**Title**
Structures and Mechanisms of Protein Channels



**Author**
Tucker, Kyle

**Publication Date**
2023

Peer reviewed|Thesis/dissertation

# Structures and Mechanisms of Protein Channels

by

Kyle C. Tucker

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular Cellular Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Stephen Brohawn, Chair
Associate Professor Roberto Zoncu
Professor Ehud Y Isacoff
Associate Professor Hillel Adesnik

Summer 2023

Structures and Mechanisms of Protein Channels

Abstract

Structures and Mechanisms of Protein Channels

by

Kyle C. Tucker

Doctor of Philosophy in Molecular Cellular Biology

University of California, Berkeley

Assistant Professor Stephen Brohawn, Chair

Proteins are biological machines composed of a single or multiple threads of amino acids which fold into three-dimensional shapes. These shapes and surfaces directly participate in many of the molecular interactions essential to life. Some proteins exist in lipid membranes and thus have been termed membrane proteins. Membranes serve as sites of import, export, and signaling and allow for compartmentalization; all critical functions for living organisms. After decades of research throughout the 20th century, many families of membrane proteins were revealed to form transport proteins that facilitate movement of substrates across membranes. The general group of transport proteins can be further divided into two groups: active and passive transporters. Passive transport proteins have been called channels, given their observed activity of opening to allow substrate diffusion down an electrochemical gradient. Channels are diverse both in their architecture and substrates, which include ions, small molecules, and other proteins. Despite the broad importance of channels in a vast array of biology many aspects of how these machines operate at a molecular level are yet to be revealed. Determining the structures of channels, which in many cases was not technically feasible until recently, has provided key insights into the form and function of these intricate machines. Presented here are studies of two different channels: the Translocase of the Outer Mitochondrial Membrane (TOM complex) and the light-gated ion channel, ChRmine.

The first study in this thesis examines the long-awaited architecture and transport mechanisms of the major protein gate of the outer mitochondrial membrane. The high-resolution structure of the TOM complex revealed its overall organization– including all its core subunits, multiple assemblies of the complex, and a glimpse at the pore which provides insights into mechanisms of protein transport. The subsequent study examines the first ever structures of a light-gated ion channel in a native-like membrane environment, using lipid nanodiscs and cryogenic electron-microscopy(cryo-EM). Determining cryo-EM structures of ChRmine was a technical achievement and revealed its surprising trimeric assembly, ion conduction path, and retinal binding pocket. Through analysis of the structure, we were

able to tune the kinetics of ChRmine using a rational engineering approach. Lastly, the latest work on ChRmine paves the path for the determination of structures previously too dynamic to capture. These elusive structures are attractive targets because they could lay the foundation for building the next generation of tools for studying the brain.

To Juanita Marie Tucker

My late mother whose life and love supported me, my pursuits, my passions and my love.

“Love liberates” - Maya Angelou

# Contents

# List of Figures

## Acknowledgments

This thesis would not have been possible without the support of many who have come before me including those who have fostered my growth. I would first like to thank my advisor, Stephen Brohawn, who has welcomed me into the lab, and nurtured my development as a scientist and a person. I want to thank Eunyong Park for introducing me into world of membrane biochemistry and structure, while teaching me to be ambitious with my scientific pursuits.

Next, I want to thank Robert Kurt, who taught me how to see biology as an experimental discipline and provided me my first glimpse into biomedical research. His belief in me flowered into my first research experience in the Espinoza lab, made possible by the Howard Hughes Medical Institute. I would like to also thank Bradley Antanaitis for giving me a place to call home in the physics department while teaching me how to frame biology as dynamic populations and ensembles using the tools of statistics and physics. My time in the Espinoza lab under the supervision of Corrie Gallant-Behm and Chris Abraham matured me into a more professional scientist and where I learned how to work in large lab with many postdocs and how to tackle challenging questions in cancer biology. I want to thank the entire lab at CU Boulder for my time there including the mentoring I received from Joaquin both as a scientist and a climber.

My time in Roberta Gottlieb's lab under the mentorship of Allen Andres was critical to me learning how to take ownership of a project and ideas. Under their guidance I truly learned how to tackle complex scientific questions from a more clinical perspective and the beauty of mitochondrial biology. It was toward the end of my time in the Gottlieb lab that I learned of Jared Rutter, who I was fortunate to meet at a conference at the NIH. Jared spurred my decision to go to graduate school and pursue questions surrounding mitochondrial biology. Can't thank him enough for also serving as mentor from a far during my time in graduate school as well.

I could not discuss my time as a graduate student without Roberto Zoncu. He has been a mentor from the day I set foot on campus and I expect that to continue beyond graduate school. I will always cherish our science and non-science discussions as often they've played a big role in my life.

I would not be the person I am today, nor be where I am without my close friends and family that loved and supported me throughout my life. Amongst these relationships, I would like to thank Hannah Weaver who has been at my side during my introduction to science and believed in me when I stumbled and weathered some of the most challenging times in my life.

As a descendant of enslaved peoples, I would also like to acknowledge the resilience, strength and sacrifices my ancestors made to give me the opportunity to present this thesis. I hope to continue to honor them as I develop as a scientist.

# Chapter 1

# Introduction

Cells and their membranes have existed for nearly 4 billion years and are an essential component of life. Cell membranes are composed primarily of protein and lipids, both of which act as more than a barrier to the external world. As cell membranes have evolved, so have their functions. Early membranes perhaps had basic functions such as forming a barrier, but it is likely evolutionary pressure resulted in these structures developing more complex functions. These include membranes that can selectively transport proteins across the lipid bilayer, and membranes that can respond to light of specific wavelengths and preferentially conduct ions with specific characteristics (charge, size, etc.). One might ask what aspect of the membrane is responsible for having so much potential for life.

It was later revealed that proteins are the components of membranes that can confer this evolutionary flexibility. However, many questions have surrounded how proteins encode function into membranes. One step forward occurred in 1972 when Singer and Nicholson presented the fluid mosaic model[1], which described proteins embedded in a two-dimensional bilayer of lipids. Combined with the work of biochemists such as Linus Pauling who predicted protein secondary structure[2, 3] and the emergence of methods of protein structure determination[4, 5, 6] we now have a better understanding of membrane proteins and how they functionalize membranes. Despite advancements in technology and the wealth of knowledge biochemists have built over the years about life, we still have many unanswered questions surrounding membranes. Some of these questions relate to health and disease[7], fundamental biology such as the origin of life[8], and the functional potential of membranes[9]. If we can understand membranes, perhaps we can provide answers to these questions and build better solutions to address related problems.

One approach to studying membranes is through understanding the inner workings of the proteins responsible for encoding their function. Membrane proteins fold into elaborate and dynamic shapes that are a product of evolution and key to understanding their function. These functions give rise to macroscopic phenomena such as touch, vision, and whole organismal survival.

While progress has been made in understanding these molecular machines through high throughput screens, biochemical reconstitution, and genome sequencing, rapid progress in

structural biology has revolutionized how we understand and probe the world of membrane proteins. Examples include PIEZO[10], TRPV1[11] and opsins[12]; these are proteins implicated in the molecular basis of touch, temperature, and vision. Microbial opsins further illustrate this point. Following their discovery and determined structures[13], scientists have learned to engineer these molecular machines for use in optogenetics, a technique for controlling neuronal electrical activity using light. The techniques of protein structure determination have allowed us to visualize these machines at atomic resolution. X-ray crystallography and early electron microscopy were invaluable. They provided initial glimpses of membrane proteins such as the first structure of bacteriorhodopsin in the 1990s[14] and a subsequent mutant structure of the same protein revealing a mechanism of vectorial proton translocation[15]. Crystallography became the prominent technique for high resolution structures of membrane proteins as evidenced by Rod Mackinon's awarded Nobel prize in 2003 for "the Atomic Basis of Selectivity Ion Conduction in K+ channels", where he succeeded in demonstrating the first structures of a potassium ion channel and the principles that govern how they function[16]. However, crystallography has significant limitations, which relate to the inherent dynamic nature of proteins. Electron microscopy (EM) had its own issues (including those related to radiation damage, image analysis, and sample preparation) but these were tractable and minimized as technology advanced. EM for protein structure determination has thoroughly matured into cryogenic electron microscopy (cryo-EM) and more examples began to emerge displaying the power of the technique[10, 11]. This was evidenced in 2017 by Richard Henderson, Joaquim Frank and Jacques Dubochet being awarded the Nobel prize for their work in developing the technique. Not only was cryo-EM producing more published protein structures than ever before, but targets for cryo-EM were often proteins deemed too dynamic or flexible to be viable options for crystallography. To note, many of these intractable targets were membrane proteins of interest to pharmaceutical companies and researchers studying fundamental biological questions[17, 18].

Cryo-EM, albeit immensely successful, has limitations and associated caveats for understanding and engineering membrane proteins. With low signal to noise, image analysis is the intrinstic a bottleneck to the reconstruction of high-resolution structures. Identifying protein particles, assigning particle view orientations, and accurately aligning particles are particularly computationally challenging.These limitations has been circumvented to some extent with better technology[19] and some benefited from a combined biochemical reconstitution-based approach such as the use of nanodiscs[20]. Nanodiscs are small, disc-shaped lipid bilayers surrounded by amphipathic scaffold proteins which allow for the reconstitution of membrane proteins in a more native environment compared to detergent micelles. The nanodisc approach offers benefits over the traditional method of keeping solubilized membrane proteins in detergent, which can be destabilizing. In contrast, reconstituting proteins back into a lipid environment provides an opportunity for investigating lipid-protein interactions and can overall stabilize the protein and facilitate high-resolution structure determination.

This thesis describes two protein channel structures and how their visualization provided a basis for understanding function and allowed us to probe their inner workings. The second chapter focuses on the Translocase of the Outer Mitochondrial Membrane (TOM) Complex

which is the major pre-protein gate responsible for the transport of 90% of all mitochondrial precursor proteins. A high resolution structure of the TOM complex had been awaited for several decades and revealed the characteristics of the pore that facilitate translocation, the nature of the interactions that connect subunits of the complex together, and the basis for oligomerized states. It also raised questions surrounding how substrates are engaged and relevance of proposed trimeric-"holo" assembly. The third chapter describes the recently discovered novel pump-like channelrhodopsin (PLC) ChRmine. Although ChRmine has been noted as perhaps the most potent opsin for optogenetic applications, our understanding of its unique properties and ability to further engineer these characteristics for optogenetics were limited due to its divergence from canonical channelrhodopsins and the lack of known structure. Shown in this thesis is the first cryo-EM structure of a channelrhodopsin in lipid nanodiscs providing insights into the divergence of ChRmine from early discovered channelrhodopsins. The structure is used to develop new variants of ChRmine with advantageous properties such as accelerated rates of closing. More importantly, this approach reveals a new path to capturing dynamic states of small, light-gated ion channels which is further discussed in the final chapter of this thesis.

# Chapter 2

# Cryo-EM structure of the mitochondrial protein-import channel TOM complex

This chapter presents work described in [21]. Eunyong conceived the project, built the atomic models and supervised the project. Eunyong and I performed experiments, interpreted results and wrote the manuscript.

## 2.1 Abstract

Nearly all mitochondrial proteins are encoded by the nuclear genome and imported into mitochondria after synthesis on cytosolic ribosomes. These precursor proteins are translocated into mitochondria by the TOM complex, a protein-conducting channel in the mitochondrial outer membrane. We have determined high-resolution cryo-EM structures of the core TOM complex from Saccharomyces cerevisiae in dimeric and tetrameric forms. Dimeric TOM consists of two copies each of five proteins arranged in two-fold symmetry: pore-forming β-barrel protein Tom40 and four auxiliary α-helical transmembrane proteins. The pore of each Tom40 has an overall negatively charged inner surface attributed to multiple functionally important acidic patches. The tetrameric complex is essentially a dimer of dimeric TOM, which may be capable of forming higher-order oligomers. Our study reveals the detailed molecular organization of the TOM complex and provides new insights about the mechanism of protein translocation into mitochondria.

## 2.2 Main

Mitochondria are double-membrane-bound organelles that perform oxidative phosphorylation and other essential cellular functions in eukaryotic cells. There are ∼1,000–1,500 mitochondrial proteins, and the vast majority (∼99%) are synthesized by cytosolic ribosomes,

initially as precursor proteins that are then imported into mitochondria[22, 23, 24]. Multiple protein complexes within the organelle mediate membrane translocation and sorting of these precursor polypeptides into four distinct compartments: the outer membrane, the inner membrane, the intermembrane space (IMS), and the matrix. The general import pore in the outer membrane is formed by the TOM complex (translocase of the outer membrane), which is responsible for initial translocation of over 90% of mitochondrial precursor proteins from the cytosol to the IMS.

Studies of the TOM complex of fungal cells have established that it consists of seven transmembrane proteins: Tom40, Tom22, Tom5, Tom6, and Tom7, as well as Tom70 and Tom20 [25, 26]. The first five proteins form a stable complex, referred to as the core TOM complex, and the latter two proteins readily dissociate from the core complex upon isolation in detergent[27, 28]. Various analyses have indicated that the detergent-solubilized TOM complex has an apparent molecular mass of ∼400–600kDa and contains multiple copies of each Tom subunit[27, 28, 29, 30, 31]. The translocation pore through which precursor polypeptides must pass is formed by Tom40[26, 32, 33, 34], a β-barrel protein structurally related to the voltage-dependent anion-selective channel VDAC, a major mitochondrial porin[35, 36]. The other Tom proteins are associated with Tom40 by their single α-helical transmembrane segments (TMs). Although functions of the α-helical Tom subunits are relatively poorly defined, they have been suggested to act as receptors for precursor proteins[37, 38, 39, 40, 41], binding sites for other factors[41, 42], and/or escorts that promote assembly and stability of the TOM complex[27, 31, 43, 44].

Current evidence indicates that translocation is a sequential process in which a precursor protein is first recruited by the cytosolic receptor domains of Tom70, Tom20, and Tom22, then threaded into the pore of Tom40, and finally handed over to the translocase of the inner membrane (TIM) complexes or IMS-resident chaperones (review in [22]). However, the underlying mechanism by which the TOM complex enables these events has been unclear. In particular, how the Tom40 channel interacts with mitochondrial targeting motifs within precursor proteins is poorly understood[32, 45, 46, 47]. The majority of matrix-targeted proteins ( 60–70% of mitochondrial precursor proteins) contain a short N-terminal cleavable sequence, termed the presequence, which typically forms a positively charged amphipathic α-helix. The amphipathic nature of the presequence is likely to be important for interaction with the Tom40 pore for initial threading. Recently, a cryo-EM structure of the dimeric core TOM complex from Neurospora crassa was reported[48], but its relatively low resolution ( 7 Å) precluded building of an atomic model and thus offered only limited insight about the pore structure and the translocation mechanism. Additionally, the oligomeric architecture of the TOM complex is a puzzle. The N. crassa structure represents a dimeric complex in which two identical pores are symmetrically arranged. However, based on previous low-resolution EM and cross-linking analyses, it has been generally thought that the TOM complex is rather dynamic and that the mature form is a trimer[26, 34, 49, 50]. The nature of the different oligomeric states remains unclear.

Here we describe near-atomic resolution structures of the core TOM complex from S. cerevisiae, determined using cryo-EM: a dimeric structure at 3.1-Å resolution and a tetrameric

structure at 4.1-Å resolution. A stable form of the complex is a dimer consisting of two copies each of Tom40, Tom22, Tom5, Tom6, and Tom7 arranged in twofold symmetry. Surface electrostatics calculations show that Tom40 forms a pore with a highly negatively charged surface, which may attract positively charged polypeptides, such as presequences, to initiate translocation. Indeed, neutralization of negatively charged patches in the pore markedly impaired the function of the TOM complex. The tetrameric structure shows that the dimeric TOM complex can further associate into larger oligomers by lateral stacking.

## 2.3 Results

### Cryo-EM analysis of a dimeric TOM complex from yeast

To enable efficient structural analysis, we first developed a new approach to overexpress and purify the S. cerevisiae TOM complex. All Tom subunits, except for the weakly associated Tom70[30, 50], were expressed in yeast cells from an inducible promoter. The complex was then isolated via affinity purification, utilizing His- and Strep-tags attached to Tom22 and Tom40, respectively. The complex was initially extracted with lauryl maltose neopentyl glycol (LMNG) detergent but was exchanged into dodecyl maltoside (DDM) during affinity purification, because free LMNG micelles often interfere with efficient single-particle cryo-EM analysis[51]. The TOM complex purified by this method was eluted in size-exclusion chromatography (SEC) as a largely monodisperse peak containing Tom40 and other Tom subunits but not Tom20 (Fig. 1a,b). The absence of Tom20 in the sample is likely because of its low-affinity association with the core complex[27, 30].

To determine the structure of the TOM complex, we used single-particle cryo-EM analysis (Table 2.1 and Extended Data Fig. 2.7). Two- and three-dimensional classifications of particle images showed that the complex is predominantly a dimer (Extended Data Fig. 2.7a,c), closely resembling the previously reported N. crassa structure[48]. After excluding empty detergent micelle and low-quality particles, 70% of particle images (160,577 out of 243,227) were used for the final 3D reconstruction of the dimeric TOM complex at 3.1-Å resolution with C2 symmetry imposed (Fig. 1c,d and Extended Data Fig. 2.7). Without imposing symmetry, the map was refined to slightly lower resolution (3.2Å) and manifested no noticeable differences from the symmetrically refined reconstruction (cross-correlation=0.99; data not shown), indicating that the dimer is highly symmetric. We note that the sample for this dimeric TOM structure additionally included a synthetic presequence peptide. However, the features of this peptide were not sufficiently resolved in our density map and therefore will not be further discussed. A separate map reconstructed at 3.5-Å resolution from a smaller data set without added presequence peptides showed an essentially identical structure (map cross-correlation=0.98; data not shown). For the sake of more-accurate modeling, we used the 3.1-Å-resolution map in this study.

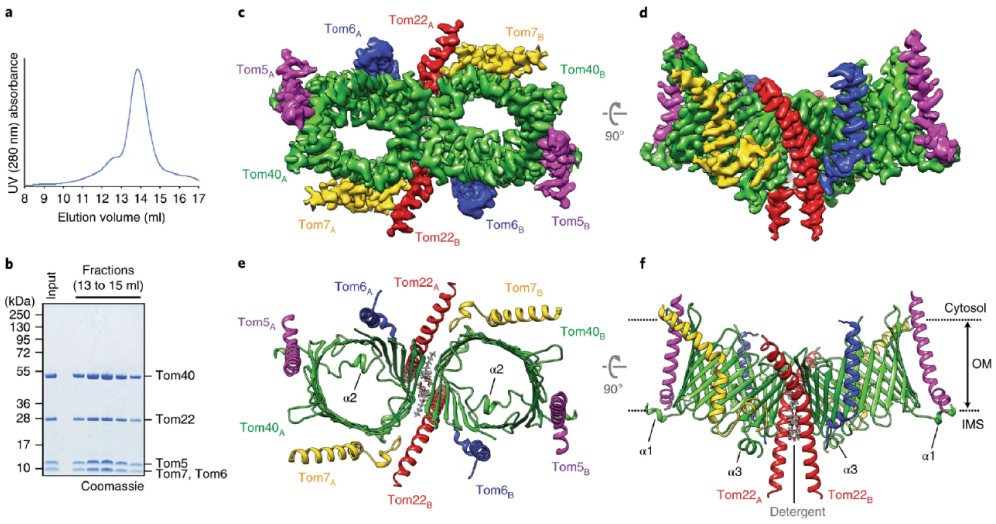


Figure 2.1: **Structure of the dimeric core TOM complex from S. cerevisiae**

**a,** SEC (Superose 6) profile of the affinity purified yeast TOM complex. **b,** Coomassie-stained SDS gel of peak fractions from Superose 6 **(a)**. **c,d,** A 3.1-Å-resolution cryo-EM reconstruction of the dimeric TOM complex. Tom subunits from each asymmetric unit are indicated by subscripts A and B. Shown are views from the cytosol **(c)** and along the membrane plane **(d)**. **e,f,** Atomic model of the TOM complex in ribbon representation. Two DDM detergent molecules between the Tom40 subunits are represented in sticks. Three α-helical segments (α1, α2, and α3) of Tom40 are indicated. Dotted lines (in **f**), approximate outer membrane (OM) boundaries.

## Overall structure of the dimeric TOM complex

The near-atomic resolution density map enabled us to build an accurate de novo model of the TOM complex (Fig. 1e,f). A local resolution estimate indicates that a large portion of the complex, especially the Tom40 subunit, is at 3.0-Å resolution or better (Extended Data Fig. 2.8a). The map resolves not only individual β-strands of Tom40 but also almost all side chains (Extended Data Fig. 2.8c). Distal segments of Tom22 and small Tom subunits, however, remain poorly resolved, likely due to intrinsic flexibility. Notably, our subunit assignment agrees with the previous assignment of the N. crassa structure[48], which was largely based on cross-linking data[34].

