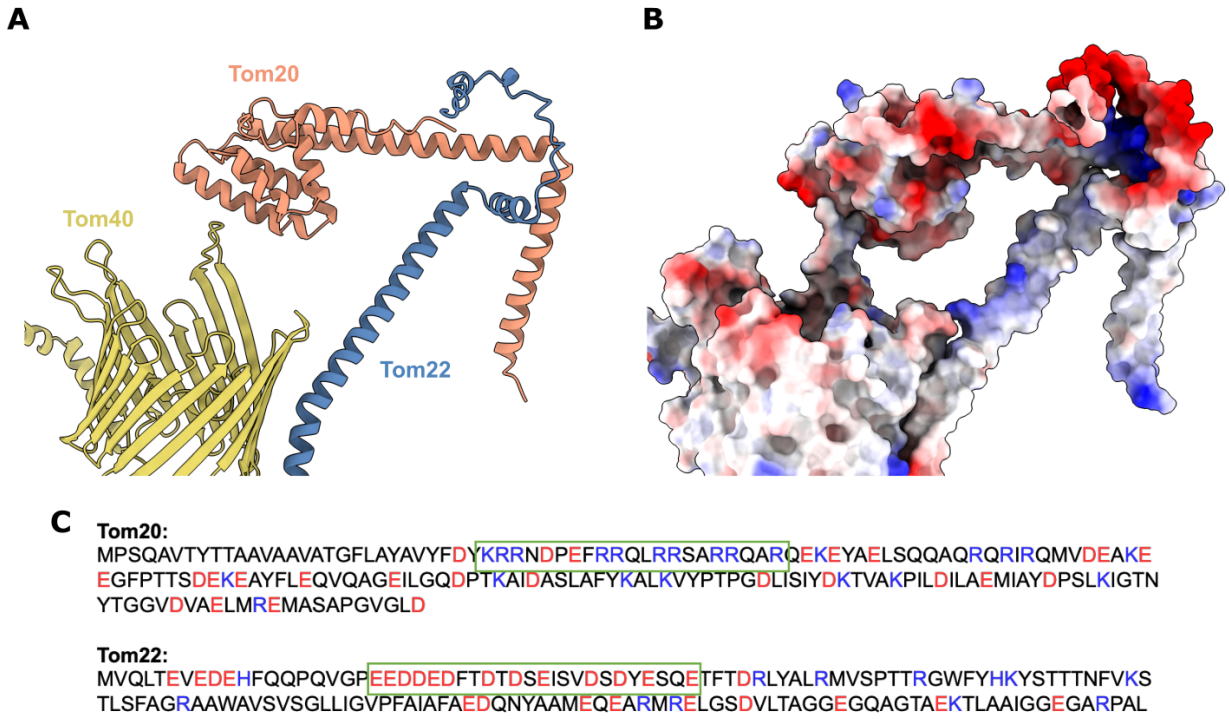

Figure 5.9: **AlphaFold docking of Tom20.** AlphaFold prediction model of $Tom20_1Tom22_1Tom40_1$ subcomplex, A) in ribbon representation, and B) with its surface colored according to the electrostatic charge of its residues. The first 15 residues of Tom22 are hidden for clarity of the region of interest. C) Sequence of Tom20 and Tom22, colored by residue charge, highlighting regions of charge complementarity. Blue indicates positive charge, and red, negative.

Both of our maps show an interesting bump protruding from Tom20, right on top of its docking site (Figure 4.53). We assigned this density to Tom22 N-terminus. Based on this assumption we generated a rigid-body-fit of our TOM core model, with an extended Tom22 on one side, as it interacts with Tom20. As shown in Figure 5.10 and in our publication, Tom20 is docked to the complex only at this site (*118*). We were able to fit a short section of Tom22 as it wraps around Tom20, based on the aforementioned $Tom20_1Tom22_1Tom40_1$ AlphaFold predictions. Figure 4.53 also shows a transversal cut of the micelle in both conformations, partially showing a transmembrane helix of Tom20 inside the micelle. The N-terminus of the subunit is not well-defined, but it does not seem to approach the complex at the IMS interface.

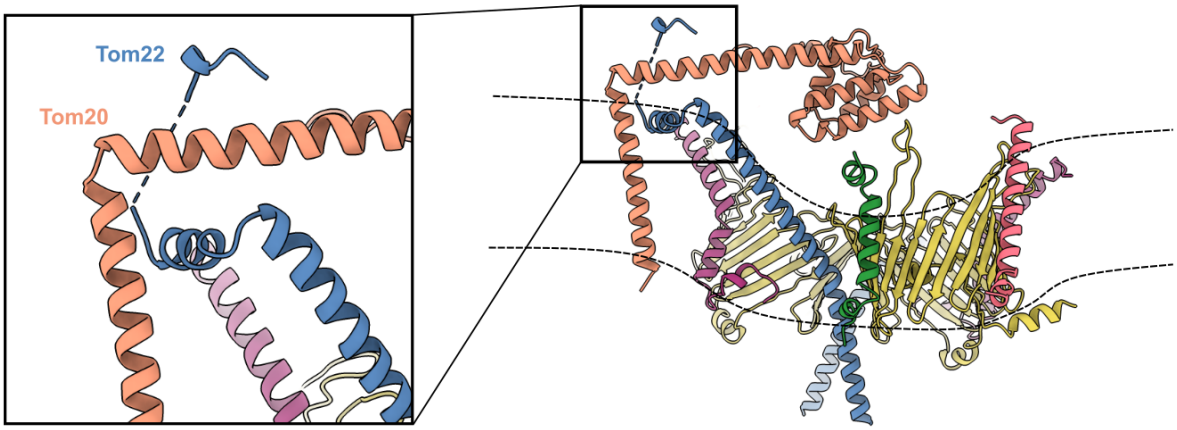

Figure 5.10: **Model of TOM holo complex.** Interaction between Tom20 and Tom22 at the docking site in our model. Tom20 is stabilized into this position by its electrostatic interactions with Tom22. The membrane plane is indicated in dashed lines. Adapted from Ornelas et al., 2023

NMR and crystalography studies have shown pALDH helix binds to the TPR domain in Tom20, suggesting that a helical structured presequence is required for this interaction (*48*, *49*). However, positive charges in the presequence are not essential for recognition by Tom20, but appear to be necessary for recognition by Tom22 (*51*, *119*). Based on this, it is considered that Tom20 and Tom22 can simultaneously recognize opposite sides of the presequence helix, with Tom20 binding to its hydrophobic region whereas Tom22 binds its positive side (*114*).

Furthermore, our maps indicate that Tom20 is present in the holo complex in at least two different conformations, and that the docking site in Tom22 serves as a pivot point for Tom20. As shown in Figure 5.11, in conformation C_1, Tom20 is positioned between the two pores. In conformation C_2, Tom20 approaches one of the pores and gets close to Tom6. This confirms the flexibility of the Tom20 receptor.

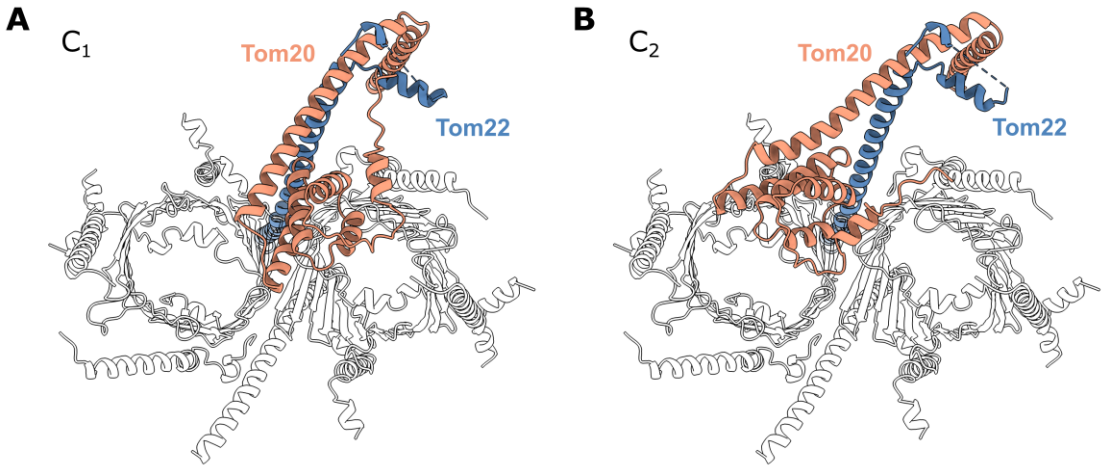

Figure 5.11: **Two conformations of Tom20**. Tom20 (orange) takes two conformations on the cytoplasmic side of TOM holo complex, as it docks to core complex at the N-terminal of Tom22 (blue). A) Conformation C_1 shows Tom20 positioned between two translocation pores. B) Conformation C_2 shows Tom20 leaning towards only one pore.