Each monomeric unit of the TOM complex contains a single copy of Tom40, Tom22, Tom5, Tom6, and Tom7, with each Tom40 forming a separate pore for polypeptide passage (Fig. 1c–f). The new structure confirms that the Tom40 barrel consists of 19 β-strands (β1–19) arranged in an antiparallel fashion, except for β1 and β19, which are parallel. Tom40

also has three short α-helical segments, α1 and α2 in the N-terminal segment and α3 near the C terminus. α1 resides on the IMS side, lying flat on the membrane surface as an amphipathic helix. Following α1, a segment containing α2 spans the interior of the Tom40 barrel, as noted previously[34, 48]. The structural features of 19 β-strands and an N-terminal segment within the pore closely resemble the structure of VDAC, despite low ( 15%) sequence identity[52] (Extended Data Fig. 2.8d). Although not resolved at high resolution, the 14-amino-acid-long C-terminal tail of Tom40 following α3 seems directed from IMS into the pore of Tom40 and loosely associated with a hydrophobic patch (referred to as HP3; described below) of the pore lining (Extended Data Fig. 2.8e). Interestingly, the same feature has also been noted with the N. crassa structure, despite poor sequence conservation at this region among different species. It is possible that the C-terminal tail may act as an autoinhibitory element that would release from the pore upon insertion of a precursor protein.

At the dimeric interface, the two Tom40 subunits directly contact each other on the cytosolic side by hydrophobic side chains in β1-β19-β18 (Fig. 1c–f and Extended Data Fig. 2.9a–c). However, a gap opens toward the IMS between the two Tom40 barrels, which are tilted away from each other by 40° (Fig. 1g). In our structure, this gap is filled by two DDM molecules and two Tom22 TMs wedged into the interface (Fig. 1c–f and Fig. 2.9c–e). In the native membrane, a phospholipid would occupy this gap in place of detergent, with its head group phosphate positioned to interact with the highly conserved Arg330 of Tom40 (Extended Data Fig. 2.8c).

Tom22 contains an unusually long ( 45 amino acid) α-helix, the middle portion (roughly positions 100–118) of which spans the membrane (Fig. 1f). The helix is longbow-shaped, because of a kink formed by Pro112 (Fig. 2a), a residue that has been reported to be important for mitochondrial targeting of Tom22 and stability of the TOM complex[34, 53]. The helix extends at least 22Å out from the membrane into the IMS, which may function as a binding site for presequences[54] or the TIM complex[55]. On the opposite cytosolic side, the Tom22 helix becomes amphipathic, lying flat on the membrane surface. Preceding the helix, the cytosolic segment (positions 1–88) of Tom22 is invisible, likely due to its flexibility. The function of this region has been suggested to be a docking site for Tom20 and Tom70 [56, 57] and/or a presequence receptor[40, 58]. The mechanism for the latter is unclear, because the domain appears to be directed away from the Tom40 pores. The other three small Tom subunits, Tom5, Tom6, and Tom7, are peripherally bound to Tom40 by interactions with different regions of Tom40 (Fig. 1c–f).

## Interactions between β-barrel and α-helical Tom subunits

The TOM complex represents a rare example of a complex consisting of both β-barrel and α-helical types of integral membrane proteins; thus, our structure offers a unique opportunity to examine interactions between the two types of membrane proteins. The structure shows that association between Tom40 and the α-helical Tom subunits is mainly mediated by hydrophobic interactions in conjunction with high surface complementarity between transmembrane domains (Fig. 2a–d and Extended Data Fig. 2.9d–h). Additionally, several polar

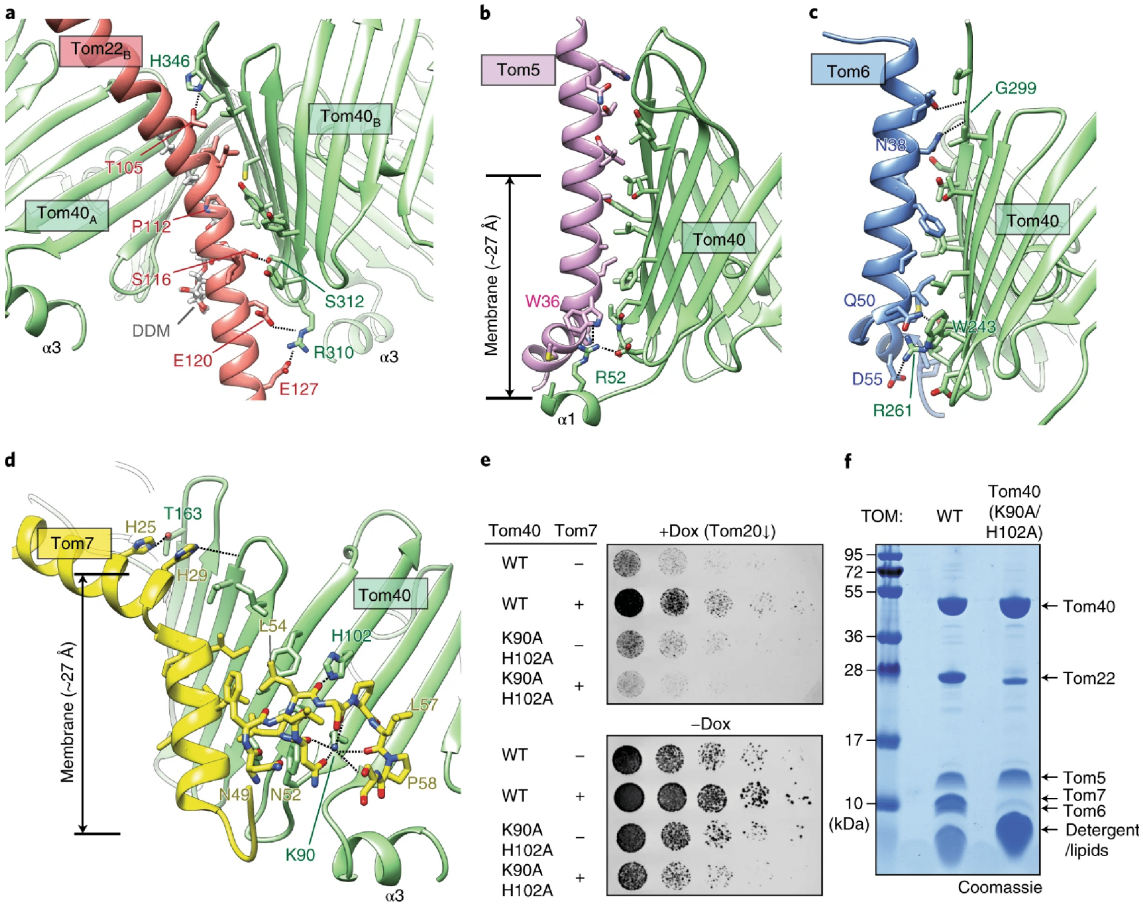


Figure 2.2: **Intersubunit contacts between Tom40 and α-helical Tom subunits**

**a,** Interactions between Tom40 and Tom22 within the same monomeric unit. The polar interactions are indicated by black dotted lines. Side views are shown. **b–d,** Interactions of Tom40 with Tom5 **(b)**, Tom6 **(c)**, and Tom7 **(d)**. Note that in **d,** N49–L54 of Tom7 is an α-helix. **e,** Growth complementation of a tom7Δ yeast strain with a Tom7-expressing plasmid (tested in a YPD medium containing 2% glucose). Where indicated, endogenous Tom20 was depleted with doxycycline (+Dox), and Tom40 contains K90A and H102A mutations. **f,** Purified TOM complex containing K90A and H102A mutations on Tom40. The experiments in **e** and **f** were repeated twice with similar results. WT, wild type.

interactions were noticed near the membrane boundaries. Conservation of these polar interactions across fungal species suggests that they may play an important role in increasing specificity and affinity of subunit interactions (Table 2.2). Indeed, mutation of Arg261 or Trp243 of Tom40, which interact with Tom6 in our structure, has been shown to decrease the stability of TOM, similar to a Tom6 knockout[59, 60].

Our structure also reveals an unusual topology of Tom7, in which its partially unstructured, hook-shaped C-terminal segment spans the IMS leaflet of the outer membrane (Fig. 2d). An unstructured polypeptide in the lipid membrane is rare, because unpaired hydrogen-bond donors and acceptors of the peptide backbone would be energetically unfavorable. In the TOM complex, this issue seems to be overcome by hydrogen bonding between backbone carbonyl oxygen atoms of Tom7 and lipid-facing side chain nitrogen atoms of conserved Lys90 and His102 of Tom40. To test the importance of this interaction, we performed a complementation assay. Previously, it has been shown that deletion of both Tom7 and Tom20 exhibits synthetic lethality[61]. Consistent with this, in the wild-type Tom40 background, exogenously expressed Tom7 rescued growth defects caused by chromosomal deletion of Tom7 and depletion of Tom20 (Fig. 2e). By contrast, with K90A H102A mutant Tom40, no such rescue was seen, likely because Tom7 cannot bind to the mutant Tom40. Instead, expression of Tom7 showed a dominant-negative phenotype in the mutant Tom40 background. Although the exact mechanism is unclear, this finding suggests that unassociated Tom7 exerts a toxic effect. To further verify a loss of the physical interaction between Tom7 and K90A H102A Tom40, we performed purification of the K90A H102A mutant TOM complex using the same procedure as for the wild-type complex. Consistent with the growth complementation experiments, Tom7 was not co-isolated (Fig. 2f). Interestingly, the amounts of copurified Tom6 and Tom22 were also reduced, and the complex seemed to be largely dissociated into monomers (Extended Data Fig. 2.9i), suggesting additional defects in assembly or stability of the complex[44].

## Pore structure of Tom40 and implications for protein translocation mechanism

To gain insight into the protein translocation mechanism by TOM, we examined the translocation pathway in Tom40. Although the Tom40 β-barrel has relatively large ( 30Å by 25Å) oval-shaped openings on both the cytosolic and IMS sides, the pore is substantially constricted ( 19Å by 13Å) halfway across the membrane by the α2 segment (Fig. 1c,e). Still, the pore would snugly fit one or perhaps two α-helices along the vertical translocation axis. Given the considerable contacts with β7–β19 of Tom40, the α2 segment appears to be a stationary feature of the pore. We also speculate that the Tom40 barrel would be unlikely to open laterally toward the lipid phase, as proposed for BamA and Sam50, which mediate membrane insertion of β-barrel proteins[62, 63, 64]. The only separable β-strand pair, β1-β19 would be energetically costly to open, because it is sealed by approximately ten hydrogen bonds and buried at the dimerization interface. Together, these findings suggest that Tom40

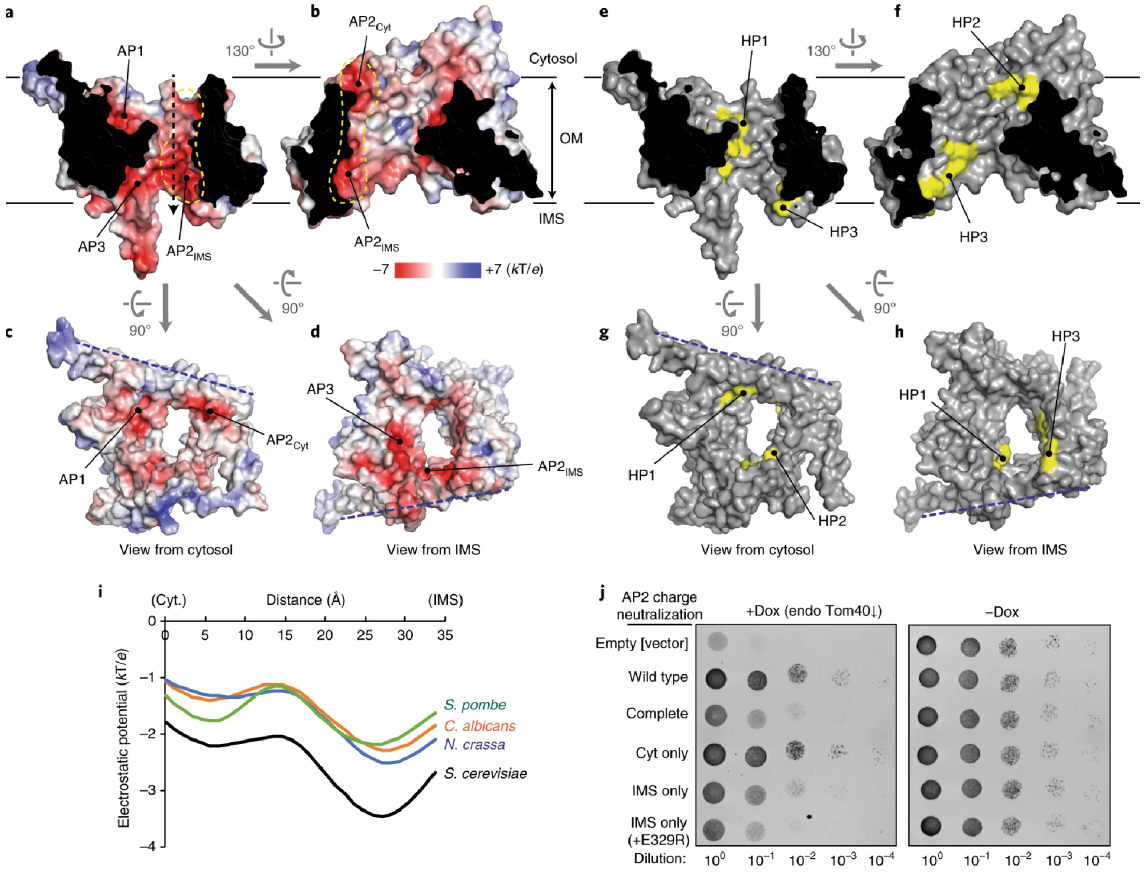


Figure 2.3: **Pore architecture of Tom40**

**a-d**, Surface electrostatics of the TOM complex shown as a heat map on a solvent-accessible surface representation. For simplicity, only one monomeric unit is shown (the dimer interface is indicated by a blue dashed line in **c and d**). Shown are cutaway side views **(a,b)** and views from the cytosol **(c)** and IMS **(d)**. Acidic patches are referred to as AP1, AP2 (also outlined by yellow dash line), and AP3. Black dashed arrow, pore axis. **e**–**h**, As in **a**–**d**, but showing hydrophobic patches (HPs) in yellow. **i,** Electrostatic potential along the pore axis (black dashed arrow in **a**). After calculating electrostatic potential using homology models for indicated species, values along the pore axis were extracted and plotted. **j,** Yeast cells expressing an indicated Tom40 mutant from a CEN plasmid were serially diluted and spotted on SC (-Leu) plates containing 2% glucose. In these strains, the presence of doxycycline (+Dox) represses expression of chromosomal Tom40. IMS only, D87N E329N E360N; Cyt only, D132N D134N; complete, a combination of IMS and Cyt. The experiment in **j** was repeated three times with similar results.

is a static pore for polypeptide passage.

To understand how Tom40 may interact with translocating polypeptides, we evaluated surface properties of its pore (Fig. 3 and Extended Data Fig. 2.10). Surface electrostatic analysis indicates that the pore has an overall negative potential, mainly attributed to multiple acidic patches (referred to as APs 1–3) on the pore lining (Fig. 3a–d and Extended Data Fig. 2.10a–d). A similar, overall negative potential is anticipated for Tom40 from other fungal species, based on homology modeling (Fig. 3i and Extended Data Fig. 2.11). This explains why Tom40 was selective for cations when ion conduction was measured by electrophysiology[32, 26]. The negative electrostatic potential likely promotes protein translocation by attracting positively charged amino acids in polypeptides, such as inner membrane proteins and presequences of matrix-targeted preproteins, both of which are often basic[65]. Interestingly, the potential seems to be more negative toward the IMS side (Fig. 3i), which may promote polypeptide movement toward IMS. The pore-lining surfaces also contain hydrophobic patches (HPs; Fig. 3e–h), which may interact with precursor proteins to facilitate translocation.

To test the functional importance of these patches, we examined cell growth defects associated with their mutations on the basis that Tom40's protein translocation function is essential for cell viability (Fig. 3j and Extended Data Fig. 2.10h). When we mutated the conserved and most prominent acidic patch AP2 by replacing five glutamate and aspartate residues with asparagine ('complete' mutant), a substantial growth retardation was observed. The defect seems largely due to the charge neutralization of AP2 residues on the IMS side ($AP2_{IMS}$). When an additional positive charge (E329R) was introduced at $AP2_{IMS}$, growth was further reduced. Complete neutralization of AP3, which is localized near the IMS opening next to $AP2_{IMS}$, also led to similar growth defects. Together, these results suggest the importance of a negative electric potential at the IMS side of the pore. We also observed impaired growth phenotypes when we mutated either HP2 or HP3 (Extended Data Fig. 2.10h). Incomplete growth inhibition by the mutation of individual patches might be due to their functional redundancy.

## Assessment of oligomeric structure of TOM

The oligomeric nature of the TOM complex is a long-standing puzzle. Our structure, as well as the N. crassa structure[48], suggests that the dimer is a stable configuration and likely to be translocation competent. However, previous low-resolution EM and cross-linking studies have proposed that the mature or holo-TOM complex is a trimer[26, 34, 49, 50]. It remains unclear whether and how the TOM complex switches between different oligomeric states.

During our purification experiments, we made a surprising observation that under a more-gentle detergent condition, the TOM complex can be purified as a larger species than a dimer (Fig. 4 and Extended Data Fig. 2.12). Although exchange of LMNG into DDM during affinity purification resulted almost exclusively in dimers that migrated as an 500-kDa species (Fig. 4a), delayed exchange into DDM at the last SEC step produced an additional peak appearing at a higher molecular size ( 1 MDa) (Fig. 4b). When DDM was substituted

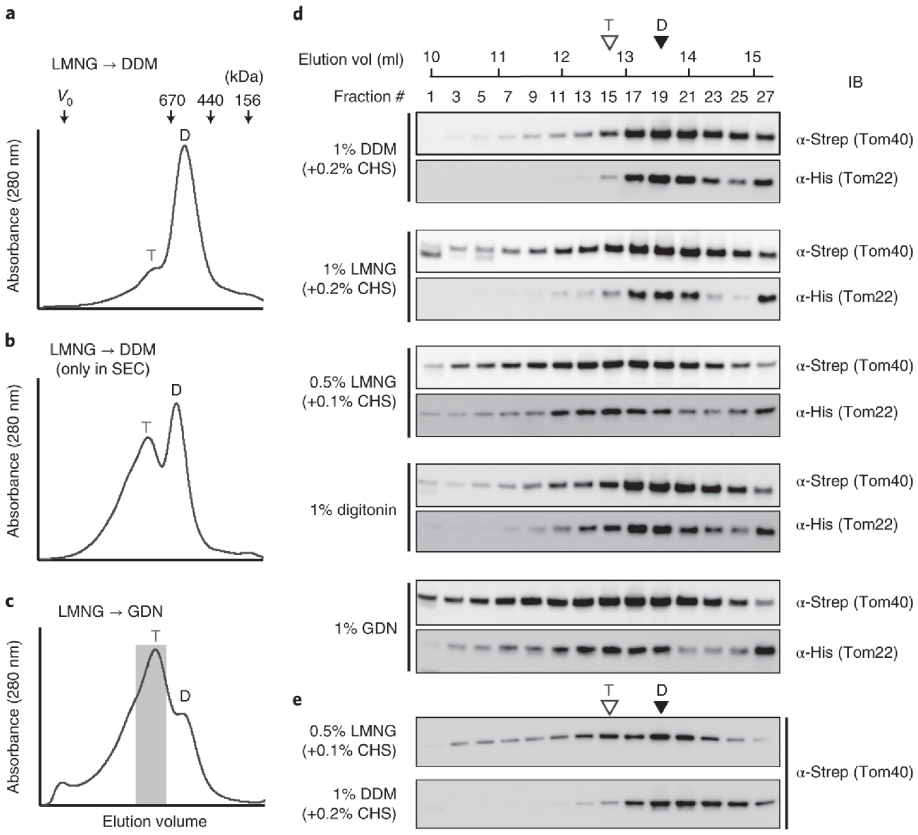


Figure 2.4: **Analysis of oligomeric states of the TOM complex**

**a–c,** SEC elution profiles of the TOM complex in different detergent conditions (details in **Extended Data Fig. 2.12a–e**). V0, void volume. In **c,** fractions in gray were used for cryo-EM analysis in **Fig. 5** and **Extended Data Fig. 2.12**. Two distinct peaks are marked as T and D. **d,** Cells overexpressing the TOM complex (cultured in a medium containing 2% ethanol and 3% glycerol as the carbon source) were lysed in indicated detergent condition, and the lysates were injected onto a Superose 6 column. The fractions were analyzed via SDS-PAGE and immunoblotting (IB). The column was equilibrated with buffer containing the same detergent used for membrane solubilization at a low concentration, as described in Methods. Approximate peak positions are marked with T and D based on the UV absorbance profiles shown in **a–c** (**Extended Data Fig. 2.12b–e**). Note that the anti-Strep-tag antibody appears to have a substantially lower detection limit (higher sensitivity) than the anti-His-tag antibody. **e,** As in **d,** but using solubilized mitochondrial fractions with endogenous (chromosomal) Tom40 tagged with a Strep-tag. The experiments in **d** and **e** were repeated twice with similar results. Source data for **d** and **e** are available with the paper online.

by glyco-diosgenin (GDN), a digitonin-like detergent that is generally considered to be more gentle than DDM, the complex eluted mostly in the 1-MDa peak (Fig. 4c). The sample also seemed to contain even larger species, as some Tom proteins eluted earlier. Importantly, SDS-PAGE analysis of peak fractions showed no changes in subunit composition (Extended Data Fig. 2.12f), indicating that the two peaks simply differ in their oligomeric states. Similar high-molecular-weight species of the TOM complex were observed when crude cell or mitochondrial extracts were analyzed via SEC under gentle-detergent conditions (Fig. 4d,e).