Our model of conformation C_2 is consistent with the model predicted by AlphaFold for $Tom20_1Tom22_1Tom40_1$. In both cases, we see Tom20 closely interacting with the longest loop in Tom40, between strands β14 and β15, in agreement with crosslinking results obtained by Shiota et al., 2015 (Figure 1.13) (*41*). Although we don't see clear interaction between Tom20 and Tom6, partly limited by the resolution, both of our maps show Tom20 in a close enough distance where it could interact with the disordered region of Tom6, in agreement with crosslinking studies suggesting Tom6 involvement at the beginning of translocation (*119*). It is possible that upon presequence recognition, Tom20 becomes active and moves towards Tom6, where it holds on to the longest loop in Tom40 for stability, and deposits the presequence into the pore. In that sense, we suggest that Tom20 remains in this position until the whole precursor protein has been translocated, ensuring that it stays unfolded by means of hydrophobic interactions. Once the protein has entered the mitochondrion, the Tom20 receptor would retreat to the center of the complex to restart the translocation process.

In Figure 4.53, the transmebrane domain of Tom20 is only partially visible inside the micelle. As Tom20 does dock to the core complex inside the membrane region, it is understandable that its transmembrane domain exhibits some flexibility as a result from the subunits movement in the cytoplasm. Although we have identified

two conformations of the TOM holo complex, it is possible that Tom20 takes on more conformations, but that these two are the most stable. This flexibility and limited interaction with the rest of the complex would explain the easy dissociation of Tom20 during purification. The fact that only a portion of our data set provided information on Tom20 might suggest that Tom20 has partially dissociated from our sample.

Both holo maps show one copy of Tom20 per core dimer. However, our LILBID analysis revealed peaks corresponding to subcomplexes formed by up to two copies of Tom20 (Figure 4.10). Our larger particle classification (Figure 4.55) shows a set of elongated low-resolution dimers with clear cytoplasmic domains. Some of these maps contain a density protruding from both sides of the complex. However, at such a low resolution, we consider it best not to over interpret the maps. Another explanation for Tom20 dimers in LILBID might be the presence of tetramers in the sample. In early stages of classification, we separated particles that were considerably larger than TOM dimer (Figure 4.54). Although we were not able to refine these potential tetramers, it is possible that the Tom20s of each TOM dimer are closer and interact strongly. A third possibility is that the second copy of Tom20 is always present in the dimer, but that its high flexibility renders it invisible in our averaging process through single-particle cryoEM.

Our AlphaFold prediction shows that TOM holo complex containing two copies of Tom20 could be viable (Figure 4.18) in a conformation close to our C_2 model. However, as shown in Figure 4.19, both copies of Tom20 would need to be in an elevated position in order to avoid clashes. We attempted to reproduce this by fitting two copies of Tom20 in our TOM core model, in conformations C_1 and C_2 (Figure 5.12), but found that their receptor domains come too close together not to clash. The presence of two Tom20s per dimer may be justified, nonetheless, because of the symmetry of the core complex, where each Tom20-Tom22 receptor could serve a single translocation pore.

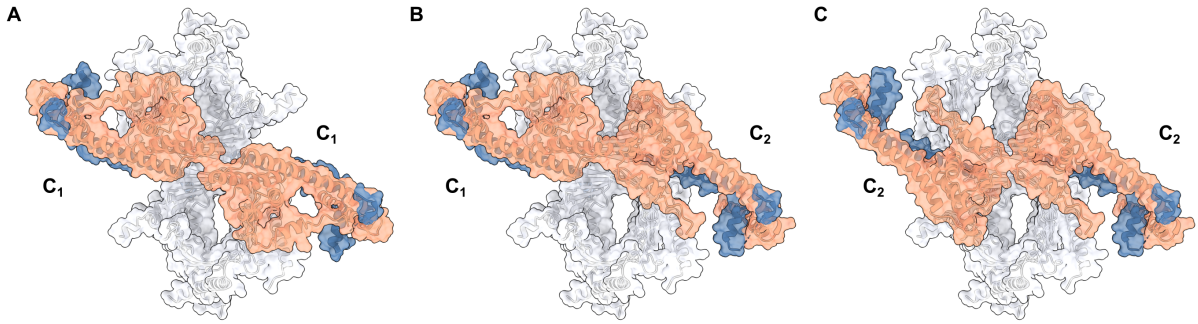

Figure 5.12: **Model of TOM with two copies of Tom20.** Ribbon and surface representation of TOM core dimer model with two copies of Tom20 in its different possible conformations: A) C_1 and C_1, B) C_1 and C_2, and C) C_2 and C_2. Taken from Ornelas et al.

Recently, a published study presented a cryoEM map and model of human TOM complex with two copies of Tom20. Su et al., (2022) crossliked TOM holo complex and obtained a high-resolution structure of TOM core complex, with a low-resolution Tom20 peaking out on the cytoplasmic side. By applying C2 symmetry and a cytoplasmic mask, the authors obtained a structure of the receptor domain of Tom20, crosslinked to Tom40, at 13 Å. However, their low-resolution Tom20 map is independent from the TOM core map and does not contain enough features to allow modelling of the TPR domain. Figure 5.13 provides a side-by-side comparison of our non-symmetrised TOM holo map and Su's symmetric, crosslinked Tom20 and TOM core maps.

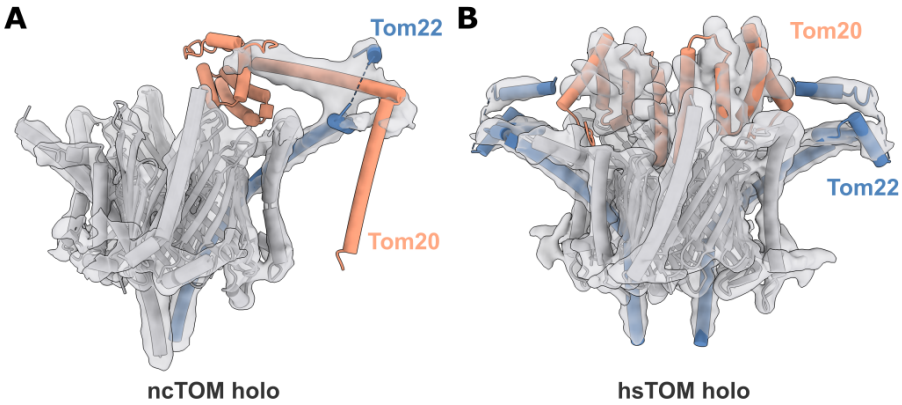

Figure 5.13: **Map comparison with human TOM.** View from the membrane side of the map and model superposition of: A) *N. crassa* TOM holo complex with Tom20 in conformation C_1 at 6.7 Å (PDB 8B4I, EMDB-15850). B) *H. sapiens* TOM core complex at 3.7 Å (PDB 7VDD, EMDB-31888) and Tom20 at 13 Å (PDB 7VC9, EMDB-31889).

In their map, Su et al. identified two helical domains at the edge of the complex, assigned as the N-terminal of Tom22. Their position, however, matches our Tom20 model, as it interacts with Tom22 and goes into the membrane. It is possible that these domains were misassigned and that they indeed are part of human Tom20.