Because many previous studies evaluating the TOM complex assembly have used blue native PAGE (BN-PAGE) analysis[27, 30, 31, 40, 43, 56], we also subjected the extracts to BN-PAGE in addition to SEC analysis (Extended Data Fig. 2.12g). This comparison, together with the new cryo-EM structure, suggest that the previously reported 400-kDa band in BN-PAGE corresponds to the dimeric complex. A discrepancy between the nominal size of dimeric TOM ( 160kDa) and its apparent size (400–500kDa) in SEC and BN-PAGE seems to originate from the complex's flat structure with hollow pores and a large detergent micelle around it. Unlike SEC analysis, BN-PAGE did not show prominent higher-oligomer species, perhaps due to dissociation into dimers in the harsh conditions of BN-PAGE[28].

## Cryo-EM structure of the tetrameric TOM complex

To understand how the larger species are organized, we analyzed the 1-MDa peak fractions using cryo-EM (Fig. 5a–c and Extended Data Fig. 2.13). As expected from the SEC analysis, particles on micrographs were much larger than those seen with the dimer sample (Extended Data Fig. 2.13b). Two- and three-dimensional classifications of particle images showed a striking tetrameric arrangement of the pores (Extended Data Fig. 2.13a,c). We also noticed that micrographs often showed particles larger than the dimensions of the tetramer, indicating that the sample included oligomers larger than tetramers (Extended Data Fig. 2.13g), consistent with the SEC profile. Notably, a minor 3D class showed three pores (Extended Data Fig. 2.13a; class 3), reminiscent of trimers seen in low-resolution EM studies[26, 49, 50]. This 'trimer' class appears to be similar to the tetramer class, lacking one monomeric unit.

The tetramer structure determined at 4.1-Å resolution reveals that it is essentially a dimer of two dimeric TOM complexes (referred to as A–B and C–D), which are arranged in a staggered parallel fashion that would allow further assembly into larger oligomers (Fig. 5a–b and Extended Data Fig. 2.13). There are only a few structural differences between the dimeric complex and dimers in the tetrameric complex, as two copies of atomic models for the dimer could be fitted into the EM map, essentially as rigid bodies. The dimer-to-dimer contact is contributed by two Tom6 subunits (Tom6B and Tom6C) as well as Tom22 and Tom5 (referred to as Tom22B and Tom5C) (Fig. 5a,b and Extended Data Fig. 2.14). Particularly, each of two Tom6 subunits at the interface interact with Tom40 from the other dimeric complex, where its flexible N-terminal segment (residues 1–25) appears to be directed to the Tom40's barrel interior next to β11 and near HP2 (Fig. 5b and Extended Data Fig.

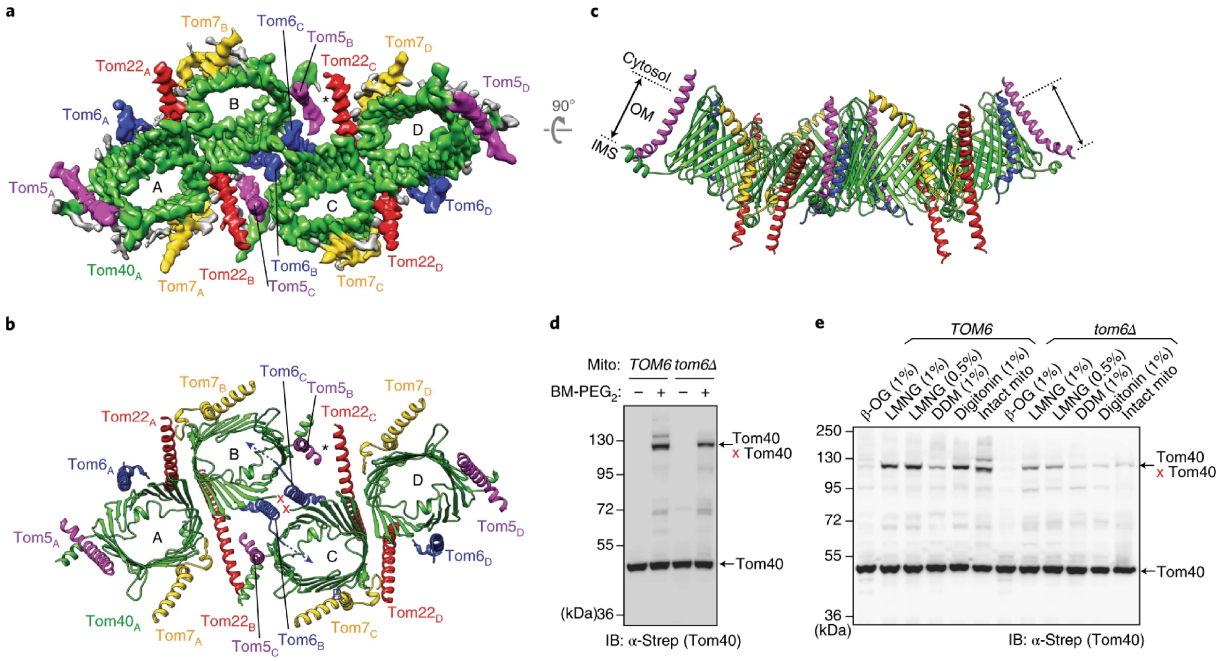


Figure 2.5: **Cryo-EM structure of the tetrameric TOM complex**

**a–c,** Cryo-EM reconstruction **(a)** and atomic model **(b,c)** of the tetrameric TOM complex. Four monomeric units are indicated by A, B, C, and D. Shown are views from the cytosol **(a,b)** and along the membrane plane **(c)**. Asterisk, gap between Tom5B and Tom22C. Red 'X' shows the approximate position of the cysteine introduced for cross-linking experiments (287C). **d,** Cross-linking between two Tom40 copies (at the endogenous level) in isolated mitochondria by bismaleimido-diethyleneglycol (BM-PEG2). Where indicated, chromosomal Tom6 was deleted (tom6Δ). Cells were grown in a YPD medium. **e,** As in **d,** but cross-linking was performed after solubilization of mitochondrial membranes with indicated detergents. The experiments in **d** and **e** were repeated twice with similar results. Source data for **d** and **e** are available with the paper online.

2.15a). This interaction readily explains the result of previous in organello cross-linking experiments that Tom6 can cross-link to two opposite sides of the Tom40 barrel[34, 66]. Lastly, it is noteworthy that the tetramer is not completely symmetric, such that a gap exists at one of the two Tom22-Tom5 contacts (Fig. 5a,b and Extended Data Fig. 2.14a–c). Furthermore, there is a considerable gap ( 7Å in width) along the dimer-dimer interface at the IMS leaflet of the membrane (Extended Data Fig. 2.14i), creating a concave curvature to the cytosolic side (Fig. 5c). In the cryo-EM map, the gaps are filled by weak density features, which should be detergent and/or lipid molecules (Extended Data Fig. 2.14h and data not shown). It is possible that in the native membrane, the gap is closed such that the complex lies relatively flat in the membrane. Looking from the IMS, protein surfaces in the interface are roughly complementary between the two TOM dimers to accommodate such a closure (Extended Data Fig. 2.14i). Nevertheless, the relatively loose interface explains why tetramers easily dissociate into dimers by excess detergent and suggests that the TOM oligomers undergo a dynamic equilibrium in the native membrane.

Examination of mitochondrial detergent extracts by SEC showed the presence of higher-order TOM oligomers at the endogenous level (Fig. 4e). To test whether their oligomeric configuration is consistent with that of the tetramer structure, we performed cross-linking experiments by introducing a cysteine to the L14–15 loop of Tom40. Although the L14–15 loop is not fully resolved in our cryo-EM maps, it is located near the dimer-dimer interface of the tetramer such that the distance between two sulfhydryl groups of introduced cysteines may become close enough (¡14Å) to be cross-linked by a bismaleimide-PEG2 (Fig. 5b and Extended Data Fig. 2.15a); however, cross-linking would not be achievable between the two Tom40 molecules within a dimeric complex (the distance is  60Å). Indeed, Tom40 could be efficiently cross-linked via cysteines at position 287 after being extracted with LMNG or digitonin, as well as in intact mitochondria, where tetramers are expected (Fig. 5d,e and Extended Data Fig. 2.15b–d). By contrast, little or no cross-linking was obtained when the complex was extracted with DDM or octyl glucoside, conditions in which the complex largely dissociates into dimers or monomers[33]. While these results do not address the previously proposed trimeric TOM complex, as a high-resolution structure of such a configuration is not available, they are consistent with tetrameric and higher oligomeric configurations observed in our structural analysis. Lastly, we tested effects of Tom6 deletion on the oligomerization of the TOM complex. Although lack of Tom6 did not prevent formation of higher oligomers in both cross-linking and SEC experiments (Figs. 4e, 5d,e, and Extended Data Fig. 2.15e), it substantially decreased the cross-linking efficiency, suggesting that Tom6 promotes formation of oligomers in the configuration revealed by our structure.

## 2.4 Discussion

Our high-resolution structures of the yeast TOM complex offer new mechanistic insights into how Tom40 mediates translocation of precursor proteins. While precursor polypeptides are first recognized by the cytosolic domains of Tom20 and Tom70, they need to be threaded into

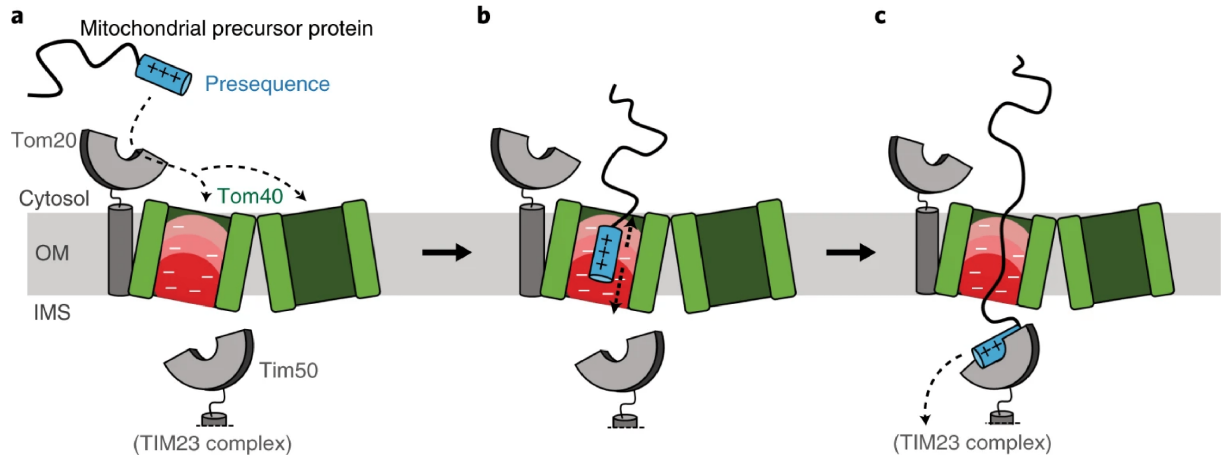


Figure 2.6: **Model for presequence engagement with the TOM complex**

**a,** The presequence is initially recruited to Tom20 by hydrophobic interactions with the cytosolic domain of Tom20. The presequence is attracted into one of two pores of the TOM complex by the negative electrostatic potential of the pores. **b,** The presequence inserts into the pore close to the IMS by electrostatic interactions. Thermal motions would allow the presequence to move vertically along the pore. **c,** Once exposed to the IMS, the presequence binds to the soluble domain of Tim50, which would further hand it over to the TIM23 complex.

the pore of Tom40. Because there is no external energy input (that is, ATP or membrane potential) involved, this early step of translocation must be driven solely by the affinity of precursor proteins toward the pore interior. Our structural and functional analyses suggest that electrostatic interactions between the Tom40 pore and the precursor protein play an important role in this process (Fig. 6). Particularly, in the cases of presequence-containing proteins, the positively charged presequence may be first attracted into the overall negatively charged Tom40 pore and then drawn toward the IMS by interaction with acidic patches on the IMS side. This mode of interaction may provide not only a driving force for presequence insertion into the pore but also an additional 'filter' for increased targeting specificity, as initial recognition of presequences by Tom20 is mediated by hydrophobic interactions[39]. It remains to be elucidated how the presequence moves out from the pore interior into the IMS. This would likely require thermal (Brownian) motion of the precursor protein, as well as some movement of the C-terminal tail of Tom40. Once exposed in the IMS, the presequence might be captured by Tim50 of the TIM23 complex, which has been shown to interact with presequences[67], and would thus be prevented from backsliding.

A highly unexpected finding was that the TOM complex can form a tetramer and larger oligomers. While the dimeric form is likely a functional unit, its clustering into larger oligomers might fine-tune the protein import activity. Unlike previous low-resolution EM

studies[26, 49, 50], we did not observe a symmetrical trimer class throughout our cryo-EM analyses. It is possible that the difference might be because our samples lacked Tom20, which has been proposed to mediate trimerization of Tom40[49], and, therefore, our study does not directly argue against the trimer model. In light of high-resolution structures, future studies will be necessary to re-evaluate the trimeric configuration and understand how Tom20 would mediate formation of trimers despite its seemingly weak association to the complex. It also remains to be elucidated what functional state the tetrameric and larger assemblies represent. One possibility is that the TOM complexes cluster into larger assemblies to increase import efficiency, potentially advantageous for the cotranslational import, where multiple precursor molecules would be produced on a polysome[68]. Lastly, our study shows that formation of tetramers and higher-order oligomers is facilitated by Tom6, which coincides well with its proposed function in stabilizing the TOM complex[27, 69]. It has been shown previously that phosphorylation of Tom6's N-terminal tail (Ser16) increases the steady-state levels of Tom6 and the TOM complex, as well as overall mitochondrial protein import[69]. Such modifications on Tom subunits could regulate the dynamics of TOM oligomerization. Our work provides a framework for further investigations to understand the structure, dynamics, and functions of the high-order TOM complex assemblies we have discovered.

## 2.5 Methods

### Constructions of plasmid and yeast strains

To generate an S. cerevisiae strain overexpressing the TOM complex components from an inducible GAL1 promoter, we used the Yeast Tool Kit (YTK) and Golden Gate assembly[70]. We first amplified coding sequences (CDS) for Tom40, Tom22, Tom20, Tom 7, Tom6, and Tom5 via PCR using genomic DNA of S. cerevisiae BY4741 as a template and cloned them individually into the pYTK1 entry plasmid. To enable affinity purification of the Tom complex, a Strep-tag (GGWSHPQFEK) and a His-tag (GGHHHHHHHH) were introduced before the stop codons of Tom40 and Tom22, respectively. The cloned Tom subunits were combined with YTK parts to generate individual expression cassettes, each containing the GAL1 promoter (YTK30), CDS of a Tom subunit, and the ENO1 terminator (YTK61). In the case of the purification in Fig. 2f and Extended Data Fig. 3i, Tom40 CDS included K90A and H102A mutations, which were introduced via site-directed mutagenesis. The six Tom expression cassettes were assembled into a single multigene plasmid, concatenating them in the order of Tom40-Tom22-Tom20-Tom7-Tom6-Tom5. The plasmid also contained a nourseothricin resistance marker (YTK78) for selection and URA3 homology arms (YTK92 and YTK86) for chromosomal integration. The resulting assembly was introduced to the yMLT62 yeast strain (a gift from J. Thorner) by a standard lithium acetate transformation method after linearizing the plasmid with the NotI endonuclease. The colonies were selected on a YPD agar plate containing 10µg/ml nourseothricin, and chromosomal integration was confirmed by PCR. The yMLT62 strain (BY4741 leu2::pACT1-GEV::HIS3MX) contains the

chimeric transcriptional activator Gal4dbd.ER.VP16 (GEV;[71]) integrated to the LEU2 locus, which induces the transcription by the GAL1 promoter upon addition of β-estradiol to the growth medium.

To generate plasmids expressing Tom40 under the native promoter, we first used PCR to amplify the endogenous Tom40 gene region (of BY4741), including the 329-bp segment upstream of the start codon and the 381-bp segment downstream of the stop codon. This fragment was then inserted into a homemade yeast CEN/ARS plasmid constructed with YTK (used parts: pYTK84, pYTK8, pYTK47, pYTK73, pYTK75, and pYTK81). The plasmid contained a LEU2 marker for selection. For immunodetection, we attached a Strep-tag to the C-terminus of Tom40 using PCR (the resulting plasmid is referred to as pe112-Tom40Strep). Where indicated, other mutations were also introduced by PCR. In some experiments, plasmids (pe115-Tom40Strep or pe115-Tom40His) contained a nourseothricin resistance marker (from pYTK78) instead of the LEU2 marker. Where a cysteine-free or single-cysteine mutant of Tom40 was used, the endogenous cysteines of Tom40 were mutated to alanine or methionine (C165A/C326A/C341A/C355M).

To replace chromosomal wild-type Tom40 with a cysteine-free or single-cysteine version (287C, 291C, or 293C), we used homologous recombination using a PCR fragment amplified from the Tom40-expressing plasmids (same as pe112-Tom40Strep except that these constructs have a shorter downstream segment (172bp instead of 381bp) following the stop codon of Tom40Strep). The DNA segment containing a 5' upstream region of Tom40, the CDS of Tom40, a 3' downstream region, and the LEU2 marker was amplified via PCR with a forward primer (CAGGGACATGGGTAAGAACTTG) and a reverse primer (gaccattgtgaaagtaaggacaaggatatgagacgtatcataactataaacaaggaattcCTGCCTATTTAACGCCAAC; lower case indicates the homologous region to the chromosomal locus). The PCR products were purified and introduced to the yeast strain BY4741 by means of lithium acetate transformation. Colonies were isolated from a synthetic complete agar medium lacking leucine (SC(-Leu)). Colonies with correct double-crossover recombination were screened by PCR of genomic DNA and Sanger sequencing.

A strain expressing Tom20 under a tetracycline-repressible promoter (replacing the native promoter of Tom20) was generated on the R1158 strain (Dharmacon) background, as described previously[72]. Deletion of chromosomal Tom7 (tom7Δ::HIS3) was carried out by transformation of a PCR product generated from YTK76 (HIS3 marker) as a template, a forward primer (agaaactagttccctcttatctctcaatatttgccaaaattagcttttaacaaataaaccCTGTGGATAACCGTAGTCG), and a reverse primer (taattcaaaattggaaatatgggcttcctctctcacccaagttgtatcgaactgatgtttGGGCGTTTTTTATTGGTC). Deletion of chromosomal Tom6 (tom6Δ::URA3) was performed similarly using pYTK76 (URA3 marker), a forward primer (ccatgtcctgtaggcttctcaagagaacaaaaacaaaacacagacaaaataattgaaaaCTGTGGATAACCGTAGTCG), and a reverse primer (caaaaaccaatatatatacaggtaagtgaaaaaatctcaactatacaagaaccaaccccGGGCGTTTTTTATTGGTC). Deletion was confirmed via PCR of genomic DNA. To introduce a K90A H102A mutation to chromosomal TOM40, we used a marker-free CRISPR-Cas9 approach[70], using a sgRNA targeting the 57–64th codons of Tom40 CDS (G CTG GTC AAT CCC GGT ACC GTG G) and a repair DNA containing the K90A H102A mutation

and CRISPR-resistant synonymous codons (G TTA GTT AAC CCT GGT ACT GTC G), which were amplified using a pe112-Tom40 template and primers (CAGGGACATGGGTAAGAACTTG; TAAACCTAAAGCTAATTGAGGAG). The successful mutation was confirmed by Sanger sequencing.

## Purification of the TOM complex

Yeast cells were grown in YPEG medium (1% yeast extract, 2% peptone, 2% ethanol and 3% glycerol) in shaker flasks at 30°C. Upon reaching an optical density (OD600) of 1.4–2, cells were induced with 50nM β-estradiol. After 9–10h of induction, cells were harvested via centrifugation at 5,000r.p.m. Cell pellets were flash frozen in liquid nitrogen and stored in -80°C until use. The TOM complex was purified via tandem affinity purification using His- and Strep-tags, as summarized in Extended Data Fig. 2.12a. Cells were first lysed by cryomilling at the liquid nitrogen temperature and resuspended in buffer (3× cell pellet volume) containing 50mM Tris-HCl, pH 8, 200mM NaCl, 10% glycerol, 20mM imidazole, and protease inhibitors (5µg/ml aprotinin, 5µg/ml leupeptin, 1µg/ml pepstatin A, and 1mM PMSF). Then, one cell pellet volume of 5% lauryl maltose neopentyl glycol (LMNG; Anatrace) and 1% cholesteryl hemisuccinate (CHS; Anatrace) was added to solubilize membranes. After a 3-h incubation at 4°C, the lysate was clarified by ultracentrifugation (Beckman Coulter rotor Type 45Ti) at 125,000g for 1h. The lysate was incubated by gentle rotation with HisPur cobalt resin (Life Technologies) for 3h at 4°C. The beads were then packed in a gravity column and washed with approximately 10 column volumes (CVs) of buffer containing 50mM Tris-HCl, pH 8, 200mM NaCl, 0.02% LMNG, 0.004% CHS, 20mM imidazole, and 10% glycerol. Resin was further washed with an additional 10 CVs of buffer containing 40mM imidazole and eluted with approximately 6 CVs of buffer containing 180mM imidazole. The eluate was then mixed with Strep-Tactin Sepharose (IBA Lifesciences) for 14h at 4°C. The beads were packed in a gravity column and washed with approximately 10 CVs of buffer containing 20mM Tris-HCl, pH 7.5, 100mM NaCl, 0.03% dodecyl-β-maltoside (DDM; Anatrace), 0.006% CHS, and 1mM dithiothreitol (DTT). In the case of purification of the tetrameric TOM complex, 0.02% glyco-diosgenin (GDN; Anatrace) was used instead of DDM and CHS. The TOM complex was eluted with buffer containing 3mM D-desthiobiotin and concentrated using AmiconUltra (100kDa cut-off, Millipore). The complex was further purified by SEC using a Superose 6 Increase 10/300 GL column (GE Lifesciences) equilibrated with 20mM Tris-HCl, pH 7.5, 100mM NaCl, 1mM DTT, and 0.03% DDM, 0.006% CHS (for the dimeric TOM complex) or 0.02% GDN (for the tetrameric TOM complex). Peak fractions were pooled, concentrated to 3.5–5mg/ml using AmiconUltra (100kDa cut-off; Millipore), and used to prepare cryo-EM grids. For experiments described in Extended Data Fig. 2.12b–f, essentially the same procedure was employed but with modified detergent conditions as indicated. Purification of the TOM complex containing K90A H102A-mutant Tom40 (Fig. 2f and Extended Data Fig. 3i) were carried out with the same procedure used for purification of the wild-type dimeric TOM complex.

## Cryo-EM specimen preparation and data acquisition

Immediately before preparing cryo-EM grids, 3mM fluorinated Fos-Choline-8 (FFC8; Anatrace) was added to the purified TOM sample. We note that the addition of 3mM FFC8 did not cause any changes in the SEC profiles of either the dimeric or tetrameric TOM complex, even after a prolonged ( 6h) incubation. To prepare cryo-EM grids, 3µl of the sample was applied to a glow-discharged Quantifoil holey carbon grid (R 1.2/1.3 Au, 400 mesh; Quantifoil). Glow discharge was carried out for 20s in 75% argon and 25% oxygen using a Gatan Solarus plasma cleaner or in air using a PELCO easiGlow glow discharge cleaner. The grid was blotted with Whatman No. 1 filter papers for 3s at 4°C and 100% humidity and plunge frozen in liquid-nitrogen-cooled liquid ethane using Vitrobot Mark IV (FEI).

A summary of image acquisition parameters is shown in Table 1. The data sets were collected on a Titan Krios electron microscope (FEI) equipped with a K2 Summit direct electron detector (Gatan) and a GIF Quantum image filter (Gatan). The microscope was operated at an acceleration voltage of 300kV. Dose-fractionated images were collected in the super-resolution mode with a physical pixel size of 1.15Å and a GIF slit width of 20eV, using SerialEM software[73]. The dose rate was 1.22 electrons/Å$^2$/frame, with a frame rate of 0.2s. For the dimeric complex, the total accumulated dose was 61 electrons/Å$^2$ (50 frames), and for the tetrameric TOM complex, it was 48.8 electrons/Å$^2$ (40 frames).

## Single-particle image analysis of the dimeric TOM complex

A summary of the single-particle analysis procedure is described in Extended Data Fig. 1a. Briefly, RELION3[74] was used for preprocessing of movies, particle picking, and Bayesian particle polishing, and then cryoSPARC v2 [75] was used for ab initio reconstruction, 3D classification, and the final 3D reconstruction. First, the movies were imported to RELION3 and corrected for motion using MotionCor2 with 5-by-5 tiling [76]. During this step, micrographs were 2×-pixel-binned (resulting in a pixel size of 1.15Å). Micrographs that were not suitable for image analysis (for example, micrographs containing crystalline ice or displaying a large drift) were removed by manual inspection. Defocus parameters were estimated using CTFFIND4 [77]. Template-based automatic particle picking was performed in RELION3 (460,148 particles from 1,587 movies). The particle templates were generated by 2D classification from Laplacian autopicking on a subset of the data. The particles were extracted from micrographs with a box size of 256 pixels. Reference-free 2D classification (Extended Data Fig. 1c) was performed to remove empty detergent micelles and obvious non-protein particle artifacts, resulting in 290,793 particles. The initial 3D model was generated by cryoSPARC (ab initio reconstruction). The first 3D refinement was carried out by RELION3 using a low-pass-filtered initial model and 290,793 particle images, yielding a 3.8-Å-resolution reconstruction. The particle images were subjected to one round of CTF refinement and Bayesian particle polishing in RELION3. These particles were subjected to a second 3D refinement, which yielded a 3.6-Å-resolution reconstruction. Then, another round of CTF refinement and particle polishing was performed. The resulting polished particles were imported to

cryoSPARC v2 for the subsequent process, as described below.

The imported particles were subjected to 2D classification in cryoSPARC to further discard artifacts and low-quality particles. The resulting 243,227 particles were used to generate four ab initio 3D reconstructions, followed by heterogeneous refinement (3D classification). A total of 179,232 (74%) particles converged to one class (class 3; Extended Data Fig. 1a) leading to a high-resolution reconstruction of the dimeric TOM complex; whereas two low-resolution classes (classes 1 and 2) appeared to have only a single pore, likely corresponding to dissociated monomers. After a second round of 3D classification to further remove low-quality particles, 160,577 particles from class 3 were refined by non-uniform refinement with C2 symmetry imposed, yielding the final map at 3.06-Å resolution (based on gold-standard Fourier shell correlation (FSC) and the 0.143 cut-off criterion; Extended Data Fig. 1e). Local resolution was estimated by cryoSPARC using default parameters (Extended Data Fig. 2a).

## Single-particle image analysis of the tetrameric TOM complex

Summaries of single-particle image analysis for the tetrameric TOM complexes are shown in Extended Data Fig. 2.12a. Essentially, motion correction, defocus estimation, particle picking, and particle extraction were performed using Warp [78], and the remaining downstream refinement process was carried out using cryoSPARC v2. Movies were corrected for motion with 8-by-8 tiling, and defocus parameters were estimated with 5-by-5 tiling. Original super-resolution micrographs were 2×-pixel-binned. Particles were automatically picked by Warp. Micrographs were manually inspected to remove unsuitable micrographs. Particle images were extracted with a box size of 400 pixels from dose-weighted frames 1–36 (skipping the last 4 frames). Particle images were then imported to cryoSPARC and subjected to one round of reference-free 2D classification to remove empty micelles. Ab initio reconstruction was performed to generate four (for tetrameric TOM) initial 3D models, which were then subjected to a heterogeneous refinement. Approximately, 80% of particle images converged into two nearly identical classes (classes 1 and 2) showing high-resolution features. These particle images were used for the final 3D reconstructions by non-uniform refinement in cryoSPARC, yielding maps at resolutions of 4.1Å. No symmetry (C1) was imposed, because the complex was found to be not completely symmetric (imposition of C2 symmetry led to artificial distortion of some density features). Local resolution was also estimated by cryoSPARC using default parameters.

## Atomic model building

A summary of model refinement and validation is shown in Table 1. The atomic model for dimeric TOM was built de novo using Coot[79] and the summed map. In addition to proteins, we also modeled several hydrophobic tails of detergent or lipid (we used DDM as a model). The model was refined in real space using Phenix [80] and the summed map with the refinement resolution limit set to 3.1Å. Different weights were tested using half maps to

check whether the used Phenix refinement protocol shows overfitting to the map (Extended Data Fig. 2b; FSCwork vs. FSCfree). To this end, we chose a weight of 2, which did not separate FSCwork and FSCfree. We also used restraints for secondary structure. The following segments were not modeled because of poor or invisible density features: N–48, 277–294, and 374–387(C) of Tom40, N–85 and 136–152(C) of Tom22, N–12 and N–26 and 48–50 (C) of Tom6, and N–10 of Tom7.

To build a model for the tetrameric TOM complex, two dimer models were fit into the tetramer map using UCSF Chimera. A few additional residues (α1 of Tom40, 81–89 of Tom22, and 25–26 of Tom6) were built using Coot, because the tetramer map shows extra densities for these segments. Additionally, we modeled 1,2-dimyristoyl-rac-glycero-3-phosphocholine (DMPC) into the density at the Tom40-Tom40 dimer interface (instead of DDM, as in the dimeric TOM complex). The model was then refined against the tetramer map in essentially the same way as that described for the dimeric TOM complex. Structural validation was done using MolProbity (ref. 61).

Protein electrostatics were calculated using PDB2PQR and the Adaptive Poisson-Boltzmann Solver (www.poissonboltzmann.org;[81]) with monovalent mobile ions (0.1M for both cation and anion) included in parameters. UCSF Chimera and PyMOL (Schrödinger) were used to prepare structural figures in the paper.

## Yeast growth assays

To test functional complementation by mutant Tom40, we used a yeast strain (TH_7610; Dharmacon) from the Yeast Tet-Promoters Hughes Collection, in which the original Tom40 promoter was replaced by a tetracycline promoter (tetprom). The cells were transformed with a CEN/ARS plasmid (pe112-Tom40Strep) constitutively expressing wild-type or mutant Tom40Strep under the endogenous promoter and selected on agar plates of a synthetic complete medium containing 2% glucose and lacking leucine (SC(-Leu)). After a 3-d incubation at 30°C, colonies were isolated. Cells were grown in 3ml of SC(-Leu) at 30°C until OD600 reached 0.7–1.5, pelleted, and resuspended in fresh medium at OD600 of 1. After tenfold serial dilution, 10µl was spotted on SC(-Leu) agar plates. Where indicated, 15µg/ml doxycycline was included in the medium to repress endogenous Tom40 expression. Plates were incubated at 30°C for 2–2.5 d before imaging. To test expression of the Tom40 mutants in cells, an equal number (2 ODs) of cells were collected from cultures in SC(-Leu) medium, and proteins were extracted by heating in NaOH/SDS solution. The samples were analyzed by SDS-PAGE and immunoblotting with anti-Strep (Genscript; A01732) and anti-PGK1 (a gift from J. Thorner) antibodies. Standard enhanced chemiluminescence reagents and a Fujifilm LAS-3000 Imager were used for detection.

For the complementation experiment in Fig. 2e, the yeast strain (R1158 tetprom-TOM20::KanMX tom7Δ::HIS3) was transformed with pe115-Tom7, which expresses wild-type Tom7 from the native promoter (the cloned region includes from 262-bp upstream to 209-bp downstream of the Tom7 CDS) or an empty pe115 vector. The transformants were selected on YPD (1% yeast extract, 2% peptone, 2% glucose) agar supplemented with

100µg/ml nourseothricin. After growth in YPD with 100µg/ml nourseothricin, cultures were diluted to OD600 of 0.1 and further diluted fivefold in serial, and then 10 µl was spotted on YPD/nourseothricin agar plates, which were incubated at 30°C for 2 d before imaging. Where indicated, 10µg/ml doxycycline was included in the medium to deplete Tom20.

## Size-exclusion chromatography and blue native PAGE analysis of extracts

Yeast cells were grown in YPEG medium and induced by β-estradiol, as previously stated. Cells from 10-ml induced culture were pelleted, washed in distilled water, frozen in liquid nitrogen and stored at -80°C until use. Pelleted cells ( 100mg) were resuspended in 400µl of lysis buffer containing 50mM Tris-HCl, pH 7.5, 200mM NaCl, 1mM EDTA, 2mM DTT, and protease inhibitors. Cells were lysed by beating with pre-chilled glass beads (two cycles of 1.5-min beating and 1-min rest). Beads were removed, and the lysate was mixed with detergent (from a 5% stock solution), as indicated. After solubilizing membranes for 1h at 4°C, samples were clarified for 1h at 13,300r.p.m. and 4°C. One-hundred microliters of the clarified sample was injected into a Superose 6 column equilibrated with 20mM Tris-HCl, pH 7.5, 150mM NaCl, 1mM EDTA, 1mM DTT, and a low concentration of detergent used for lysis (that is, 0.03% DDM, 0.006% CHS; 0.02% LMNG, 0.004% CHS; 0.02% GDN; or 0.08% digitonin). Fractions were collected and analyzed by SDS-PAGE and immunoblotting analyses. For immunoblotting, anti-Strep-tag and anti-His-tag (Life Technologies; MA1-21315) monoclonal antibodies were used.

Samples for BN-PAGE were prepared essentially the same way but with a minor modification. The lysis buffer contained 50mM Tris pH 7.5, 50mM NaCl, 10% glycerol, 1mM DTT, and protease inhibitors. Detergent-solubilized lysates were clarified by ultracentrifugation for 30min at 250,000g (Beckman TLA-100 rotor) and 4°C. Coomassie Blue G-250 (prepared as a 5% stock in 0.5M 6-aminohexanoic acid; 1/4 amount of added detergent by weight) was added to the lysate. BN-PAGE was performed using a 4–16% Novex Native PAGE gel (Life Technologies) according to the manufacturer's instructions.

Where crude mitochondria fractions were used instead of whole cell lysates for SEC analysis (Fig. 4e and Extended Data Fig. 2.15e), detergents were added directly to mitochondria (400µg in 100µL of 10mM MOPS pH 7.2 and 250mM sucrose) for 1.5h on ice with intermittent mixing. After clarification by centrifugation, the sample was injected into a Superose 6 column, and fractions were analyzed as described above.

## Tom40-Tom40 crosslinking

A total of 0.2mM BM-PEG2 was added to 50µg of crude mitochondria in 50µl of 10mM MOPS, pH 7.2, and 250mM sucrose for 4 min at 23°C (20-min incubation was used for experiments in Extended Data Fig. 2.15b,c). Where detergent extracts were used, mitochondria were first solubilized on ice for 1.5h with indicated detergent (when LMNG or DDM was used, 0.2× CHS was supplemented) before adding bismaleimido-diethyleneglycol

(BM-PEG2; Thermo Pierce). Reactions were quenched with the addition of 50mM DTT (or 20mM NEM in Extended Data Fig. 2.15b,c) on ice for 20min. Proteins were precipitated with 10% TCA, washed with cold acetone, and resuspended in SDS sample buffer prior to SDS-PAGE and immunoblotting. For the immunoprecipitation experiment in Extended Data Fig. 2.15c, mitochondria were pelleted and solubilized in buffer containing 20mM Tris-HCl, pH 7.5, 150mM NaCl, 100mM β-OG for 1h on ice after cross-linking with BM-PEG2. The extract was clarified and incubated first with 2µg anti-Strep antibody (or no antibody in mock) for 2h at 4°C and additionally with 25µl of Protein A beads (Thermo/Pierce) for 2h. The beads were washed with the solubilization buffer containing 50mM β-OG, and bound proteins were eluted with SDS sample buffer. The samples were analyzed by SDS-PAGE and immunoblotting using anti-His antibody conjugated to horse radish peroxidase (Proteintech; HRP-66005). For cross-linking after SEC (Extended Data Fig. 2.15d), 850µg of mitochondria was solubilized in 100µl buffer containing 10mM MOPS, pH 7.2, 250mM sucrose, 0.5% LMNG and 0.1% CHS for 1.5h on ice. The clarified extract was then injected into a Superose 6 column. Fractions were incubated with 0.2mM BM-PEG2, quenched with 50mM DTT and analyzed by SDS-PAGE and immunoblotting.

## 2.6 Data availability

The cryo-EM density maps and atomic model are available through EM DataBank (EMD-20728, EMD-20729) and Protein Data Bank (PDB 6UCU, PDB 6UCV), respectively. Source data for Figs. 4d,e and 5d,e are available with the paper online.

## 2.7 Acknowledgements

We thank D. Toso for help with electron microscope operation and J. Thorner for yeast strains and antibodies. We thank J. Thorner, J. Hurley, S. Brohawn, and S. Itskanov for critical reading of the manuscript. This work was funded by UC Berkeley (E.P. and J. T.), Vallee Scholars Program (E.P.), and NSF Graduate Research Fellowship Program (K.T.; DGE-1752814).

## 2.8 Author information

### Authors and Affiliations

Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA

Kyle Tucker & Eunyong Park

California Institute for Quantitative Biosciences, University of California, Berkeley, Berkeley, CA, USA

Eunyong Park

## Contributions

E.P. conceived the project. K.T. and E.P. performed experiments. E.P. built the atomic models. K.T. and E.P. interpreted results and wrote the manuscript. E.P. supervised the project.

## Corresponding author

Correspondence to Eunyong Park.

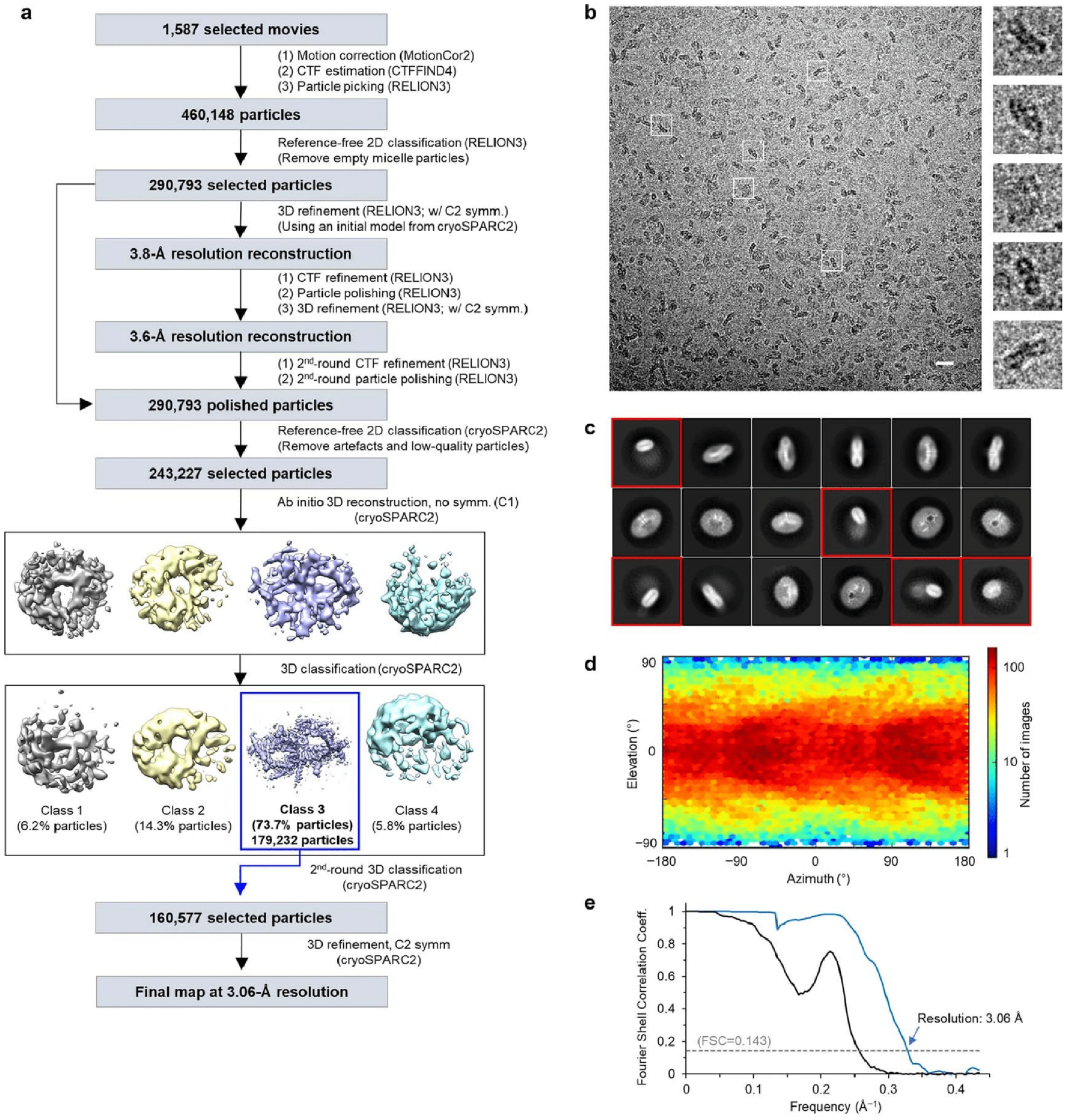


Figure 2.7: **Single-particle cryo-EM analysis of the dimeric core TOM complex**

**a,** Summary of single-particle image analysis procedure. **b,** A representative motion-corrected micrograph. Scale bar, 20 nm. Right panels show magnified images of selected particles outlined with white squares. The particle image size is 209 Å (width) by 209 Å (height). **c,** Representative class averages from 2D classification by RELION3. The box dimensions are 297 Å (width) by 297 Å (height). Classes in red boxes are likely empty micelles and thus excluded in subsequent analysis. **d,** Heat map showing particle orientation distribution (produced in the final 3D reconstruction by cryoSPARC2). **e,** Fourier shell correlation (FSC) of two independently refined half maps. Blue line, corrected masked FSC. Solid black line, unmasked FSC.