Unlike our TOM holo dimer models, fitted with two copies of Tom20 (Figure 5.12), the crosslinked structure of Su et al., (2022) fits two Tom20 receptor domains without any clashes. In their case, each receptor domain hovers on top of one translocation pore, in a position closer to our C_1 conformation rather than to C_2 (Figure 5.14). Although they make no mention of the different conformations of Tom20 over the pores, they suggest a translocation mechanism in which Tom20 receptor domain stands completely perpendicular to the membrane and bends down to the pore after contact with a presequence. They, however, do not provide any further structural evidence to support this, nor do they propose a location for the transmembrane helix of Tom20.

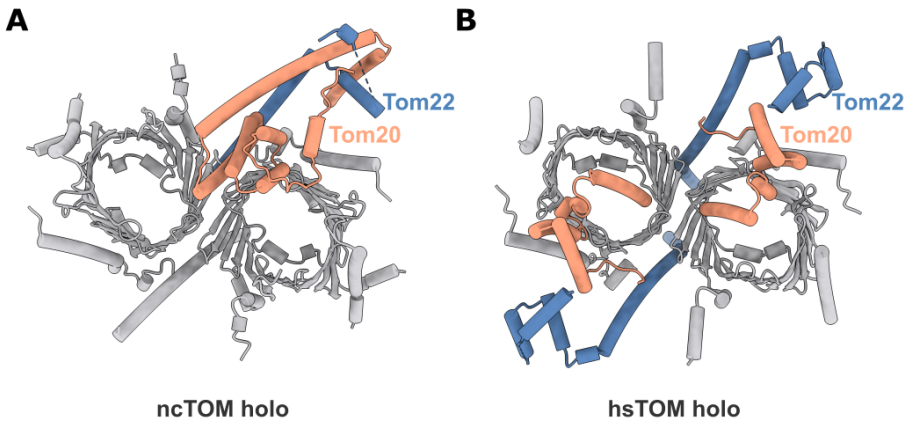

Figure 5.14: **Model comparison with human TOM.** Cytoplasmic view of the A) *N. crassa* (PDB 8B4I) and B) *H. sapiens* (PDB 7VDD + 7VC9) TOM complex. Tom20 and Tom22 are colored in orange and blue respectively. The remaining core subunits are shown in gray.

TOM trimer

As mentioned in previous chapters, the structure of a TOM trimer has been reported using cryoEM (*40*) and mass spectrometry (*41*) techniques. During 2D particle classification in our TOM with pSu9-MBP dataset, we observed a class average with

an additional apparent pore next to the well-known TOM core dimer density (Figure 4.28, class 12). Despite considerable computational strategies to separate TOM trimer particles, we were unable resolve a three pore TOM map. Upon 2D classification of our final TOM holo C_1 dataset, we identified a new class average with an apparent density hole close to the dimer (Figure 4.51, class 4). Analysing the validation report provided by EMDB of our published C_1 map (*120*), we noticed a similar lack of density in the orthogonal projection of the map, next to Tom20. At this high resolution, it is evident that the absence in density does not correspond to a third pore. Notably, the hole lacks the clear features of the Tom40 barrel and is considerably smaller 5.15. In light of this, we propose that in our map, the lack of lipids between Tom20 transmembrane helix and the core complex after solubilization may explain such gap in the density map. Likewise, detergents present in that area might not pack as densely as lipid molecules would in the membrane, resulting in what appears to be a pore.

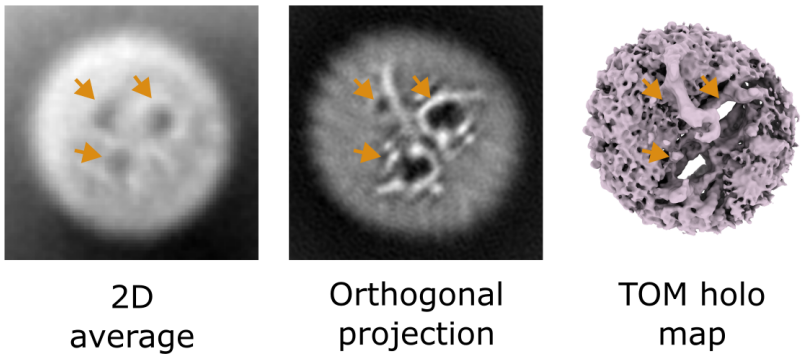

Figure 5.15: **Apparent third pore in our TOM holo complex.** Comparison between a 2D class average depicting an apparent third pore in the TOM + pSu9-MPB dataset, an orthogonal projection obtained from the EMDB validation report of our published TOM holo C_1 map (EMDB-15856), and the TOM holo C_1 map. Orange arrows point at two TOM pores and the apparent thrid pore.

Tom70

Our successfully purified of TOM holo complex in GDN contains the Tom70 subunit. Our SDS-PAGE analysis indicates the presence of Tom70 with a clear band near the 75 kDa marker (Figure 4.3). We demonstrated a stable composition of the complex with our nanoDSF experiments, where saw no indication of the partial disassambly of

the complex prior to melting temperature. This suggests that Tom70 is unlikely to detach from the complex after purification (Figure 4.6). Despite this, we were not able to locate Tom70 in our TOM holo complex maps.

Structure prediction analysis showed that Tom70 is has 28 α-helices, including a transmembrane helix at its N-terminal, followed by a disordered domain that connects to its receptor domain. The predictions we obtained with both, Quick2D and AlphaFold (Figures 4.13 and 4.14) are in agreement with the published crystal and cryoEM structures of Tom70 (*44–46*). However, AlphaFold multimer predictions with Tom70 were tricky, as they consistently showed its transmembrane region siting close to the receptor groove. When running the predictions with its soluble domain alone, we saw Tom70 located on top of the translocation pores. However, when Tom20 was included in the prediction, it would displace Tom70, leaving Tom70 on the side, independent from the rest of the complex. This may suggest that Tom20 and Tom70 may compete for interaction with the pores in order to commence precursor protein translocation.

Our MS experiments reveal nonetheless interesting interactions of Tom70 with other TOM subunits. In our crosslinking studies we saw multiple links between Tom20 and Tom70 (Figure 4.23), hinting at a possible collaboration between both receptors, which has been proposed in the past (*114*). In our LILBID results, we were able to identify multiple peaks corresponding to Tom70, alone and with other subunits (Figure 4.10). Once more, we saw Tom70 interacting with Tom20, but also with Tom40 and Tom22, which further hints at the synergy of the three receptors in translocation. We also saw Tom70 interacting with two smaller TOM subunits, which might be helpful in identifying the Tom70 transmembrane helix in the holo complex, but because of their small and similar sizes, we are not able to identify them.

In both our cryoEM datasets, we identified one or two small dark features attached to some of the extracted particles prior to classification (Figures 4.27 and 4.34). Figure 5.16 shows particles from both cryoEM datasets containing these features. They appear consistently at a short distance from the translocation pores, mostly on the cytoplasmic

side of the complex. These features may correspond to the large soluble domain of Tom70, confirming that it is stably attached to the complex after purification. In that case, we may argue that because of the long, unstructured linker joining the soluble and transmembrane domains of Tom70, Tom70 was too flexible to be resolved in our single-particle analysis.

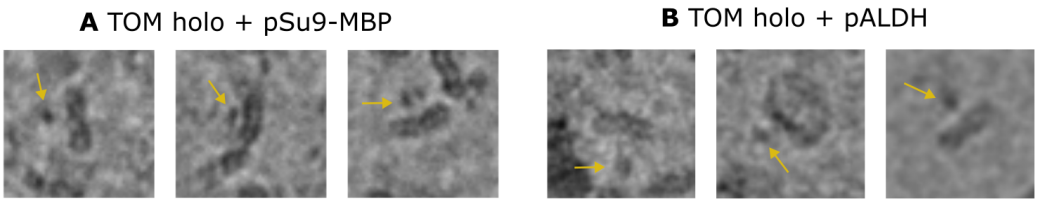

Figure 5.16: **TOM particles containing Tom70.** Extracted particles taken from the A) TOM holo + pSu9-MBP and B) TOM holo + pALDH datasets, which display a round feature at a short distance from the complex (yellow arrows).