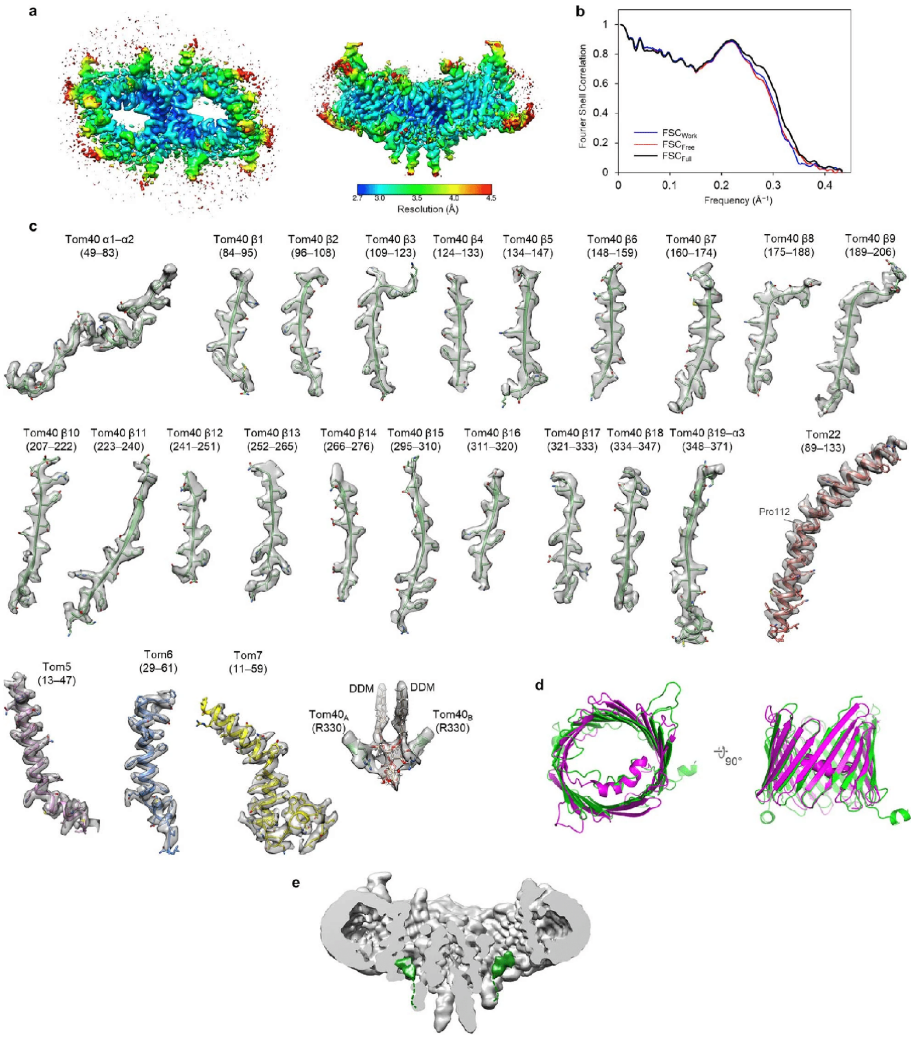


Figure 2.8: **Cryo-EM map and atomic model quality of the dimeric TOM complex**

**a,** Local resolution represented by a heat map on the density contour (unsharpened, summed map). **b,** FSC between the EM map and the atomic model. Blue curve, FSCwork (FSC between half map 1 and a model refined against half map 1). Red curve, FSCfree (FSC between half map 2 and the model refined against half map 1). Black curve, FSCfull (FSC between the combined map and the final atomic model refined against the combined map. All refinements were performed by Phenix with the same weight. **c,** Examples of the density map and the atomic model for indicated segments. Numbers in the brackets indicate ranges of amino acid residues shown. **d,** Structural comparison of Tom40 and VDAC. Structures of Tom40 (this study; green) and murine VDAC (PDB 3EMN; magenta) are superimposed. Left, view from cytosol. Right, side view. **e,** Density features (green) for the C-terminal tails of Tom40 are shown in a 5-Å low-pass-filtered map. Shown is vertical cross-section along the Tom40 pores. A weak connection (not shown) between the density in green and α3 of Tom40 is indicated by a dashed line.

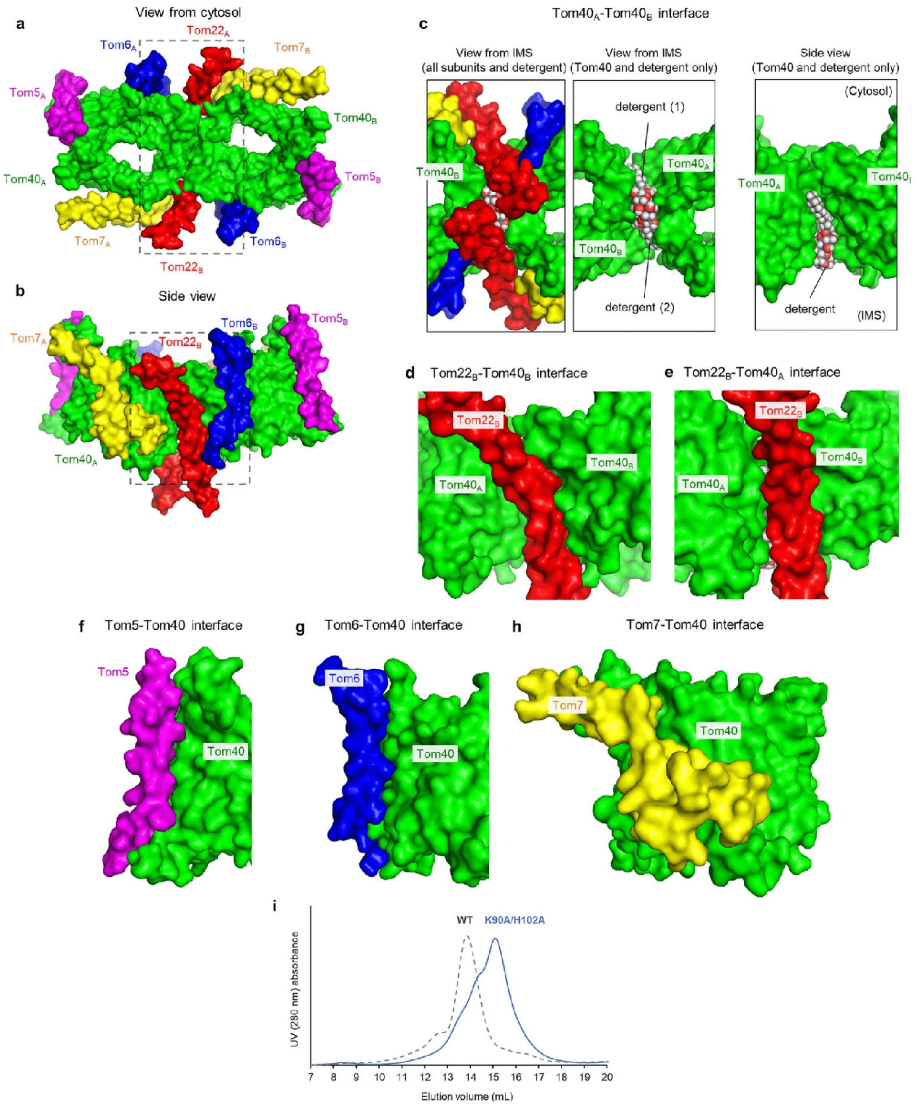


Figure 2.9: **Surface complementarity of Tom subunits at interfaces and purification of the TOM complex with a K90A H102A mutation**

**a, b,** Overview of the dimeric TOM complex in solvent-accessible surface representation. Shown are views from the cytosol **(a)** and along the membrane plane **(b)**. The color scheme is the same as in **Fig. 1**. The regions marked by a dashed line are magnified in **c** (with a 180° rotation) and **d,** respectively. **c,** Interface between the two Tom40 subunits. Left, a view from IMS (showing all subunits). Middle, as in the left panel but showing only Tom40 and DDM detergent molecules. Right, as in the middle panel but showing a side view. **d,e,** Side views showing the Tom40 and Tom22 interfaces within the same asymmetric unit **(d)** and between the two asymmetric units **(e)**. f–h, Side view showing interfaces between Tom40 and other small Tom subunits. The viewing angles are the same as in Fig. **2b–d,** respectively. **i,** Superose 6 SEC profile of the mutant TOM complex with K90A H102A Tom40 (solid blue) purified as the wild-type complex (dashed gray line; also **see Fig. 1a**).

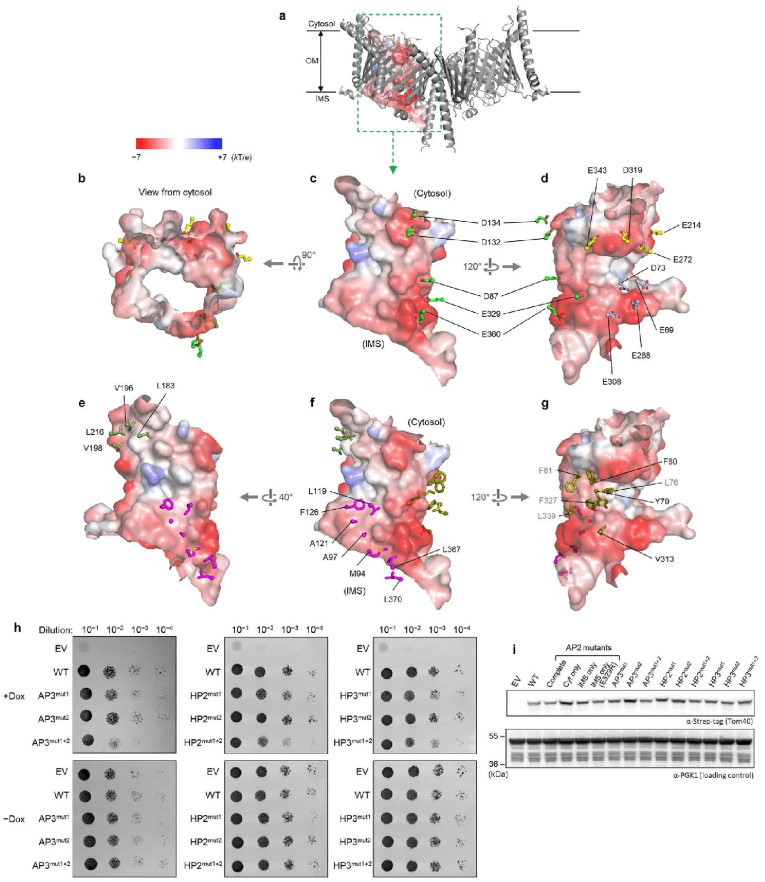


Figure 2.10: **Acidic and hydrophobic patches on the Tom40 pore surface**

**a,** Overview (side view) of the dimeric TOM complex (gray ribbons) and the Tom40 pore cavity (surface representation; shown for only one Tom40 subunit). b–d, Surface electrostatics is shown as a heat map overlaid on the pore cavity shown in surface representation. Side chains of acidic amino acids are shown in stick representation (AP1, AP2, and AP3 are in yellow, green, and pale blue, respectively). In c, only AP2 side chains are shown for clarity. e–g, As in b–d, but side chains of hydrophobic patches are shown in stick representation (HP1, HP2, are HP3 are in olive, green, and magenta, respectively). Note that some hydrophobic side chains in HP1 (labeled in gray; F81, F327, and L76) are only partially exposed, as they are involved in interactions between the α2 segment and the β sheets. h, As in Fig. 3j, but with mutants of AP3, HP2, and HP3. AP3mut1= E268N E308N; AP3mut2= E69N D73N; HP2mut1= L183S L216S; HP2mut2= V196N V198S; HP3mut1= L119S A121N F126N; HP3mut2= M94N A97S. Dox, doxycycline. i, Expression of Tom40 pore mutants (contains a C-terminal Strep-tag) was examined by SDS-PAGE and immunoblotting analyses of whole-cell lysates. PGK1, loading controls. Source data for i are available with the paper online. The experiments in h and i were repeated at least twice with similar results.

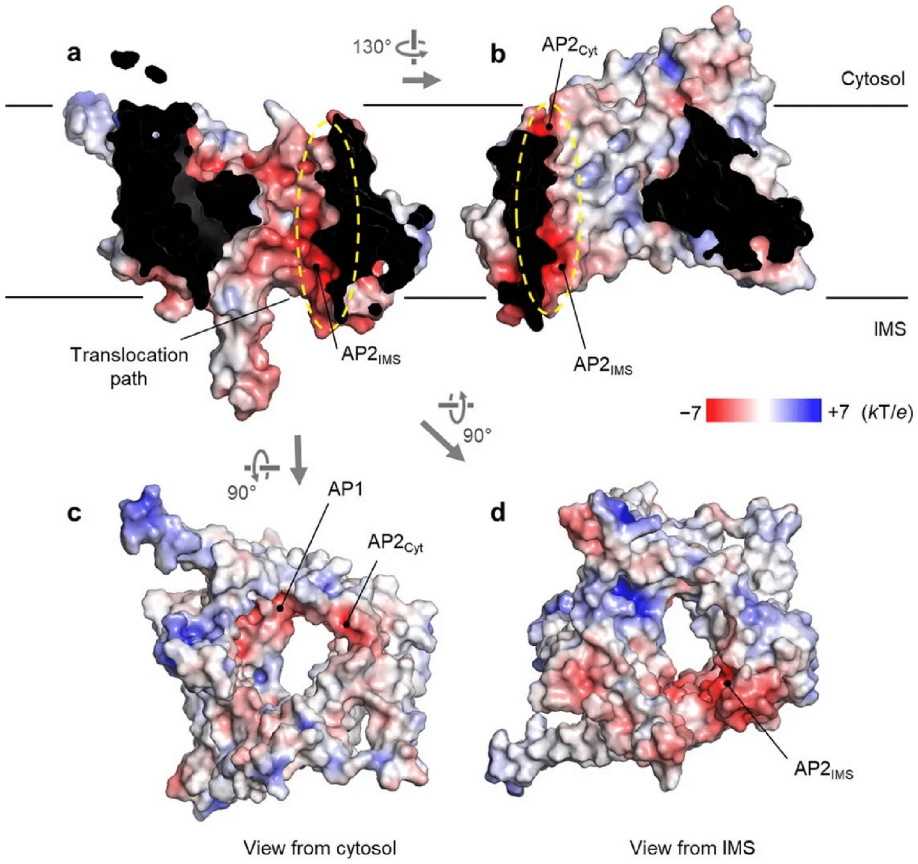


Figure 2.11: **Homology modeled pore architecture of N. crassa Tom40**

As in Fig. 3 a–d, but with N. crassa TOM complex. An N. crassa homology model was generated by SWISS-MODEL using the S. cerevisiae structure as a template, and electrostatic potential was calculated by Adaptive Poisson-Boltzmann Solver (APBS). The dashed yellow line indicates AP2. Note that unlike the S. cerevisiae TOM complex AP3 is not prominent in N. crassa.

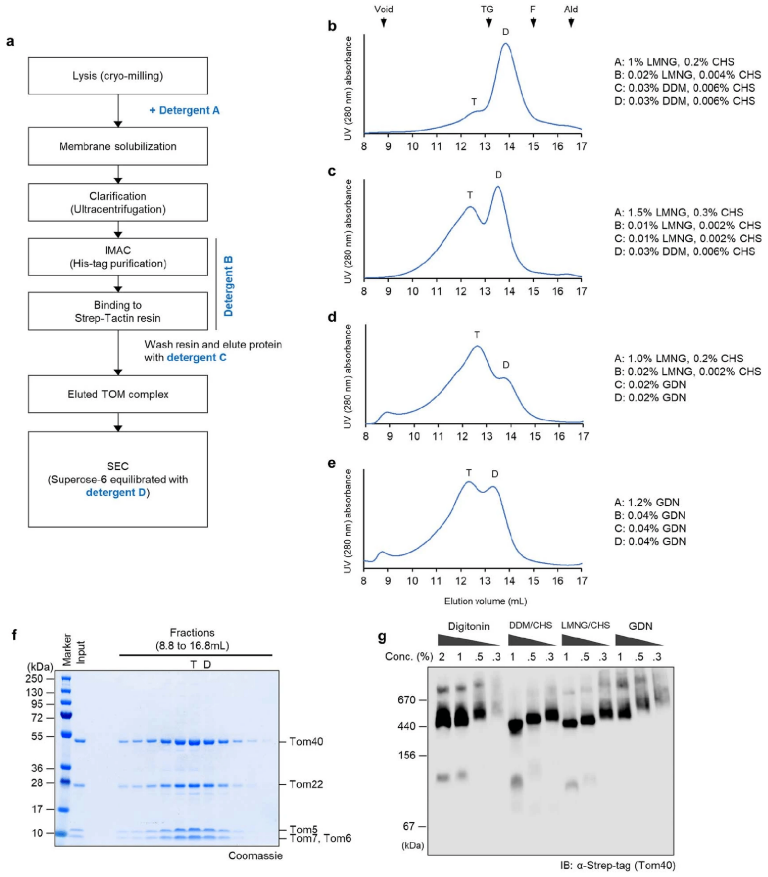


Figure 2.12: **Effects of detergent on the oligomeric state of the TOM complex**

**a,** Schematic diagram of the TOM complex purification procedure. Different detergent conditions (indicated by blue texts) were tested (specific conditions in **b**–**e**). **b**–**e,** Detailed SEC profiles of the purified TOM complex purified under different detergent conditions. "D" indicates the dimer peak, and "T" indicates the tetramer peak. Positions of the void peak (void) and peaks of molecular weight standards are indicated by arrowheads. TG, thyroglobulin (670 kDa); F, ferritin (440 kDa); ald, aldolase (156 kDa). Note that **b**–**d** is the same as in **Fig. 4a**–**c,** and b is the same experiment shown in **Fig. 1a. f,** SDS-PAGE analysis of peak fractions from the SEC purification shown in **c**. The peak positions are marked with "T" and "D". The SDS gel was stained by Coomassie. **g,** Crude lysates prepared from cells overexpressing the TOM complex were solubilized with indicated detergent and subjected to BN-PAGE, followed by immunoblotting using an anti-Strep-tag antibody (detecting Tom40-Strep). A gradual decrease of mobility of the TOM complex accompanied by lowered detergent concentrations is likely due to an increased detergent micelle size. Source data for g are available with the paper online. The experiment in g was repeated twice with similar results.

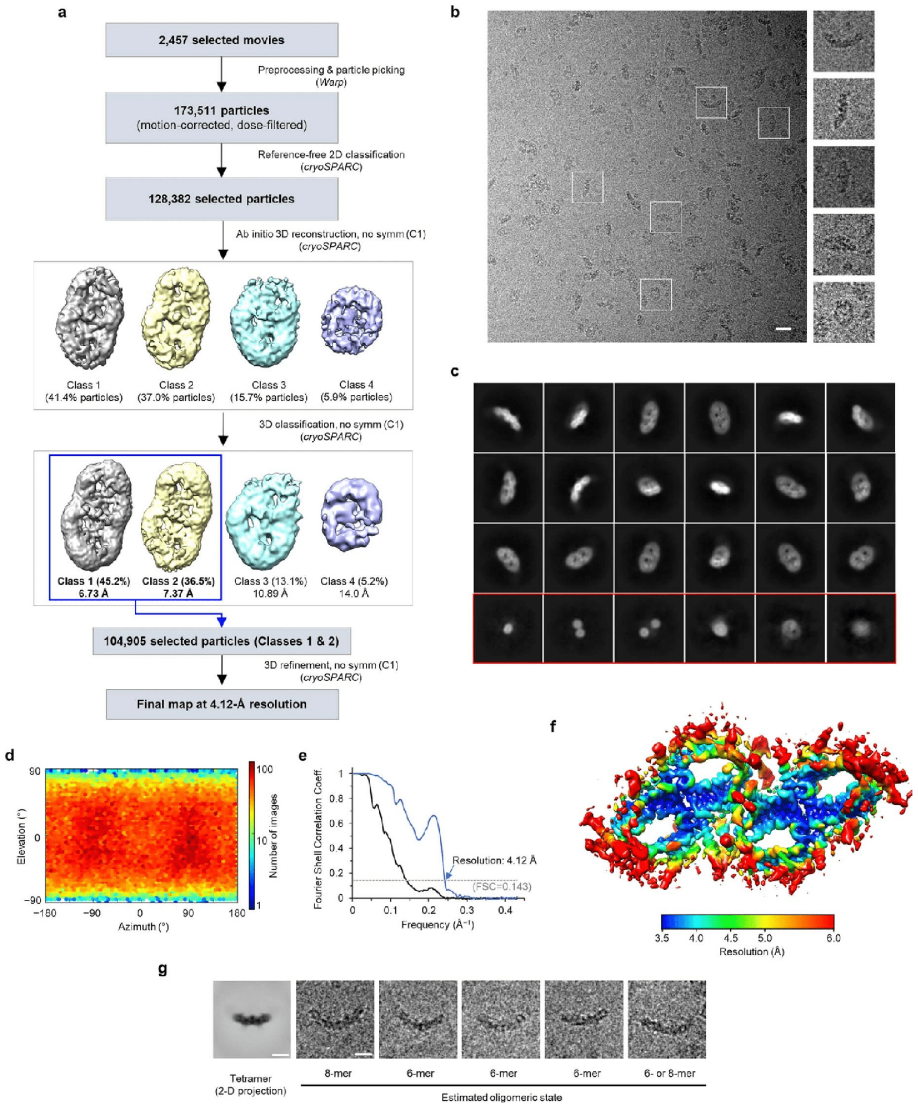


Figure 2.13: **Cryo-EM analysis of the tetrameric complex**

**a,** Summary of single-particle image analysis procedure. **b,** A representative micrograph. Scale bar, 20 nm. The dimensions of magnified images are 414 Å (width) by 414 Å (height). **c,** Examples of selected 2D class averages. The box dimensions are 460 Å (width) by 460 Å (height). **d,** Particle orientation distribution. **e,** Fourier shell correlation (FSC). Blue line, corrected masked FSC. Solid black line, unmasked FSC. **f,** Local resolution map. **g,** Example images of particles larger than the tetramer. The leftmost image shows a 2D projection (side view with the longest width) of the 3D reconstruction of the tetrameric TOM complex. The other images show examples of large particles on micrographs. Estimated oligomeric states are indicated. Scale bar, 100 Å.