A second explanation is that Tom70 is mostly hidden in the larger particles that we found in our TOM holo with pALDH dataset (Figure 4.54). It is possible that Tom70 is more stable when it is part of tetrameric TOM holo or that it is present in larger TOM holo dimers shown in Figure 4.55. As mentioned in the previous section however, despite different processing approaches applied to these particles, they only produced lower-resolution maps with limited information. It is likely that the flexibility of both Tom20 and Tom70 increased the amount of noise in the data, complicating the overall alignment process.

The remaining question is the mechanism by which Tom70 feeds presequences into the translocation pore. It is unclear if Tom70 is able to move towards Tom40 on its own, or if it is accompanied by chaperones. Another possibility agrees with the proposal made by Su et al. with Tom20 protruding vertically from the membrane, bending down towards the pores after receiving the precursor from Tom70. Either way, our structural model provides no discernible feature to infer such movement in the receptors.

Chapter 6

Conclusions and Outlook

The TOM complex is essential to cell viability because of its role as main entry gate into mitochondria. Basic mitochondrial activities, such as cellular respiration, depend on its correct function. Disruption of the complex can hinder mitophagy and lead to a variety of diseases. Unveiling the structure of TOM complex is of great consequence, because it leads to a better understanding of its translocation mechanism, and more importantly, of how it can fail. Structural biology, in particular single particle cryoEM provide us with great insights into the structure of proteins. However, handling of protein samples during isolation limit the information that we can obtain from them. Furthermore, flexible proteins, such as the subunits of TOM holo complex, can be difficult to resolve using single-particle, with subunit movement generating noise and interfering with the resolution of obtained maps. Therefore, we found great support in MS techniques during the development of this project.

Thanks to our diverse experimentation, we confirmed the presence of seven TOM holo subunits in our sample and identified their interactions. Using cryoEM, we obtained the first high-resolution structure of *N. crassa* TOM core complex, and modeled a homodimer containing its five characteristic subunits. Additionally, we identified a lipid siting at the symmetry axis of the dimer, between two Tom40 β-barrels, and eight detergent molecules that indicating lipid binding sites in the membrane. Using our model, we were able to compare *N. crassa* TOM complex with that of human and

yeast mitochondria, and identified key differences in their structure. These differences hint at evolutionary adaptations of translocation pathways and diverse roles that TOM subunits can take.

Our cryoEM data also allowed us to visualize the translocation of pALDH in our TOM core map. Despite its lower resolution, its position inside the pore agrees with published research and brings new insight on the residues that guide the precursors through Tom40. Interestingly, we identified a pair of phenylalanines at the exit pathway that may be key actors in the gating mechanism of the complex. We also obtained the first TOM holo structure that shows how Tom20 associates with the complex, and two different conformations that it can take. These two conformations give us insight of changes that occur in TOM when incubated with a presequence. In the analysis of Tom20, XL-MS and LILBID were of great help for model building. Evidence of the interactions of Tom20 with other TOM subunits, particularly the $Tom20_1Tom22_1Tom40_1$ subcomplex, in combination with statistical predictions of AlphaFold, allowed us to confidently fit Tom20 into our two maps.

In our TOM holo maps we saw that Tom20 clearly binds to Tom22. We found that Tom22 not only serves as a pivot point for Tom20, but that its disordered N-terminal is wrapped around Tom20, displaying complementary charge interactions. As this marks the only point of contact between Tom20 and the complex, the Tom22 N-terminal region is fundamental for stability of the holo complex. All evidence now points towards Tom22 not only acting as a receptor, but as a cooperation partner in presequence recognition together with Tom20. Unfortunately, we were not able to observe a high-resolution interaction between the receptors and our precursor proteins. More efficient crosslinking, and possibly further optimization of presequence design, will be necessary to see this interaction through cryoEM. An *in silico* analysis of the translocation mechanism would also be helpful in visualizing these interactions. The TOM complex is a relatively small system and would be a target for molecular dynamics

simulations. This would allow us to observe the binding mechanism of presequences to the Tom20-Tom22 receptor and how they are deposited into the translocation pore.

XL-MS and LILBID also provided many hints on Tom70. Both techniques demonstrated abundant interactions between Tom20 and Tom70, which points towards their cooperation in the early stages of translocation. We also saw Tom70 interacting with Tom22 and Tom40, which indicates a close proximity of Tom70 transmembrane helix to Tom22 and Tom40, or that the Tom70 soluble domain comes close enough to the pore to form a strong interaction with them. Further crosslinking analysis can help us identify a binding site of Tom70 to the complex and understand its collaboration with Tom20.

Knowing all this, we suggest the following translocation mechanism. In a first instance, we propose that Tom20 and Tom70 make the initial contact with the helical mitochondrial targeting sequences. Tom70 hands down the precursors to Tom20, which can also detect incoming presequences independently.

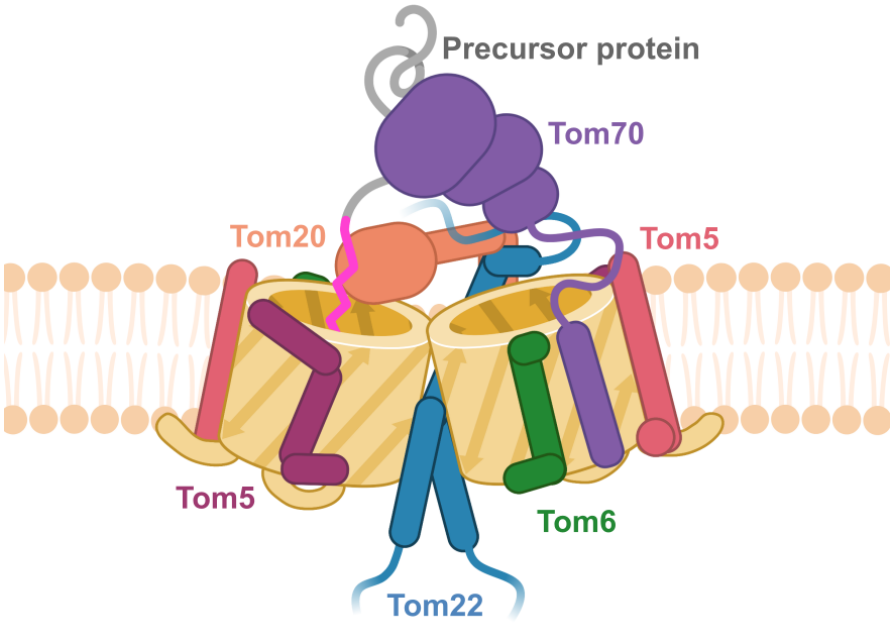

Figure 6.1: **Hypothetical translocation model.** Proposed model of TOM's translocation mechanism, showing cooperation between Tom20, Tom22 and Tom70 in precursor protein detection and insertion into the Tom40 translocation pore. Created with BioRender.com.

Tom22 stabilizes Tom20 in its position as it directs precursor proteins towards a free translocation pore. There, the Tom20 receptor domain makes contact with the longest loop in Tom40, possibly stabilized by the disordered N-terminal of Tom6. This interaction might be necessary to keep precursor proteins unfolded during translocation. Presequences then insert into the Tom40 pore, and are translocated by means of electrostatic interactions, and held in place by hydrophobic interactions with the inner helices of the pore. It remains unclear whether Tom40 alone is capable of completing the translocation process. It is likely that at this point, the C-terminal of Tom22, in cooperation with the small TIMs or TIM23, fish presequences out of the pore from inside the IMS.

An open question remaining is whether both pores can operate simultaneously. Recent studies have shown that the complex is dynamic and has different levels of activation, suggesting independent translocation through each pore. However, the exact stoichiometry of the TOM holo complex remains unknown and it is likely that the number of receptors per complex has an effect on its translocation efficiency. Likewise, we have seen the existence of higher oligomers formed by TOM. It is unclear if they are an artifact caused by isolation or whether the complex would also form these same clusters in its native environment.

Further work will be necessary to study Tom70. Despite its large size, we were not able to identify it as a clear protein density in our analysis, which points towards it being part of our holo particles. It seems that its flexibility makes it difficult to resolve in single'particle analysis. However, the evidence of its attachment to some of the extracted particles makes us hopeful. It is possible that studying mitochondria or OMVs by cryoET sheds new light into the role of Tom70 in the complex.