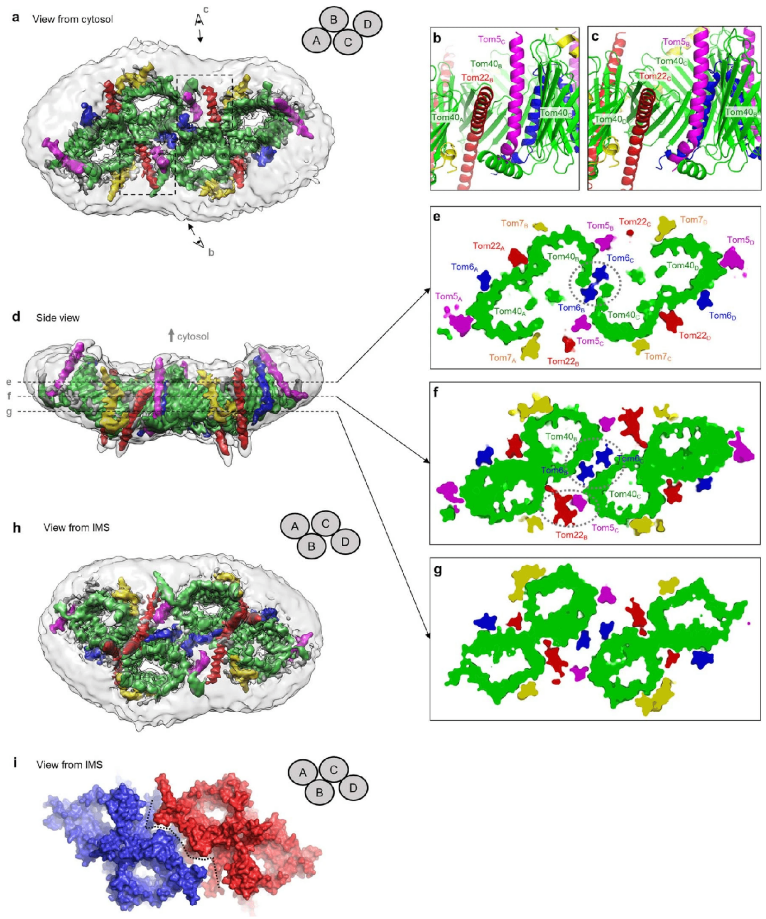


Figure 2.14: **Dimer-dimer interface in the tetrameric TOM complex**

**a,** Overview (cytosolic view) of the tetrameric TOM complex. The 4.1-Å-resolution 3D reconstruction was represented with a composite map showing two different contour levels to show the protein features (colored contour; low-pass-filtered at 4.1 Å) and the detergent micelle (semitransparent gray contour; low-pass-filtered according to local resolution values). Organization of monomeric units are schematized in the upper right corner. Areas marked by dashed rectangles are shown in **b** and **c** (after rotating for a side view) with arrows and eye symbols indicating the viewing directions. **b,c,** Side views showing the dimer-dimer contacts between units B and C. Note that the tetramer is not symmetric and that there is a sizeable gap between Tom5B and Tom22C **(c)** in contrast to Tom5C and Tom22B **(b)**. **d,** As in **a,** but showing a side view. Dashed lines indicate cross-sectional planes for cutaway views shown in **e**–**g**. **e**–**g,** Cutaway views (views from cytosol) at different positions along the membrane axis. In **e** and **f,** major interactions mediating the tetramerization are indicated by dashed ovals. Note that in **g,** there is a gap along the interface (also see **h** and **i**). **h,** As in a and **d,** but showing a view from IMS. **i,** Solvent-accessible surface of the tetrameric TOM complex. The dashed line indicates the interfacial gap. The two dimers (A–B and C–D) are in blue and red, respectively.

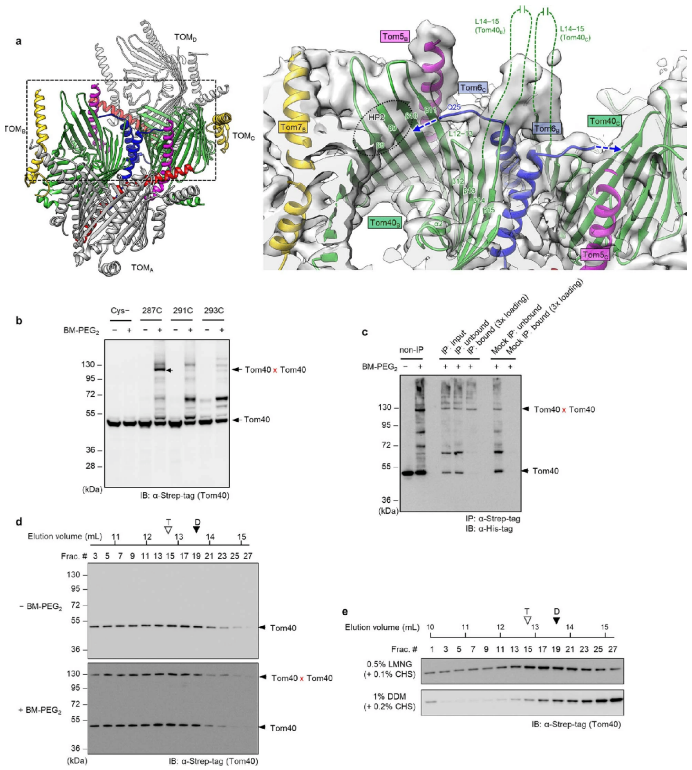


Figure 2.15: **Biochemical validation of higher oligomeric TOM complexes**

**a,** Left, overview (angled cytosolic view) of the tetrameric TOM complex. Monomeric units B and C are shown in color, and A and D are in gray. The region in the black dashed box is magnified and shown in the right panel. Right, the cryo-EM density map (semitransparent gray) and the atomic model (in color) are shown for the B–C interface. The blue dashed arrows indicate the directions of the unmodeled N-terminal segments (residues 1–24) of the Tom6C and Tom6B subunits. The black dotted oval indicates the hydrophobic patch HP2. The green dashed lines indicate the unmodeled loop (L14-15; residues 277–294) between β14 and β15 of Tom40. **b,** Mitochondria were treated with BM-PEG2 and analyzed by SDS-PAGE and immunoblotting (IB). Tom40 contained no or an indicated single cysteine. **c,** Mitochondria were purified from cells expressing Tom40Strep (M287C) from the chromosomal locus and Tom40His (M287C) from a CEN plasmid. After treating with BM-PEG2, mitochondria were solubilized with octyl glucoside and subjected to immunoprecipitation (IP) using anti-Strep-tag antibodies (mock: IP without anti-Strep-tag antibodies). **d,** Mitochondria with Tom40Strep (M287C) expressed from the endogenous promoter were solubilized in 0.5% LMNG and 0.1% CHS and then injected to Superose 6 column. Fractions were treated with BM-PEG2 before SDS-PAGE and immunoblotting. **e,** As in **Fig. 4e,** but with mitochondria isolated from the tom6Δ mutant background. "T" and "D" indicate the peak positions of tetramers and dimers, respectively. Yeast were grown in YPEG (**b and e**) or YPD (**c and d**). Source data for panels **b**–**e** are available with the paper online. The experiments in **b**–**e** were repeated at least twice with similar results.

## 2.9 Extended data and Supplementary information

| Dimeric TOM complex (EMDB-20728) (PDB 6UCU) | Tetrameric TOM complex (EMDB-20729) (PDB 6UCV) | |
|---|---|---|
| **Data collection and processing** | | |
| Magnification | 43,478× | 43,478× |
| Voltage (kV) | 300 kV | 300 kV |
| Electron exposure (e–/Å2) | 61 | 43.9 |
| Defocus range (μm) | −0.8 to −2.5 | −0.9 to −3.0 |
| Pixel size (Å) | 1.15 | 1.15 |
| Symmetry imposed | *C* 2 | *C* 1 |
| Initial particle images (no.) | 460,148 | 173,511 |
| Final particle images (no.) | 160,577 | 104,905 |
| Map resolution (Å) | 3.06 | 4.12 |
| FSC threshold | -0.143 | -0.143 |
| Map resolution range (Å) | 2.6–8.5 | 3.4–15 |
| **Refinement** | | |
| Initial model used (PDB code) | De novo | Dimeric complex (6UCU) |
| Model resolution (Å) | 3.06 | 4.12 |
| FSC threshold | -0.143 | -0.143 |
| Map sharpening *B* factor (Å2) | −60 | −60 |
| Model composition | | |
| Nonhydrogen atoms | 8,414 | 15,103 |
| Protein residues | 7,438 | 15,011 |
| Ligands | 976 | 92 |
| *B* factors (Å2) | | |
| Protein | 59.81 | 125.97 |
| Ligand | 58.05 | 71.71 |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.006 | 0.004 |
| Bond angles (°) | 0.955 | 0.825 |
| Validation | | |
| MolProbity score | 1.24 | 1.34 |
| Clashscore | 3.02 | 3.97 |
| Poor rotamers (%) | 0.12 | 0 |
| Ramachandran plot | | |
| Favored (%) | 97.22 | 97.16 |
| Allowed (%) | 2.78 | 2.84 |
| Disallowed (%) | 0 | 0 |

**Supplementary Table 1. Conservation of inter-subunit polar interactions in fungal TOM complexes**

| *S. cerevisiae* | | | Type of interaction | Other species* | | | | |
|---|---|---|---|---|---|---|---|---|
| Position in α-helical Tom subunit† | | Position in Tom40† | | Kp | Ca | Sp | Af | Nc |
| Tom22 | W100 | K113 | cation-π | W91–P116 | W96–A117 | W93–P83 | W97–P93 | W87–P93 |
| Tom22 | T105 | H346 | H-bonding | **S96–H346** | **S101–H346** | **S98–Q345** | **S102–Q318** | **S92–H315** |
| Tom22 | S116 | S312 | H-bonding | **S107–S312** | A112–A311 | **S109–S276** | A113–S284 | A103-S280 |
| Tom22 | E120 | R310 | electrostatic | **E111–R310** | **E116–R309** | E113–A274 | **E117–R282** | **D107–R278** |
| Tom22 | E127 | R310 | H-bonding, electrostatic | **E118–R310** | **E123–R309** | E120–A274 | **E124–R282** | **E114–R278** |
| Tom5 | W36 | R52 | cation-π | **W35–R55** | **W34–R56** | L35–K22 | **Y36–R31** | **Y37–R33** |
| Tom6 | T34 | G299 (N) | Weak electrostatic | **Q28–G299**‡ | **T29–G298**‡ | **S21–G263**‡ | **T34–G271**‡ | **S34–G267**‡ |
| Tom6 | N38 | G299 (O) | H-binding | **N32–G299**‡ | **Q33–G298**‡ | **S25–G263**‡ | **S38–G271**‡ | **S38–G267**‡ |
| Tom6 | I49 (O) | R261 | H-bonding | **I43**‡**–R261** | **I44**‡**–R262** | **L36**‡**–R230** | **L49**‡**–R239** | **L49**‡**–R237** |
| Tom6 | Q50 | W243 | NH-π | **Q44–W243** | **Q45–W244** | **K37–W211** | **H50–W221** | S50–W219 |
| Tom6 | D55 | R261 | H-bonding, electrostatic | **D49–R261** | **D50–R262** | **N42–R230** | **E55–R239** | **E55–K237** |
| Tom6 | L57 (O) | Y307 | H-bonding | **L51**‡**–Y307** | **L52**‡**–Y306** | **L44**‡**–Y271** | **L57**‡**–Y279** | **L57**‡**–Y275** |
| Tom7 | H25 | T163 | H-bonding | **T22–T166** | **K30–S167** | **K19–N134** | R20–G144 | **R20–D142** |
| Tom7 | H29 | L135 (O) | H-bonding | **H26–F138**‡ | **H34–L139**‡ | **H23–G105**‡ | **H24–G115**‡ | **H24–G115**‡ |
| Tom7 | N52 | K90 | H-bonding | **N49–K93** | **N57–K94** | R46–K60 | K47–K69 | R47–K71 |
| Tom7 | N52 (O) | K90 | H-bonding | **N49**‡**–K93** | **N57**‡**–K94** | **R46**‡**–K60** | **K47**‡**–K69** | **R47**‡**–K71** |
| Tom7 | L54 (O) | H102 | H-bonding | **L51**‡**–H105** | **L59**‡**–H106** | **I48**‡**–H72** | **F49**‡**–H81** | **L49**‡**–H83** |
| Tom7 | S55 (O) | K90 | H-bonding | **S52**‡**–K93** | **S60**‡**–K94** | **N49**‡**–K60** | **S50**‡**–K69** | **S50**‡**–K71** |
| Tom7 | P56 (O) | T361 | H-bonding | **P53**‡**–S361** | P61‡–A360 | P50‡–L325 | P51‡–I333 | P51‡–A330 |
| Tom7 | L57 (O) | K90 | H-bonding | **L54**‡**–K93** | **F62**‡**–K94** | **L51**‡**–K60** | **L52**‡**–K69** | **L52**‡**–K71** |

* Species names (% amino acid sequence identity of Tom40 to *S. cervisiae* Tom40): Kp, *Komagataella phaffii* (61.7%); Ca, *Candida albicans* (61.1%); Sp, *Schizosaccharomyces pombe* (42.1%); Af, *Aspergillus fumigatus* (40.1%); Nc, *Neurospora crassa* (38.4%). Note,

**Extended Data Table 2. Cryo-EM image process and atomic model refinement**

| | TOM-pALDH (dimer) | *apo* dimeric TOM | *apo* tetrameric TOM |
|---|---|---|---|
| **Accession codes** | | | |
| EM Data Bank | **** | **** | **** |
| Protein Data Bank | **** | **** | **** |
| **Cryo-EM data acquisition and single-particle analysis** | | | |
| Data acquisition | | | |
| Microscope | Titan Krios | Talos Arctica | Titan Krios |
| Acceleration voltage | 300 kV | 200 kV | 300 kV |
| Camera (recording mode) | K2 Summit + GIF (super-resolution mode) | K2 Summit (super-resolution mode) | K2 Summit + GIF (super-resolution |
| Magnification | 43,478x | 43,103x | 43,478x |
| Physical pixel size (Å) | 1.15 | 1.16 | 1.15 |
| Electron dose rate ($e^-$/ Å$^2$/frame) | 1.22 | 1.25 | 1.22 |
| Frame rate (s/frame) | 0.2 | 0.2 | 0.2 |
| Total electron dose ($e^-$/Å$^2$) | 61 | 42.5 (used frames: 1–34) | 43.9 (used frames: 1–36) |
| Defocus range (μm) | −0.8 to −2.5 | −0.9 to −2.5 | −0.9 to −3.0 |
| Number of micrographs collected | 1,766 | 1,528 | 3,104 |
| Number of micrographs used | 1,587 | 976 | 2,457 |
| Image processing and reconstruction | | | |
| Number of extracted particles | 460,148 | 247,202 | 173,511 |
| Box size (pixels) | 256 | 256 | 400 |
| No. of particles in reconstruction | 160,577 | 103,127 | 104,905 |
| Symmetry used for reconstruction | C2 | C2 | C1 |
| Resolution, unmasked (Å) | 4.33 (0.5 FSC)<br>3.93 (0.143 FSC) | 7.1 (0.5 FSC)<br>4.1 (0.143 FSC) | 10.8 (0.5 FSC)<br>7.1 (0.143 FSC) |
| Resolution, masked, corrected (Å) | 3.42 (0.5 FSC)<br>3.06 (0.143 FSC) | 3.90 (0.5 FSC)<br>3.53 (0.143 FSC) | 4.36 (0.5 FSC)<br>4.12 (0.143 FSC) |
| Estimated B-factor (Å$^2$) (cryoSPARC) | 99.7 | 89.0 | 60.8 |
| **Model Refinement (Phenix)** | | | |
| Map pixel size (Å) | 1.15 | 1.16 | 1.15 |
| Map sharpening B-factor (Å$^2$) | -60 | -50 | -60 |
| Map lowpass filter (Å) | 3.06 | 3.53 | 4.12 |
| Refinement resolution limit (Å) | 3.06 | 3.53 | 4.12 |
| Number of atoms, protein | 14,632 (incl. hydrogen) | 7,378 | 14,891 |
| Number of atoms, non-protein | 2,172 (DDM molecules) | 70 (two DDM molecules) | 92 (two DMPC molecules) |
| **Model Statistics** | | | |
| Average B-factor (Å$^2$) | 63.81 | 39.67 | 131.83 |
| r.m.s deviations, bond length (Å) | 0.008 | 0.006 | 0.005 |
| r.m.s deviations, bond angle (°) | 1.25 | 1.22 | 1.17 |
| Ramachandran Plot | | | |
| Favored (%) | 97.09 | 96.34 | 97.20 |
| Outliers (%) | 0.00 | 0.00 | 0.00 |
| Rotamers | | | |
| Favored (%) | 96.62 | 96.74 | 97.11 |
| Outliers (%) | 0.13 | 0.00 | 0.00 |
| MolProbity scores | | | |
| Clash score / percentile | 2.74 / 98% | 4.66 / 95% | 3.83 / 96% |
| Overall score / percentile | 1.22 / 99% | 1.48 / 96% | 1.33 / 98% |

# Chapter 3

# Cryo-EM structures of the channelrhodopsin ChRmine in lipid nanodiscs

This chapter presents work described in [82]. The cryo-EM data processing, cryo-EM experiments, and the biochemistry were performed and designed by myself. The structure was modeled and refined by Steve and myself. Savitha designed, performed, and analyzed all electrophysiology experiments. Hillel and Steve conceived of and supervised the project. The manuscript, with input from all authors, was written by Steve and myself.

## 3.1 Abstract

Microbial channelrhodopsins are light-gated ion channels widely used for optogenetic manipulation of neuronal activity. ChRmine is a bacteriorhodopsin-like cation channelrhodopsin (BCCR) more closely related to ion pump rhodopsins than other channelrhodopsins. ChRmine displays unique properties favorable for optogenetics including high light sensitivity, a broad, red-shifted activation spectrum, cation selectivity, and large photocurrents, while its slow closing kinetics impedes some applications. The structural basis for ChRmine function, or that of any other BCCR, is unknown. Here, we present cryo-EM structures of ChRmine in lipid nanodiscs in apo (opsin) and retinal-bound (rhodopsin) forms. The structures reveal an unprecedented trimeric architecture with a lipid filled central pore. Large electronegative cavities on either side of the membrane facilitate high conductance and selectivity for cations over protons. The retinal binding pocket structure suggests channel properties could be tuned with mutations and we identify ChRmine variants with ten-fold decreased and two-fold increased closing rates. A T119A mutant shows favorable properties relative to wild-type and previously reported ChRmine variants for optogenetics. These results provide insight into structural features that generate an ultra-potent microbial opsin and provide a platform for rational engineering of channelrhodopsins with improved properties that could

expand the scale, depth, and precision of optogenetic experiments.

## 3.2 Introduction

Rhodopsins are photoreceptor proteins divided into type I (microbial) and type II (animal) groups[83, 84, 85]. Rhodopsins consist of an opsin protein with seven transmembrane helices (TM1-TM7) rendered photosensitive by a covalently bound retinal chromophore[84, 85, 86, 87, 88]. Light-induced retinal isomerization results in protein conformational changes coupled to downstream functions. Animal type II rhodopsins are G-protein coupled receptors in which isomerization of 11-cis-retinal to all-trans-retinal initiates signaling through heterotrimeric G-proteins and arrestin. Microbial type I rhodopsins are more diverse; isomerization of all-trans retinal to 13-cis-retinal results in protein conformational changes that are coupled to enzymatic, ion pump, or ion channel activities in different proteins[84, 85, 86, 87, 88].