Bibliography

(*1*) Chinnery, P. F. (2003). Mitochondria. *Journal of Neurology, Neurosurgery & Psychiatry 74*, 1188–1199.

(*2*) Campo, M. L., Peixoto, P. M., and Martínez-Caballero, S. (2017). Revisiting trends on mitochondrial mega-channels for the import of proteins and nucleic acids. *Journal of Bioenergetics and Biomembranes 49*, 75–99.

(*3*) Kühlbrandt, W. (2015). Structure and function of mitochondrial membrane protein complexes. *BMC Biology 13*, 89.

(*4*) Kühlbrandt, W. (2019). Structure and Mechanisms of F-Type ATP Synthases. *Annual Review of Biochemistry 88*, 515–549.

(*5*) Zimorski, V., Ku, C., Martin, W. F., and Gould, S. B. (2014). Endosymbiotic theory for organelle origins. *Current Opinion in Microbiology 22*, 38–48.

(*6*) Schmidt, O., Pfanner, N., and Meisinger, C. (2010). Mitochondrial protein import: from proteomics to functional mechanisms. *Nature Reviews Molecular Cell Biology 11*, Number: 9 Publisher: Nature Publishing Group, 655–667.

(*7*) Neupert, W., and Herrmann, J. M. (2007). Translocation of Proteins into Mitochondria. *Annual Review of Biochemistry 76*, 723–749.

(*8*) Wiedemann, N., and Pfanner, N. (2017). Mitochondrial Machineries for Protein Import and Assembly. *Annual Review of Biochemistry 86*, 685–714.

(*9*) Hansen, K. G., and Herrmann, J. M. (2019). Transport of Proteins into Mitochondria. *The Protein Journal 38*, 330–342.

(*10*) Vögtle, F. .-., Wortelkamp, S., Zahedi, R. P., Becker, D., Leidhold, C., Gevaert, K., Kellermann, J., Voos, W., Sickmann, A., Pfanner, N., and Meisinger, C. (2009). Global Analysis of the Mitochondrial N-Proteome Identifies a Processing Peptidase Critical for Protein Stability. *Cell 139*, 428–439.

(*11*) Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T., and Pfanner, N. (2009). Importing Mitochondrial Proteins: Machineries and Mechanisms. *Cell 138*, 628–644.

(*12*) Dolezal, P., Likic, V., Tachezy, J., and Lithgow, T. (2006). Evolution of the Molecular Machines for Protein Import into Mitochondria. *Science 313*, 314–318.

(*13*) Pfanner, N., Warscheid, B., and Wiedemann, N. (2019). Mitochondrial proteins: from biogenesis to functional networks. *Nature Reviews Molecular Cell Biology 20*, 267–284.

(*14*) Gupta, A., and Becker, T. (2021). Mechanisms and pathways of mitochondrial outer membrane protein biogenesis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics 1862*, 148323.

(*15*) Harbauer, A. B., Zahedi, R. P., Sickmann, A., Pfanner, N., and Meisinger, C. (2014). The Protein Import Machinery of Mitochondria—A Regulatory Hub in Metabolism, Stress, and Disease. *Cell Metabolism 19*, 357–372.

(*16*) Model, K., Prinz, T., Ruiz, T., Radermacher, M., Krimmer, T., Kühlbrandt, W., Pfanner, N., and Meisinger, C. (2002). Protein translocase of the outer mitochondrial membrane: role of import receptors in the structural organization of the TOM complex. *Journal of Molecular Biology 316*, 657–666.

(*17*) Pitt, A. S., and Buchanan, S. K. (2021). A Biochemical and Structural Understanding of TOM Complex Interactions and Implications for Human Health and Disease. *Cells 10*, 1164.

(*18*) Künkele, K.-P., Heins, S., Dembowski, M., Nargang, F. E., Benz, R., Thieffry, M., Walz, J., Lill, R., Nussberger, S., and Neupert, W. (1998). The Preprotein Translocation Channel of the Outer Membrane of Mitochondria. *Cell 93*, 1009–1019.

(*19*) Ahting, U., Thun, C., Hegerl, R., Typke, D., Nargang, F. E., Neupert, W., and Nussberger, S. (1999). The Tom Core Complex. *Journal of Cell Biology 147*, 959–968.

(*20*) Mayer, A., Neupert, W., and Lill, R. (1995). Mitochondrial Protein Import: Reversible Binding of the Presequence at the trans Side of the Outer Membrane Drives Partial Translocation and Unfolding. *80*, 127–137.

(*21*) Stan, T., Ahting, U., Dembowski, M., Künkele, K.-P., Nussberger, S., Neupert, W., and Rapaport, D. (2000). Recognition of preproteins by the isolated TOM complex of mitochondria. *The EMBO Journal 19*, 4895–4902.

(*22*) Gold, V. A., Ieva, R., Walter, A., Pfanner, N., van der Laan, M., and Kühlbrandt, W. (2014). Visualizing active membrane protein complexes by electron cryotomography. *Nature Communications 5*, 4129.

(*23*) Bausewein, T., Mills, D. J., Langer, J. D., Nitschke, B., Nussberger, S., and Kühlbrandt, W. (2017). Cryo-EM Structure of the TOM Core Complex from Neurospora crassa. *Cell 170*, 693–700.

(*24*) Mager, F., Sokolova, L., Lintzel, J., Brutschy, B., and Nussberger, S. (2010). LILBID-mass spectrometry of the mitochondrial preprotein translocase TOM. *Journal of Physics: Condensed Matter 22*, 454132.

(*25*) Tucker, K., and Park, E. (2019). Cryo-EM structure of the mitochondrial protein-import channel TOM complex at near-atomic resolution. *Nature Structural & Molecular Biology 26*, 1158–1166.

(*26*) Araiso, Y. et al. (2019). Structure of the mitochondrial import gate reveals distinct preprotein paths. *Nature 575*, 395–401.

(*27*) Wang, W., Chen, X., Zhang, L., Yi, J., Ma, Q., Yin, J., Zhuo, W., Gu, J., and Yang, M. (2020). Atomic structure of human TOM core complex. *Cell Discovery 6*, 1–10.

(*28*) Guan, Z., Yan, L., Wang, Q., Qi, L., Hong, S., Gong, Z., Yan, C., and Yin, P. (2021). Structural insights into assembly of human mitochondrial translocase TOM complex. *Cell Discovery 7*, 22.

(*29*) Wang, S., Findeisen, L., Leptihn, S., Wallace, M. I., Hörning, M., and Nussberger, S. (2022). Spatiotemporal stop-and-go dynamics of the mitochondrial TOM core complex correlates with channel activity. *Communications Biology 5*, 471.

(*30*) Ahting, U., Thieffry, M., Engelhardt, H., Hegerl, R., Neupert, W., and Nussberger, S. (2001). Tom40, the Pore-Forming Component of the Protein-Conducting Tom Channel in the Outer Membrane of Mitochondria. *Journal of Cell Biology 153*, 1151–1160.

(*31*) Gessmann, D., Flinner, N., Pfannstiel, J., Schlösinger, A., Schleiff, E., Nussberger, S., and Mirus, O. (2011). Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry. *Biochimica et Biophysica Acta (BBA) - Bioenergetics 1807*, 1647–1657.

(*32*) Vasiljev, A., Ahting, U., Nargang, F., Go, N., Habib, S., Kozany, C., Panneels, V., Sinning, I., Prokisch, H., Nussberger, S., and Rapaport, D. (2004). Reconstituted TOM Core Complex and Tim9/Tim10 Complex of Mitochondria Are Sufficient for Translocation of the ADP/ATP Carrier across Membranes. *Molecular Biology of the Cell 15*, 1445–1458.

(*33*) Naveed, H., Jackups, R., and Liang, J. (2009). Predicting weakly stable regions, oligomerization state, and protein–protein interfaces in transmembrane domains of outer membrane proteins. *Proceedings of the National Academy of Sciences 106*, 12735–12740.

(*34*) van Wilpe, S., Ryan, M. T., Hill, K., Maarse, A. C., Meisinger, C., Brix, J., Dekker, P. J. T., Moczko, M., Wagner, R., Meijer, M., Guiard, B., Hönlinger, A., and Pfanner, N. (1999). Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. *Nature 401*, 485–489.

(*35*) Araiso, Y., Imai, K., and Endo, T. (2022). Role of the TOM Complex in Protein Import into Mitochondria: Structural Views. *Annual Review of Biochemistry 91*, 679–703.

(*36*) Bolliger, L., Junne, T., Schatz, G., and Lithgow, T. (1995). Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria. *The EMBO Journal 14*, 6318–6326.

(*37*) Hönlinger, A., Bömer, U., Alconada, A., Eckerskorn, C., Lottspeich, F., Dietmeier, K., and Pfanner, N. (1996). Tom7 modulates the dynamics of the mitochondrial outer membrane translocase and plays a pathway-related role in protein import. *The EMBO Journal 15*, 2125–2137.

(*38*) Dembowski, M., Künkele, K.-P., Nargang, F. E., Neupert, W., and Rapaport, D. (2001). Assembly of Tom6 and Tom7 into the TOM Core Complex ofNeurospora crassa *. *Journal of Biological Chemistry 276*, 17679–17685.

(*39*) Dietmeier, K., Hönlinger, A., Bömer, U., Dekker, P. J. T., Eckerskorn, C., Lottspeich, F., Kübrich, M., and Pfanner, N. (1997). Tom5 functionally links mitochondrial preprotein receptors to the general import pore. *Nature 388*, 195–200.

(*40*) Model, K., Meisinger, C., and Kühlbrandt, W. (2008). Cryo-Electron Microscopy Structure of a Yeast Mitochondrial Preprotein Translocase. *Journal of Molecular Biology 383*, 1049–1057.

(*41*) Shiota, T. et al. (2015). Molecular architecture of the active mitochondrial protein gate. *Science 349*, 1544–1548.

(*42*) Araiso, Y., Imai, K., and Endo, T. (2021). Structural snapshot of the mitochondrial protein import gate. *The FEBS Journal 288*, 5300–5310.

(*43*) Su, J., Liu, D., Yang, F., Zuo, M.-Q., Li, C., Dong, M.-Q., Sun, S., and Sui, S.-F. (2022). Structural basis of Tom20 and Tom22 cytosolic domains as the human TOM complex receptors. *Proceedings of the National Academy of Sciences 119*, DOI: `10.1073/pnas.2200158119`.

(*44*) Wu, Y., and Sha, B. (2006). Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p. *Nature Structural & Molecular Biology 13*, 589–593.

(*45*) Gordon, D. E. et al. (2020). Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science 370*.

(*46*) Gao, X., Zhu, K., Qin, B., Olieric, V., Wang, M., and Cui, S. (2021). Crystal structure of SARS-CoV-2 Orf9b in complex with human TOM70 suggests unusual virus-host interactions. *Nature Communications 12*, 2843.

(*47*) Brix, J., Dietmeier, K., and Pfanner, N. (1997). Differential Recognition of Preproteins by the Purified Cytosolic Domains of the Mitochondrial Import Receptors Tom20, Tom22, and Tom70. *Journal of Biological Chemistry 272*, 20730–20735.

(*48*) Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S.-i., Endo, T., and Kohda, D. (2000). Structural Basis of Presequence Recognition by the Mitochondrial Protein Import Receptor Tom20. *Cell 100*, 551–560.

(*49*) Saitoh, T., Igura, M., Obita, T., Ose, T., Kojima, R., Maenaka, K., Endo, T., and Kohda, D. (2007). Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. *The EMBO Journal 26*, 4777–4787.

(*50*) Mayer, A., Nargang, F. E., Neupert, W., and Lill, R. (1995). MOM22 is a receptor for mitochondrial targeting sequences and cooperates with MOM19. *The EMBO Journal 14*, 4204–4211.

(*51*) Shiota, T., Mabuchi, H., Tanaka-Yamano, S., Yamano, K., and Endo, T. (2011). In vivo protein-interaction mapping of a mitochondrial translocator protein Tom22 at work. *Proceedings of the National Academy of Sciences 108*, 15179–15183.

(*52*) Rochow, E., *An Introduction to Microscopy by Means of Light, Electrons, X-Rays, or Ultrasound*, Google-Books-ID: g1jhBwAAQBAJ; Springer Science & Business Media: 2012; 379 pp.

(*53*) Williams, D. B., and Carter, C. B., *Transmission electron microscopy: a textbook for materials science*, 2nd ed; Springer: New York, 2008; 4 pp.

(*54*) *Light Microscopy: Methods and Protocols*; Markaki, Y., and Harz, H., Eds.; Methods in Molecular Biology, Vol. 1563; Springer New York: New York, NY, 2017.

(*55*) Kasai, N., Kakudo, M., and Kasai, N., *X-ray diffraction by macromolecules*; Springer series in chemical physics 80; Kodansha ; Springer: Tokyo : Berlin ; New York, 2005; 504 pp.

(*56*) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research 28*, 235–242.

(*57*) Marion, D. (2013). An Introduction to Biological NMR Spectroscopy. *Molecular & Cellular Proteomics 12*, 3006–3025.

(*58*) Jiang, Y., Chen, Z., Han, Y., Deb, P., Gao, H., Xie, S., Purohit, P., Tate, M. W., Park, J., Gruner, S. M., Elser, V., and Muller, D. A. (2018). Electron ptychography of 2D materials to deep sub-ångström resolution. *Nature 559*, 343–349.

(*59*) Nakane, T. et al. (2020). Single-particle cryo-EM at atomic resolution. *Nature 587*, 152–156.

(*60*) Reimer, L., *Transmission Electron Microscopy: Physics of Image Formation and Microanalysis*, Google-Books-ID: CUfsCAAAQBAJ; Springer: 2013; 532 pp.

(*61*) Orlova, E. V., and Saibil, H. R. (2011). Structural Analysis of Macromolecular Assemblies by Electron Microscopy. *Chemical Reviews 111*, 7710–7748.

(*62*) Kühlbrandt, W. (2014). The Resolution Revolution. *Science 343*, Publisher: American Association for the Advancement of Science, 1443–1444.

(*63*) Henderson, R., and Clarke, B. C. (1997). Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction. *Proceedings of the Royal Society of London. Series B: Biological Sciences 241*, 6–8.

(*64*) Fernandez-Leiro, R., and Scheres, S. H. W. (2016). Unravelling biological macromolecules with cryo-electron microscopy. *Nature 537*, Number: 7620 Publisher: Nature Publishing Group, 339–346.

(*65*) Dubochet, J., Chang, J. .-., Freeman, R., Lepault, J., and McDowall, A. W. (1982). Frozen aqueous suspensions. *Ultramicroscopy 10*, 55–61.

(*66*) Marston, J. O., Vakarelski, I. U., and Thoroddsen, S. T. (2012). Cavity formation by the impact of Leidenfrost spheres. *Journal of Fluid Mechanics 699*, 465–488.

(*67*) Gross, J. H. In *Mass Spectrometry: A Textbook*, Gross, J. H., Ed.; Springer International Publishing: Cham, 2017, pp 1–28.

(*68*) Leney, A. C., and Heck, A. J. R. (2017). Native Mass Spectrometry: What is in the Name? *Journal of the American Society for Mass Spectrometry 28*, 5–13.

(*69*) Peetz, O., Hellwig, N., Henrich, E., Mezhyrova, J., Dötsch, V., Bernhard, F., and Morgner, N. (2019). LILBID and nESI: Different Native Mass Spectrometry Techniques as Tools in Structural Biology. *Journal of the American Society for Mass Spectrometry 30*, 181–191.

(*70*) Steigenberger, B., Albanese, P., Heck, A. J. R., and Scheltema, R. A. (2020). To Cleave or Not To Cleave in XL-MS? *Journal of the American Society for Mass Spectrometry 31*, 196–206.

(*71*) Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., and Agard, D. A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods 14*, 331–332.

(*72*) Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology 192*, 216–221.

(*73*) Nogales, E., and Scheres, S. H. W. (2015). Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Molecular Cell 58*, 677–689.

(*74*) Penczek, P., Radermacher, M., and Frank, J. (1992). Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy 40*, 33–53.

(*75*) Jumper, J. et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature 596*, Number: 7873 Publisher: Nature Publishing Group, 583–589.