Ion pump and ion channel rhodopsins are widely utilized in optogenetics where their heterologous expression in targeted cells enables control of membrane potential and electrical excitability with light[85, 89]. Targeted neuronal activation is typically achieved with depolarizing cation-selective channelrhodopsins, while silencing is achieved with hyperpolarizing pumps or anion-selective channelrhodopsins[85]. Two decades since the initial characterization of a cation-selective channelrhodopsin (CCR) from the chlorophyte algae Chlamydomonas reinhardtii ChR1[90], the optogenetic toolbox has been greatly expanded by efforts to discover or engineer novel channelrhodopsins with varied properties[85]. Key features distinguishing channelrhodopsins include selectivity for different ions, kinetics of channel opening and closing, light sensitivity, spectral properties, and photocurrent magnitude. Most of the currently utilized depolarizing channelrhodopsins are derived from genes from chlorophyte algae (including ChR2/CHETA[91, 92], CoChR[93], Chronos/ChroME[93, 94, 95], and Chrimson[93]). However, a distinct family of channelrhodopsins was recently identified in cryptophyte algae and termed bacteriorhodopsin-like cation channelrhodopsins (BCCRs). BCCRs more closely resemble archaeal pump rhodopsins, such as Halobacterium salinarum bacteriorhodopsin, than CCRs in sequence and mechanistic properties[96, 97, 98]. Gating of the BCCR GtCCR2 involves proton transfer between the retinal Schiff base and residues in a manner analogous to how proton transport is achieved in bacteriorhodopsin[98]. Still, the mechanistic basis for BCCR channel activity in the context of higher homology to pump rhodopsins is not fully understood.

ChRmine is a BCCR identified through structure-based sequence database mining that is distinct among optogenetic tools for its combination of high light sensitivity, cation over proton selectivity, large photocurrents, and broad, red-shifted action spectrum[99]. These properties have enabled previously difficult or intractable experiments including holographic stimulation of large numbers of cortical neurons and transcranial stimulation of deep brain circuits[99, 100, 101]. Still, some properties of ChRmine limit its utility in certain applications, including a slow closing rate that precludes accurate command of high frequency spike

trains and a broad absorbance spectrum that limits spectral multiplexing with other tools[95, 99]. Structure-based engineering of other channelrhodopsins has successfully tuned channel properties[94, 95, 102, 103, 104, 105], but this approach is currently limited for ChRmine by the divergence of BCCRs from CCRs and absence of an experimentally determined ChRmine (or any BCCR) structure.

Here, we present cryo-EM structures of ChRmine in apo (opsin) and retinal-bound (rhodopsin) forms to provide insight into the molecular determinants of channel activity in a BCCR, to better understand unique ChRmine properties, and to facilitate rational engineering of ChRmine variants to create new optogenetic tools.

## 3.3 Results and discussion

We determined cryo-EM structures of ChRmine (residues 1–309 of Rhodomonas lens cation channelrhodopsin 1) in lipid nanodiscs (Fig. 3.1, Table 1). ChRmine was cloned with a cleavable C-terminal GFP tag, expressed in insect cells grown in the presence of 5$\mu$M all-trans retinal, extracted and purified in detergent, and reconstituted into lipid nanodiscs formed by the scaffold protein MSP1E3D1 and a mixture of DOPE, POPC, and POPS lipids.

We observed two populations of ChRmine particles distinguished by nanodisc diameter. Particles with small and large diameter nanodiscs were processed separately and generated reconstructions to 2.7Åand 3.1Åresolution, respectively (Fig. 3.1, Supplementary Figs. 3.6 and 3.7; Table 1). ChRmine residues 2–285 were de novo modeled into each map. Residue 2 is best modeled as an acetylalanine, consistent with common posttranslational enzymatic cleavage of the initiating methionine and acetylation of adjacent small hydrophobic residues (Supplementary Fig. 3.8)[106]. The cytoplasmic C-terminal 19 amino acids are not visible in the cryo-EM map, likely due to disorder. Both maps display clear density for one retinal and multiple lipids per subunit, while the higher resolution map from small diameter ChRmine nanodiscs showed additional density for nineteen ordered water molecules per subunit. As there are no other substantial differences between the two ATR-bound ChRmine structures (overall r.m.s.d.=0.28Å), we focus our discussion below on the higher resolution structure unless otherwise noted.

ChRmine is a homotrimer with a central pore between subunits (Fig. 3.1), unlike all other channelrhodopsins of known structure which are homodimers[85, 107]. The general topology of each ChRmine subunit is typical of rhodopsins, with an extracellular N-terminal region (amino acids 2–26), seven transmembrane helices (TM1-7, amino acids 27–270), intracellular and extracellular linkers (ICL1-3 and ECL1-3), a cytoplasmic C-terminal region (Fig. 3.1c, d). A conserved lysine (K257) on TM7 is covalently linked to a molecule of all-trans retinal through a Schiff base. Trimerization in ChRmine is achieved through interactions between transmembrane helices, ECL1 linkers, and N-terminal regions (Fig. 3.1a, b, e). Association of TM1 and TM2 from one subunit with TM4 and TM5 from the neighboring subunit creates a continuous protein interface across the lipid membrane (Fig. 3.1e). Intersubunit grooves on

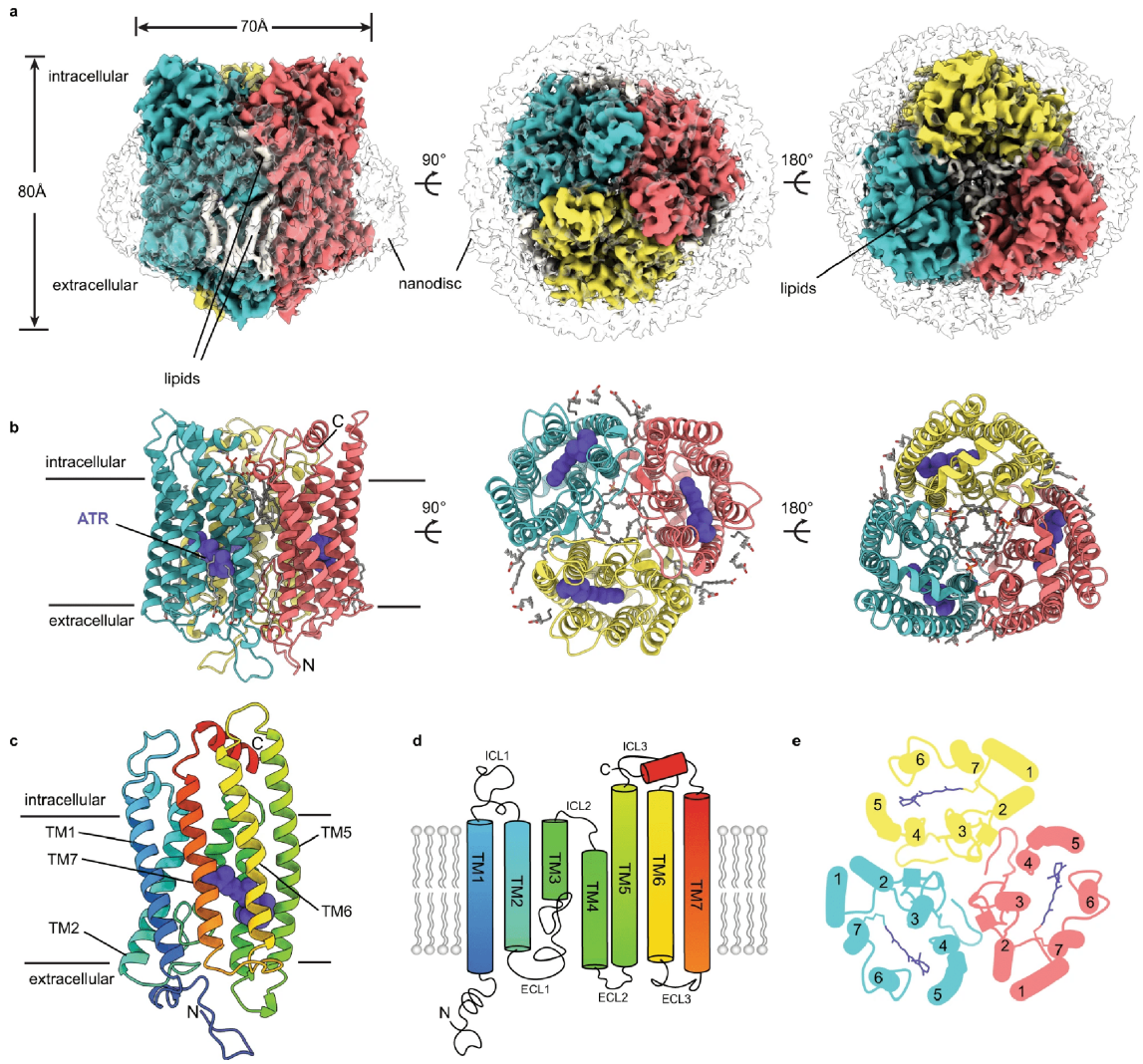


Figure 3.1: **Structure of ChRmine in lipid nanodiscs**

**a,** Cryo-EM map of retinal-bound ChRmine in a MSP1E3D1 nanodisc at 2.7Å resolution viewed from (left) the membrane plane, (middle) the extracellular side, and (right) the intracellular side. Subunits are colored red, teal, or yellow; lipids are white; and the nanodisc is transparent. **b** Model shown in the same views as **(a)** with all-trans retinal shown as purple spheres. **c** Model and **(d)** cartoon illustration of a single ChRmine subunit colored blue (N-terminus) to red (C-terminus). **e** Helical arrangement and packing in the ChRmine trimer viewed from the intracellular side and clipped above the all-trans retinal.

the extracellular side and a pocket on the intracellular side of the transmembrane interface are occupied by ordered lipid acyl chains (Fig. 3.1a, b). This arrangement is analogous to ion pump rhodopsins[108, 109] and contrasts with CCR homodimers that assemble through interactions between TM3, TM4, and the N-terminal region of each subunit[13, 110, 111, 112].

While central pores are also observed in bacteriorhodopsin pump trimers[108], the ChRmine central pore is distinctly shaped with a pair of constrictions separating polar extracellular and nonpolar intracellular central cavities (Fig. 3.2a). The hourglass-shaped pore is a consequence of uniquely structured ECL1s and N-terminal regions in ChRmine (Fig. 3.2b). In other channelrhodopsins of known structure, including C1C2[13] and Chrimson[112], ECL1 forms a short two-stranded β-sheet connecting TM2 and TM3 (Fig. 3.2b). In ChRmine, TM2 is extended by one helical turn and TM3 is partially unwound at its N-terminal end (Fig. 3.2b). The intervening ECL1 is kinked into a staple-like "C" shape that projects towards the central trimer axis with residues 102–107 lining 15Åof the central pore between subunits (Fig. 3.2b, c). G14-T23 from the N-terminal region forms a short helix and linker that caps ECL1, while A2-V13 extend another 15Åinto the cytoplasmic solution to form a mouth below the central pore (Fig. 3.2b, Supplementary Fig. 3.8).

Could the central pore form a conduction pathway in ChRmine and perhaps help explain its large photocurrents relative to other channelrhodopsins? The extracellular central cavity is indeed wide and electronegative as might be expected for a cation permeation path (Fig. 3.2e). Acetylated A2, G14, G102, and F108 carbonyls and the D17 side chain line the walls of the extracellular central cavity as it tapers from 5 to 1.4Åin radius at a polar constriction formed by F104 carbonyls about one-third of the way across the membrane bilayer (Fig. 3.2a, c, d). However, the intracellular central cavity is nonpolar and packed with lipids that preclude ion passage (Fig. 3.2a, c, d, e). Hydrophobic side chains from TM2, TM3, and TM4 (including F76, F80, L118, P121, M122, Y125, Y129, A141, F144) line the walls of the intracellular central cavity as it tapers from 6Åto 1.8Åin radius at a hydrophobic constriction formed by I106 side chains about two-thirds of the way across the membrane. Consistent with the nonpolar environment (and analogous to ion pump rhodopsin trimers[108, 109, 113]), we observe strong tube-shaped features in this region that are well fit by three complete and tightly packed DOPE lipid molecules (Fig. 3.2c, Supplementary Figs. 4, 5). Association of lipid acyl chains results in constrictions less than 1Åin radius, effectively sealing the pore. The lipids are fenced in the central pore away from bulk lipids due to cytoplasmic extensions of ChRmine ICLs and C-terminal regions that form an intracellular ring connecting the three subunits (Supplementary Fig. 3.9). Opening a central pore would therefore require dilation of the two central constrictions and breaking intersubunit interactions to permit lipids to diffuse out of the pore. We conclude the central pore is unlikely to serve as a conduction pathway in ChRmine.

Instead, ChRmine displays features consistent with a single conduction pathway within each subunit (Fig. 3). As expected, the channel is captured in a closed ground state conformation, but the relative positions of TM1, TM2, TM3, and TM7 create deep extracellular and intracellular cavities for ion conduction that are only sealed from one another at the

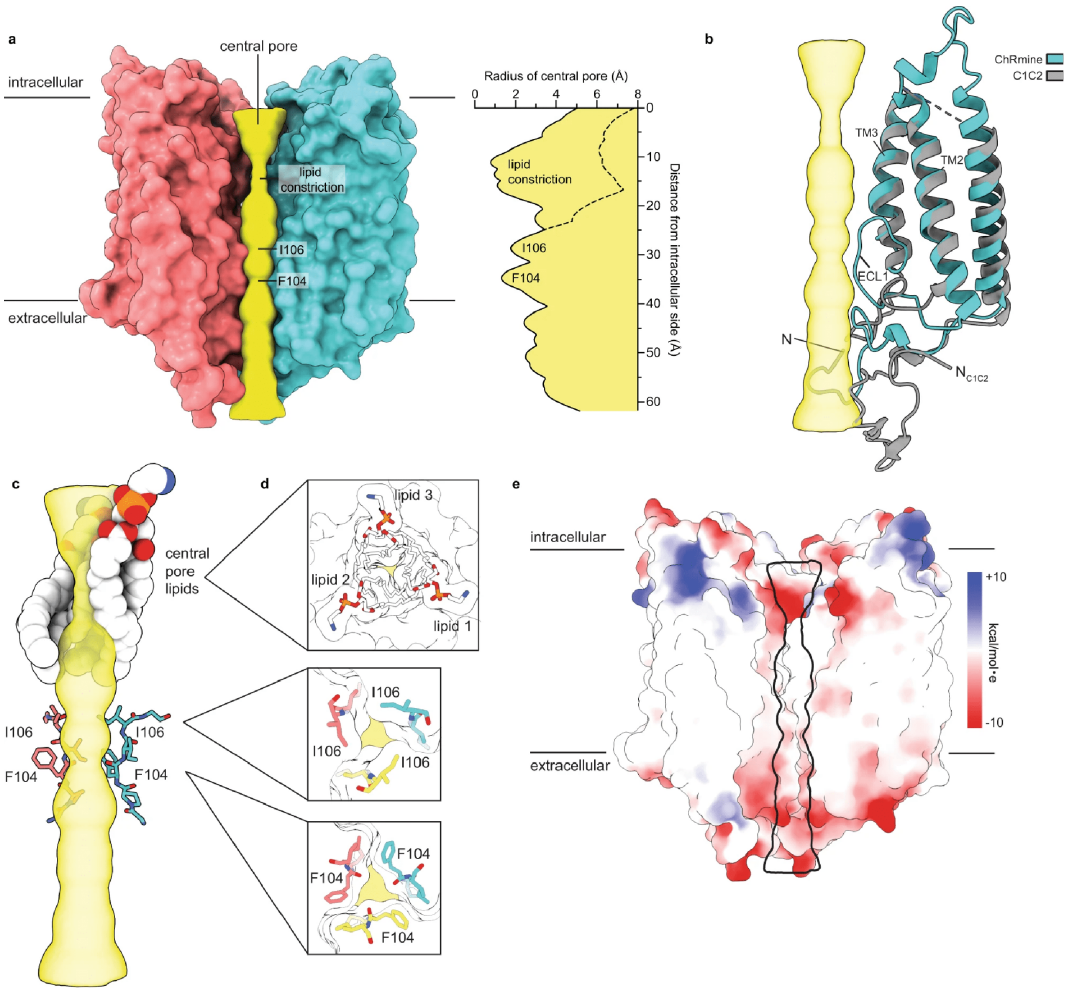


Figure 3.2: **A lipid-filled central pore in ChRmine**

**a,** View from the membrane plane of the ChRmine central pore (yellow) against the surface of two channel subunits (the front subunit is not displayed). Corresponding pore radius as a function of distance from the intracellular side is shown on the same scale. b Overlay of a ChRmine and a C1C2 subunit. The uniquely structured ECL1 and N-terminal region of ChRmine line the central pore. c Constrictions of the central pore formed by lipids and ECL1 viewed from the membrane plane and d from the intracellular side. e Electrostatic character of the central pore drawn on the ChRmine surface in the same view as (a).

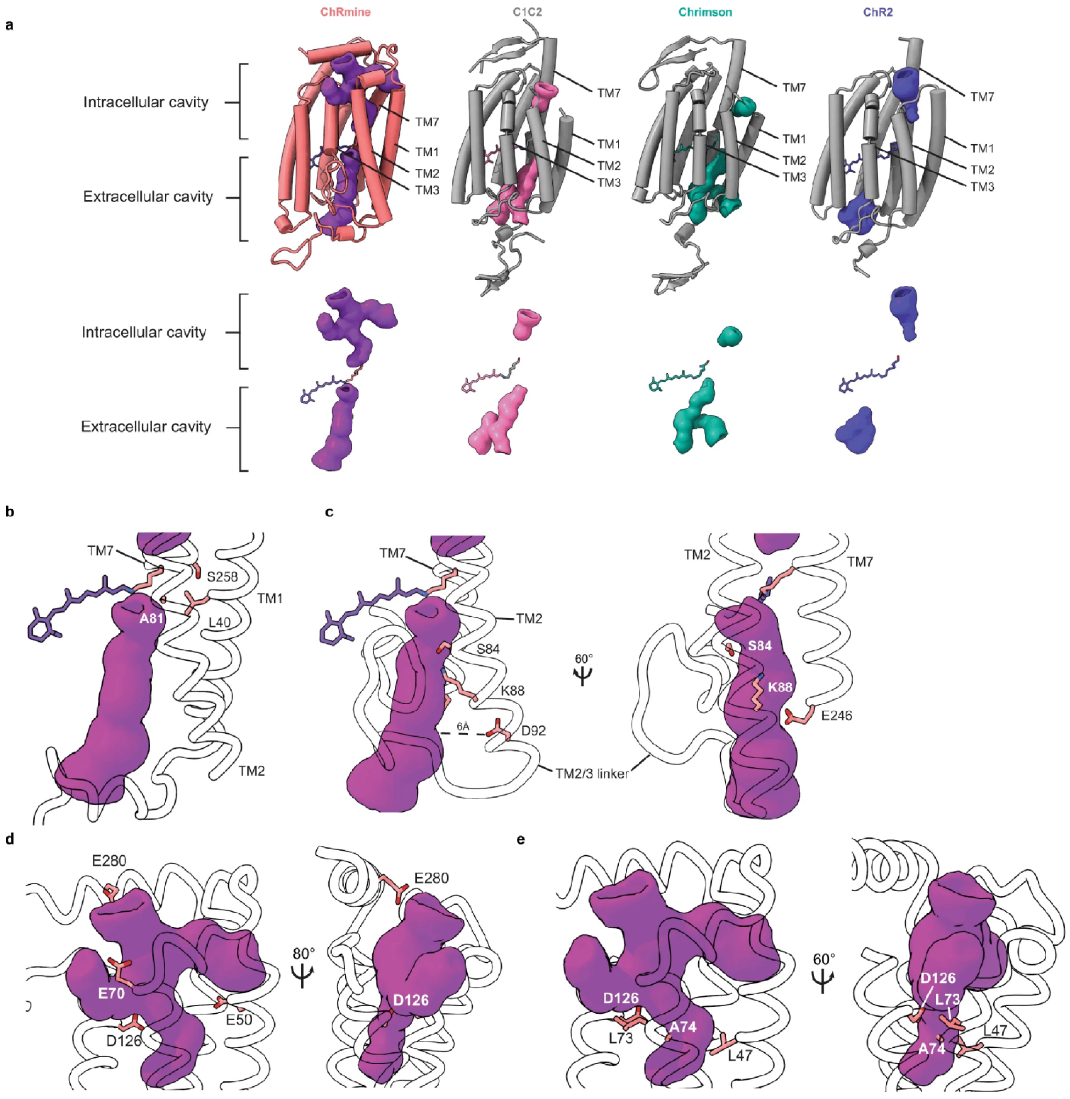


Figure 3.3: **The ChRmine ion conduction pathway**

**a,** Ion conducting pathways in ChRmine, C1C2, Chrimson, and ChR2. A single subunit is shown with helices as tubes (upper). Surfaces of ion conducting pathways are purple, pink, green, or blue. Zoomed in views of the ChRmine (b, c) extracellular and d, e intracellular cavities with key residues indicated.

Schiff base (Fig. 3a). Comparing the presumed conduction pathway in ChRmine to other channelrhodopsins reveals differences on both sides of the membrane that could contribute to properties of ChRmine that differ from other channelrhodopsins.