(*76*) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of *Coot*. *Acta Crystallographica Section D Biological Crystallography 66*, 486–501.

(*77*) Liebschner, D. et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology 75*, 861–877.

(*78*) MacKenzie, J. A., and Payne, R. M. (2007). Mitochondrial protein import and human health and disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1772*, 509–523.

(*79*) Sokol, A. M., Sztolsztener, M. E., Wasilewski, M., Heinz, E., and Chacinska, A. (2014). Mitochondrial protein translocases for survival and wellbeing. *FEBS Letters 588*, 2484–2495.

(*80*) Baker, K. P., and Schatz, G. (1991). Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria. *Nature 349*, 205–208.

(*81*) Bigi, A., Cascella, R., and Cecchi, C. (2023). -Synuclein oligomers and fibrils: partners in crime in synucleinopathies. *Neural Regeneration Research 0*, 0.

(*82*) Galagan, J. E. et al. (2003). The genome sequence of the filamentous fungus Neurospora crassa. *Nature 422*, Number: 6934 Publisher: Nature Publishing Group, 859–868.

(*83*) Schmitt, S., Prokisch, H., Schlunck, T., Camp, D. G., Ahting, U., Waizenegger, T., Scharfe, C., Meitinger, T., Imhof, A., Neupert, W., Oefner, P. J., and Rapaport, D. (2006). Proteome analysis of mitochondrial outer membrane from Neurospora crassa. *Proteomics 6*, 72–80.

(*84*) Maschberger, M., Resch, H. M., Duhr, S., and Breitsprecher, D. Exploring Protein Stability by nanoDSF. Prometheus NT.48 – The Stability Expert. NanoTemper Technologies GmbH.

(*85*) Morgner, N., Barth, H.-D., Brutschy, B., Morgner, N., Barth, H.-D., and Brutschy, B. (2006). A New Way To Detect Noncovalently Bonded Complexes of Biomolecules from Liquid Micro-Droplets by Laser Mass Spectrometry. *Australian Journal of Chemistry 59*, 109–114.

(*86*) Morgner, N., and Robinson, C. V. (2012). *Mass* ign: An Assignment Strategy for Maximizing Information from the Mass Spectra of Heterogeneous Protein Assemblies. *Analytical Chemistry 84*, 2939–2948.

(*87*) Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research 31*, 3784–3788.

(*88*) Combe, C. W., Fischer, L., and Rappsilber, J. (2015). xiNET: cross-link network maps with residue resolution. *Molecular & cellular proteomics: MCP 14*, 1137–1147.

(*89*) Kimanius, D., Dong, L., Sharov, G., Nakane, T., and Scheres, S. H. W. (2021). New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal 478*, 4169–4185.

(*90*) Punjani, A., Rubinstein, J. L., Fleet, D. J., and Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods 14*, mog, 290–296.

(*91*) Punjani, A., and Fleet, D. J. (2021). 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology 213*, 107702.

(*92*) Wagner, T. et al. (2019). SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology 2*, Number: 1 Publisher: Nature Publishing Group, 1–13.

(*93*) Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., and Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science: A Publication of the Protein Society 30*, 70–82.

(*94*) Armenteros, J. J. A., Salvatore, M., Emanuelsson, O., Winther, O., Heijne, G. v., Elofsson, A., and Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance 2*, DOI: `10.26508/lsa.201900429`.

(*95*) Gabler, F., Nam, S.-Z., Till, S., Mirdita, M., Steinegger, M., Söding, J., Lupas, A. N., and Alva, V. (2020). Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Current Protocols in Bioinformatics 72*, e108.

(*96*) Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A. N., and Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology 430*, 2237–2243.

(*97*) Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices11Edited by G. Von Heijne. *Journal of Molecular Biology 292*, 195–202.

(*98*) Heffernan, R., Yang, Y., Paliwal, K., and Zhou, Y. (2017). Capturing non-local interactions by long short-term memory bidirectional recurrent neural networks for improving prediction of protein secondary structure, backbone angles, contact numbers and solvent accessibility. *Bioinformatics 33*, 2842–2849.

(*99*) Yan, R., Xu, D., Yang, J., Walker, S., and Zhang, Y. (2013). A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Scientific Reports 3*, Number: 1 Publisher: Nature Publishing Group, 2619.

(*100*) Wang, S., Peng, J., Ma, J., and Xu, J. (2016). Protein Secondary Structure Prediction Using Deep Convolutional Neural Fields. *Scientific Reports 6*, Number: 1 Publisher: Nature Publishing Group, 18962.

(*101*) Gruber, M., Söding, J., and Lupas, A. N. (2006). Comparative analysis of coiled-coil prediction methods. *Journal of Structural Biology 155*, 140–145.

(*102*) Delorenzi, M., and Speed, T. (2002). An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics 18*, 617–625.

(*103*) Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *Journal of Molecular Biology 305*, 567–580.

(*104*) Käll, L., Krogh, A., and Sonnhammer, E. L. L. (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology 338*, 1027–1036.

(*105*) Käll, L., Krogh, A., and Sonnhammer, E. L. L. (2005). An HMM posterior decoder for sequence feature prediction that includes homology information. *Bioinformatics 21*, i251–i257.

(*106*) Jones, D. T., and Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics 31*, 857–863.

(*107*) Hanson, J., Yang, Y., Paliwal, K., and Zhou, Y. (2017). Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks. *Bioinformatics 33*, 685–692.

(*108*) Dosztányi, Z., Csizmók, V., Tompa, P., and Simon, I. (2005). The Pairwise Energy Content Estimated from Amino Acid Composition Discriminates between Folded and Intrinsically Unstructured Proteins. *Journal of Molecular Biology 347*, 827–839.

(*109*) Evans, R. et al. (2022). Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.10.04.463034.

(*110*) Croll, T. I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology 74*, Number: 6 Publisher: International Union of Crystallography, 519–530.

(*111*) Jurrus, E. et al. (2018). Improvements to the APBS biomolecular solvation software suite. *Protein Science 27*, 112–128.

(*112*) Pfanner, N., Tropschug, M., and Neupert, W. (1987). Mitochondrial protein import: Nucleoside triphosphates are involved in conferring import-competence to precursors. *Cell 49*, 815–823.

(*113*) Wedde, S., Kleusch, C., Bakonyi, D., and Gröger, H. (2017). High-Throughput Feasible Screening Tool for Determining Enzyme Stabilities against Organic Solvents Directly from Crude Extracts. *ChemBioChem 18*, 2399–2403.

(*114*) Yamano, K., Yatsukawa, Y.-i., Esaki, M., Hobbs, A. E. A., Jensen, R. E., and Endo, T. (2008). Tom20 and Tom22 Share the Common Signal Recognition Pathway in Mitochondrial Protein Import. *Journal of Biological Chemistry 283*, 3799–3807.

(*115*) Hallermayer, G., and Neupert, W. (1974). Lipid Composition of Mitochondrial Outer and Inner Membranes of Neurospora crassa. *Biological Chemistry 355*, Publisher: De Gruyter, 279–288.

(*116*) Dekker, P. J. T., Ryan, M. T., Brix, J., Müller, H., Hönlinger, A., and Pfanner, N. (1998). Preprotein Translocase of the Outer Mitochondrial Membrane: Molecular Dissection and Assembly of the General Import Pore Complex. *Molecular and Cellular Biology 18*, 6515–6524.

(*117*) Sato, T. K., Kawano, S., and Endo, T. (2019). Role of the membrane potential in mitochondrial protein unfolding and import. *Scientific Reports 9*, 7637.

(*118*) Ornelas, P., Bausewein, T., Martin, J., Morgner, N., Nussberger, S., and Kühlbrandt, W. (2023). Two conformations of the Tom20 preprotein receptor in the TOM holo complex. *Proceedings of the National Academy of Sciences 120*, Publisher: Proceedings of the National Academy of Sciences, e2301447120.

(*119*) Yamamoto, H., Itoh, N., Kawano, S., Yatsukawa, Y.-i., Momose, T., Makio, T., Matsunaga, M., Yokota, M., Esaki, M., Shodai, T., Kohda, D., Aiken Hobbs, A. E., Jensen, R. E., and Endo, T. (2011). Dual role of the receptor Tom20 in specificity and efficiency of protein import into mitochondria. *Proceedings of the National Academy of Sciences 108*, Publisher: Proceedings of the National Academy of Sciences, 91–96.

(*120*) EMDB Electron Microscopy Data Bank, Electron Microscopy Data Bank `https://www.ebi.ac.uk/emdb/EMD-15856` (accessed Aug. 15, 2023).

Supplementary material

1) DNA sequence of the pSU9-BC2NB insert in the pUC57 synthetic plasmid provided by Dr. Thomas Bausewein:

CCATGGCTTC TACCCGTGTT CTGGCTTCTC GTCTGGCTTC TCAGATGGCT GCTTCTGCTA AAGTTGCTCG TCCGGCTGTT CGTGTTGCTC AGGTTTCTAA ACGTACCATC CAGACCGGTT CTCCGCTGCA GACCCTGAAA CGTACCCAGA TGACCTCTAT CGTTAACGCT ACCACCCGTC AGGCTTTCCA GAAACGTGCT TACTCTTCTG GTCAGGTTCA GCTGGTTGAA TCTGGTGGTG GTCTGGTTCA GCCGGGTGGT TCTCTGACCC TGTCTTGCAC CGCTTCTGGT TTCACCCTGG ACCACTACGA CATCGGTTGG TTCCGTCAGG CTCCGGGTAA AGAACGTGAA GGTGTTTCTT GCATCAACAA CTCTGACGAC GACACCTACT ACGCTGACTC TGTTAAAGGT CGTTTCACCA TCTTCATGAA CAACGCTAAA GACACCGTTT ACCTGCAGAT GAACTCTCTG AAACCGGAAG ACACCGCTAT CTACTACTGC GCTGAAGCTC GTGGTTGCAA ACGTGGTCGT TACGAATACG ACTTCTGGGG TCAGGGTACC CAGGTTACCG TTTCTTCTAA AAAAAAACAC CACCACCACC ACCACCACCA CTAATAACTC GAG

2) Amino acid sequence of the pSu9-MPB precursor protein:

MASTRVLASR LASQMAASAK VARPAVRVAQ VSKRTIQTGS PLQTLKRTQM TSIVNATTRQ AFQKRAYSSG MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDE ALKDAQTNSS WSHPQFEK-

3) Amino acid sequence of the pALDH precursor protein:

MLRAALSTAR RGPRLSRLLS GGGSWSHPQF EK

Declaration of scientific collaboration

Except where stated otherwise by reference or acknowledgment, the work presented was generated by myself, Pamela Ornelas[1], under supervision of my advisors during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research:

Figure 4.2: Outer membrane vesicles. Prof. Stephan Nussberger[2] produced mitochondrial OMVs from *N. crassa* hyphae, from which I isolated TOM holo complex. Dr. Lea Dietrich[1] collected and provided a tomographic slice of the OMVs.

Figure 4.9: LILBID of TOM core and holo complexes at high-laser intensity. and **Figure 4.10: LILBID of TOM core and holo complexes at low-laser intensity.** Dr. Janosch Martin[3] collected LILBID-MS data under supervision of Prof. Nina Morgner[3]. Dr. Janosch Martin[3] and I performed the analysis and peak assignment.

Figure 4.20: Native-PAGE of TOM holo with pSu9-MBP. Dr. Thomas Bausewein[1] designed the experiment. Prof. Stephan Nussberger[2] provided the α-Tom22 antibody. Dr. Thomas Bausewein[1] and I carried out the experiment.

Figure 4.23: Crosslinking of TOM holo with pSu9-MBP. Dr. Thomas Bausewein[1] and I designed and carried out the crosslinking experiment. Dr. Julian Langer[1] performed XL-MS on both EDC and DSSO crosslinked samples.

Table 4.2: TOM holo complex, subcomplexes and subunits. Dr. Janosch Martin[3] collected LILBID-MS data under supervision of Prof. Nina Morgner[3]. Dr. Janosch Martin[3] and I performed the analysis and peak assignment.

Affiliations:
1 Max-Planck-Institute of Biophysics, Frankfurt, Germany, 60438
2 Department of Biophysics, Institute of Biomaterials and Biomolecular Systems, University of Stuttgart, Stuttgart, Germany, 70569
3 Institute of Physical and Theoretical Chemistry, Goethe University of Frankfurt, Frankfurt, Germany, 60439

Publication statement

The following publication resulted from this work:

Ornelas, P., Bausewein, T., Martin, J., Morgner, N., Nussberger, S., and Kühlbrandt, W. (2023). Two conformations of the Tom20 preprotein receptor in the TOM holo complex. *Proceedings of the National Academy of Sciences.* Volume 120, issue 34, e2301447120. DOI: https://doi.org/10.1073/pnas.2301447120

Permission rights

Whenever a figure, table or text is identical to a previous publication, copyright permission and/or co-author agreement has been obtained as listed below.

The following parts of the thesis have been previously published by PNAS under Creative Commons Attribution License 4.0 (CC BY):

Ornelas, P., Bausewein, T., Martin, J., Morgner, N., Nussberger, S., and Kühlbrandt, W. (2023). *Two conformations of the Tom20 preprotein receptor in the TOM holo complex.* Proceedings of the National Academy of Sciences Volume 120, issue 34, e2301447120.

- **Figure 5.2:** Structure of TOM core complex.
- **Figure 5.10:** Model of TOM holo complex.
- **Figure 5.12:** Model of TOM with two copies of Tom20.

The following parts of the thesis have been previously published by other journals:

- **Figures 1.4 and 1.12A** Taken from Kunkele et al., 1998 (*18*).
  Elsevier license number 5577700761274.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=c626d1d0-8769-4dc4-a70e-4c1cee4ac89f
- **Figure 1.5.** Adapted from Gold et al., 2014 (*22*).
  This is an open access article distributed under the terms of the Creative Commons CC BY-NC-ND 4.0 license (https://creativecommons.org/licenses/by-nc-nd/4.0/).
- **Figures 1.6 and 1.10** Adapted from Bausewein et al., 2017 (*23*).
  Elsevier license number 5578080745045.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=bdd85f9b-f931-4a78-8a43-6055e300b111
- **Figure 1.9.** Adapted from Wang et al., 2022 (*29*).
  This is an open access article distributed under the terms of the Creative Commons CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

- **Figure 1.12B.** Taken from Model et al., 2008 (*40*).
  Elsevier license number 5578670273621.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=ab326d2c-72a6-4efc-b26d-6f63921995b5

- **Figure 1.13.** Adapted from Araiso et al., 2019 (*26*).
  Springer Nature license number 5578670975624.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=376ec177-b804-4352-93da-99eea59db5f5

- **Figure 1.15.** Taken from Williams and Carter, 2008 (*53*).
  Springer Nature license number 5578780319570.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=584c7aec-aab6-4049-9111-f53b071f89ae

- **Figure 1.17.** Taken from Williams and Carter, 2008 (*53*).
  Springer Nature license number 5578780468559.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=1484473b-353e-482c-a599-9ead8e4ec19a

- **Figure 1.16**. Taken from Orlova and Saibil, 2011 (*61*).
  Adapted with permission from the American Chemical Society (ACS) to access, download, copy, display and redistribute, under the stated conditions in the Standard ACS AuthorChoice Usage Agreement. Any further permission related to the material excerpted should be directed to the ACS through https://pubs.acs.org/doi/10.1021/cr100353t.

- **Figure 1.20.** Taken from Fernandez-Leiro et al., 2016 (*64*).
  Springer Nature license number 5578810511450.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=24a9d60a-7c12-41a4-8db0-2e5ea3d8c96e

- **Figure 1.21.** Adapted from Peetz et al., 2019 (*69*).
  Reprinted (adapted) with permission from J. Am. Soc. Mass Spectrom. 2019, 30, 1, 181–191. Copyright 2019 American Chemical Society.

- **Figure 1.23.** Taken from Nogales et al., 2015 (*73*).
  Elsevier license number 5578820737576.
  License: https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=83f905cd-62ba-4dca-90b2-12efe0e9e6d8