First, the ChRmine extracellular cavity is wider, deeper, and more electronegative than in other channelrhodopsins including C1C2, Chrimson, and ChR2 (Fig. 3a). The funnel-

shaped outer cavity extends from the extracellular solution to K257 and the Schiff base (Fig. 3a–c). Acidic residues including D100, D115, E154, E158, D242, E246, and D253 line the cavity walls to create a highly electronegative environment that favors interactions with cations. The funnel narrows to a single constriction 3Åin diameter formed by side chains of D115, D253, Y85, Y116, and K88 just extracellular to the Schiff base (Fig. 3b, c). This contrasts with other channelrhodopsins including C1C2 and Chrimson, which have two narrow extracellular tunnels that terminate 5Åbelow the Schiff base(ref 28,31). The differences in diameter are predominantly due to changes at three sites that correspond to the central and extracellular gates of other channelrhodopsins. First, the central gate in C1C2 (that involves S102, E129 ("E3"), and N297) is absent in ChRmine due to smaller analogous side chains (L40, A81, and S258) and differences in their relative position(Fig. 3b). Second, an outer gate constriction in C1C2 (that involves K132 and E136 ("E4")) is structured differently in ChRmine due to smaller side chains (S84 and K88) and kinking of the adjacent TM3, which together open a larger path towards the Schiff base(Fig. 3c). Third, an additional outer gate constriction in C1C2 (that involves E140 ("E5") and T285) is dramatically dilated in ChRmine due to extension and displacement of TM2 and rearrangement of the TM2-TM3 linker. ChRmine's D92 is positioned 6Åfurther away from the conduction path than the analogous residue C1C2 E140, while electronegativity of the wider ChRmine cavity is increased by the substitution of E246 in place of C1C2 T285(Fig. 3c).

The ChRmine intracellular cavity is also larger, deeper, and more forked relative to other channelrhodopsins (Fig. 3a). In ChRmine, three openings at the intracellular surface (between TM2 and TM3, between TM1 and TM7, and above TM2) merge into a single cavity. Acidic residues including E50, E70, D126, and E280 create an electronegative surface that is expected to favor interactions with cations (Fig. 3d, e). The cavity extends deep into the channel before terminating just intracellular to K257 and the Schiff base. This contrasts with other channelrhodopsins including C1C2, Chrimson, and ChR2 which have a shorter intracellular cavity terminating at an inner gate 7Åcloser to the intracellular solution. The difference is predominantly due to rearrangement of the region corresponding to the intracellular gate in other channelrhodopsins. In C1C2, the intracellular gate involves Y109, E121 ("E1"), E122 ("E2"), and H173 and creates a tight seal at the end of the intracellular cavity. In ChRmine, the smaller corresponding residues L47, L73, A74, and D126 constrict the cavity to a radius of 1.5Å, but do not form a tight seal (Fig. 3a, d, e).

We next consider the structure around the retinal chromophore to better understand spectral and kinetic properties of ChRmine. We focus on three aspects of this region that are unique to ChRmine and have been shown to impact light absorption and channel closing rates in other channelrhodopsins: (i) electrostatic interactions around the Schiff base, (ii) polarity around the retinal β-ionone, and (iii) packing around the retinal polyene and β-ionone moieties.

Electrostatic interactions around the protonated Schiff base are determinants of rhodopsin spectral properties[84, 85, 86, 87, 88]. Generally, negative charge provided by the protein (a counterion) stabilizes the protonated ground state and results in blue light absorption, while the absence of negative charge or increased distance between negative charge and the Schiff

base destabilizes the ground state and results in red shifted absorption. A negatively charged group typically also acts as a proton acceptor during the photocycle. ChRmine has negatively charged groups extracellular to the Schiff base like other rhodopsins, but their positioning and interaction network is distinct (Fig. 3.4a). In ChRmine, D115 and D253 correspond to negative charges in other channel (E162 and D292 in C1C2) and pump rhodopsins (D85 and D212 in HsBr). In C1C2, D292 acts as the proton acceptor and E162 is not required for channel function[13, 114] (Fig. 3.4b). In HsBr, D85 acts as the proton acceptor and D212 is positioned by hydrogen bonds with Y57 and Y185[115] (Fig. 3.4c). ChRmine D253 is held by hydrogen bonds with two tyrosines (Y85 and Y116) like HsBr D212, although the second tyrosine is contributed by TM6 in HsBR and TM3 in ChRmine (Fig. 3.4a, c). However, due to bending and unwinding of the N-terminal end of TM3 in ChRmine, D115 is further away from the Schiff base (7.0Å) and rather forms an interaction through a water molecule positioned where the side chain carboxylates C1C2 E162 or HsBr D85 are found (Fig. 3.4a, b, d). This charge separation could contribute to red-shifted spectra of ChRmine.

A second determinant of rhodopsin spectral characteristics is the polarity around the retinal β-ionone ring. Electronegative binding pockets generally result in red-shifted absorption spectra, perhaps due to stabilization of a retinal isomerization intermediate in which charge is transferred towards the β-ionone ring[112, 116]. In ChRmine, the pocket around the β-ionone ring is distinctly electronegative due to contributions from S149, E154, and Y113 residues (Fig. 3.5a–c). The red-shifted channlerhodopsin Chrimson has a comparably electronegative β-ionone pocket due to similarly positioned serine (S223) contributed from TM5 (compared to S149 contributed from TM4 in ChRmine)[112]. In the relatively blue-shifted C1C2, however, two of three corresponding residues (W201 and A206) are nonpolar or positioned further away from the retinal, resulting in a less electronegative retinal binding pocket[13, 112].

The shape of the retinal binding pocket has also been shown to influence closing rate of other channelrhodopsins. Tight packing around the retinal polyene and β-ionone ring is correlated with fast closure, while mutations that loosen packing have been shown to slow closing, presumably by impairing coupling between retinal and protein conformational changes[112, 117]. The retinal binding pocket is substantially less tightly packed on three sides compared to C1C2 or Chrimson, consistent with the 5-10-fold slower closing rate ChRmine. The intracellular and extracellular sides of the retinal binding pocket are formed by aromatic residues in all three opsins. In C1C2 and Chrimson, the polyene chain is tightly sandwiched between tryptophans (W163/W262 and W166/W265, respectively). In ChRmine, the extracellular tryptophan is substituted by a smaller tyrosine (Y116) and the intracellular tryptophan (W223) is displaced away from the retinal due to kinking of TM6 (Fig. 3.5a–c). Along the lateral axis of the pocket facing TM3 and TM4, M201 in Chrimson closely abuts the retinal and its mutation to smaller residues has been shown to decrease closing kinetics[112]. In ChRmine, I146 in this position does not extend as far along the polyene ring (Fig. 3.5d). The consequence of these three differences in the ChRmine retinal binding pocket are large gaps on three sides of the retinal between C8-C13 not observed in other channelrhodopsins.

During preparation of this manuscript, an independent report of a retinal-bound ChRmine

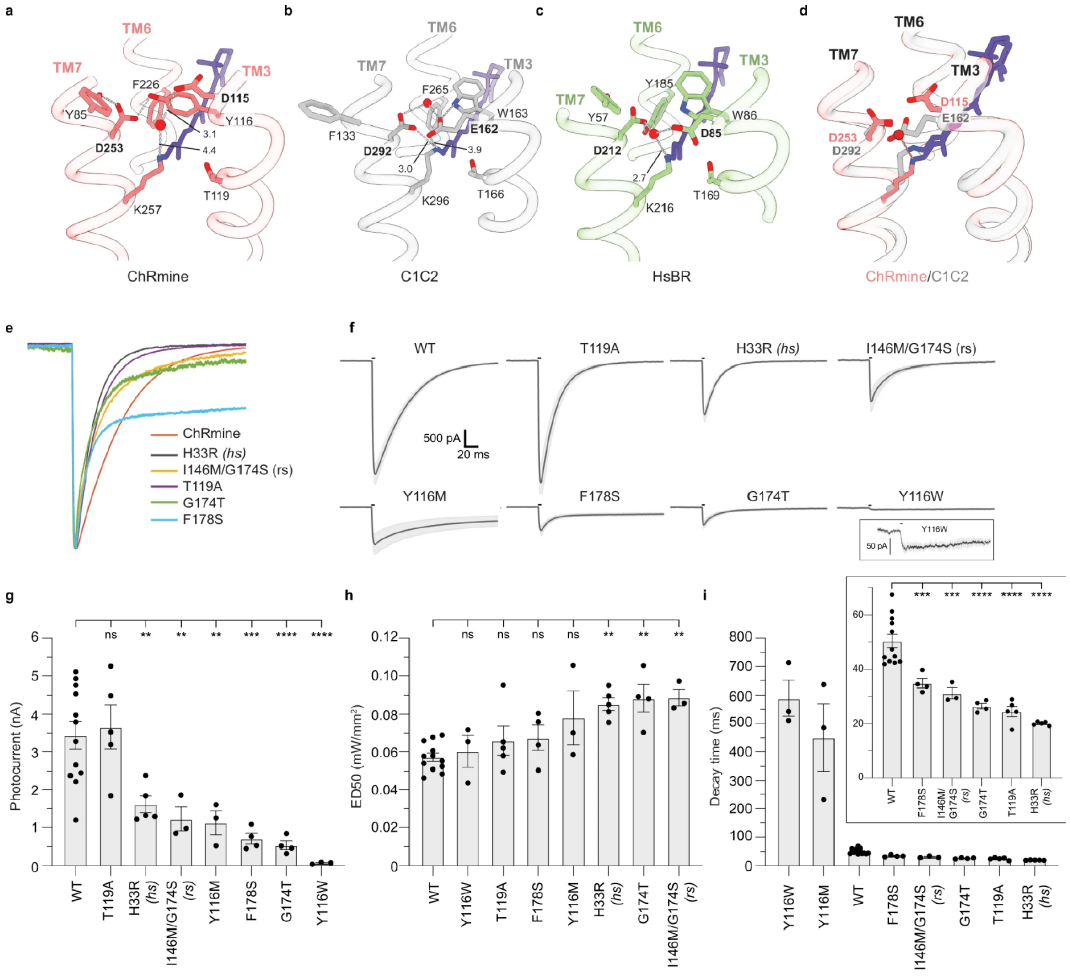


Figure 3.4: **Schiff base chemistry and function of ChRmine variants**

**a,** Views of the area around the protonated Schiff base in a ChRmine, (b) C1C2, (c) HsBr, and d ChRmine and C1C2 overlaid. Negatively charged counterions and proton acceptors, interacting residues, and key water molecules are indicated. e Peak-normalized currents for ChRmine and five mutants with faster closing rate. Note long-lived F178S steady-state current. f Average photocurrents for ChRmine and each mutant evaluated. The bar above each trace indicates a 5 ms light pulse. Y116W is shown in inset with ten-fold expanded current scale. g Peak photocurrents, (h) light sensitivity, and i decay time constant for each channel. Values are mean ± sem for n = 12, 5, 5, 3, 3, 4, 4, and 3 cells for wild-type ChRmine, T119A, H33R, I146M/G174S, Y116M, F178S, G174T, and Y116W, respectively. Differences assessed with one-way analysis of variance (ANOVA) with Dunnett correction for multiple comparisons (ns, not significant (P = 0.9994 for T119A in (i) and 0.9995, 0.7446, 0.6626, and 0.0957 for Y116W, T119A, F178S, and Y116M in h, respectively); **P ¡ 0.01; ***P ¡ 0.001; ****P ¡ 0.0001). i, inset Expanded scale comparing wild-type ChRmine and mutants with accelerated closing rate. Source data for g–i are provided as a Source Data file.

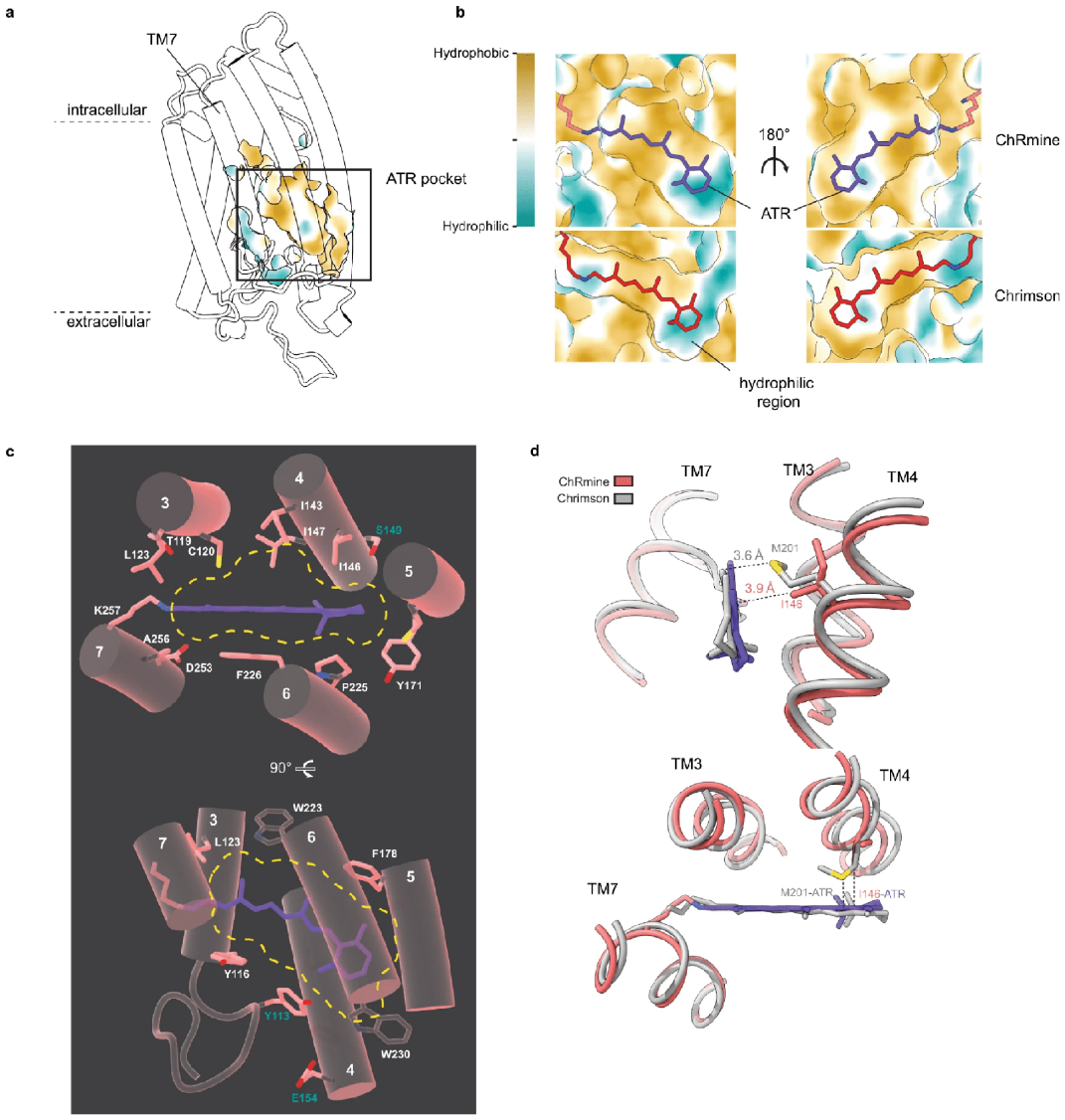


Figure 3.5: **The ChRmine retinal binding pocket**

**a,** a ChRmine viewed from the membrane plane with position of retinal binding pocket highlighted. b Views of the retinal binding pockets of (top) ChRmine and (bottom) Chrimson. Protein surface is colored from hydrophobic (brown) to hydrophilic (teal). c ChRmine retinal binding pocket with key residues drawn as sticks. d Overlay of ChRmine and Chrimson illustrating differences in packing environment.

structure determined in detergent micelles and in complex with antibody fragments to 2.0Åresolution was posted and published[118]. Two substantial differences are observed between that structure and the ones reported here. The first is around the extracellular N-terminal region and does not have obvious functional consequences. We expressed ChRmine without N-terminal modifications and found the N-terminus from each subunit forms a closed ring at the extracellular side of the central pore. This region is not observed in the detergent-solubilized structure[118], probably because additional residues from an N-terminal epitope tag sterically clash with the N-terminal ring, resulting in local disorder.

The second difference is in the central pore and has major consequences for functional interpretation. We determined ChRmine structures in lipid nanodiscs and observed tightly packed lipids in the hydrophobic intracellular cavity of the central pore that preclude ion conduction. This observation is consistent with structures of trimeric ion pump rhodopsins that have similar lipid-filled central pores26. In contrast, lipids are not modeled in the central pore of the detergent solubilized structure[118]. The most likely reason for this difference is that centrally bound lipids are extracted or replaced by less ordered detergent molecules when ChRmine is solubilized in detergent micelles. In the absence of centrally bound lipids, the central pore was predicted to form an ion conduction path that could open upon retinal isomerization and dilation of the F104 and I106 constrictions[118]. We conclude central conduction in trimeric ChRmine is unlikely for several reasons: (1) similarity to other trimeric pump rhodopsins, (2) the hydrophobic nature of the pore, and (3) the presence of clearly resolved lipids in the structure of ChRmine reconstituted into lipid nanodiscs reported here.

We generated mutations in ChRmine to test the functional impact of residues implicated as potential determinants of channel properties (Fig. 3.4e–i). We reasoned mutations at Y116 would decelerate ChRmine closing because Y116 hydrogen bonds with D253 in the counterion network, forms part of the floor of the retinal binding pocket, and contributes to the extracellular constriction. Indeed, Y116M or Y116W mutants that disrupt hydrogen bonding and packing interactions slow closing 10-fold (Fig. 3.4i). Both mutants also reduce photocurrents ( 4- and 50-fold compared to wild-type ChRmine, respectively), though we cannot rule differences in expression level from contributing to this effect (Fig. 3.4g).

We then reasoned other manipulations to the counterion network or conduction path could have opposing effects and accelerate ChRmine closing. S169 in Chrimson contacts the Schiff base and an S169A mutant accelerates channel closing[112]. Consistently, we found the analogous T119A mutation in ChRmine accelerated channel closing 2-fold without significantly impacting photocurrents or light sensitivity relative to wild-type (Fig. 3.4e–i). We confirmed that the reported mutation H33R (called “high speed” or hsChRmine)[118], which lines the extracellular cavity, also accelerated closing. In contrast to T119A though, H33R showed significantly decreased photocurrents and light sensitivity (Fig. 3.4g, h). ChRmine T119A could therefore be preferable to wild-type or H33R ChRmine for optogenetic applications.

We next tested mutations that alter retinal binding pocket properties. A ChRmine I146M/G174S mutation (called “red-shifted” or rsChRmine) was reported to alter spectral

properties[118]. We reasoned it would also accelerate closing rate because the mutations result in a tighter retinal binding pocket: I146 is analogous to M201 in Chrimson[112] that packs against the retinal face and a gap exists between G174 and the retinal indole. Indeed, I146M/G174S ChRmine closed 1.6-fold faster than wild-type. We generated additional mutations predicted to either directly (G174T) or indirectly (F178S, by making space for Y182 to pack against the retinal as seen in all other channelrhodopsin structures) create a tighter retinal pocket. Consistently, G174T and F178S ChRmine displayed accelerated closing (1.9- and 1.4-fold), though we note F178S closing is biphasic with a second slower step.

To date, there are no channelrhodopsin structures without bound retinal and the structural consequences of retinal binding, which could have implications for the design of functional variants, are unknown. We purified apo-ChRmine from cells grown without addition of all-trans retinal, reconstituted it into lipid nanodiscs, and determined its structure to 4.1Åresolution (Supplementary Fig. 3.10, Table 1). The lower resolution of the apo structure is due in part to fewer particles and anisotropy of the reconstruction due to preferred particle orientation and potentially an increase in overall ChRmine flexibility. The apo-ChRmine structure (consisting of residues 3–285) was modeled by combining portions de novo built (into well-defined regions) with docked portions of the ATR-bound structure (into less well-defined regions) prior to structure refinement. The structures are similar overall (r.m.s.d. = 0.8Å); no major rearrangements in the overall fold are observed (Supplementary Fig. 3.11). However, subtle conformational changes are focused to the intracellular halves of TM1, TM5, TM6, and TM7, which bow out toward the membrane 0.5Å, 1Å, 2Å, and 1Å, respectively (Supplementary Fig. 3.11). The overall effect is to expand the retinal binding pocket and loosen packing between TM6 and its neighboring helices TM5 and TM7 on the membrane facing sides. These changes may facilitate loading apo-ChRmine with retinal.

Together, our results provide insight into the unique structure and functional properties of ChRmine and suggest design principles for new and improved channelrhodopsins. ChRmine displays large photocurrents compared to other channelrhodopsins when expressed in neurons or cultured cells[95, 99]. It is unknown how differences in single channel conductance and/or expression levels contribute to large photocurrents for ChRmine compared to other CCRs. R. lens CCR1 (from which ChRmine is derived) does not show substantially larger photocurrents compared to other BCCRs38 in HEK293 cells, suggesting some contribution of expression level differences. We speculate large photocurrents in neurons could be due, at least in part, to differences along the canonical conduction pathway within each subunit that expand and deepen cavities on either side of the membrane. This would be consistent with hsChRmine (H33R)[118], which is predicted to constrict the extracellular pathway, reducing photocurrents 2-fold in addition to accelerating closing rate (Fig. 3.4). We further speculate that chemistry around the Schiff base, counterion network structure, and a loosely-packed retinal binding pocket contribute to slow closing in ChRmine. Mutations that markedly slow (Y116W, Y116M) or accelerate closing rate (T119A, G174T, F178S, and I146M/G174S (rsChRmine[118])) support this hypothesis.

We identify one mutation, ChRmine T119A, with properties of a promising optogenetic tool. ChRmine T119A retains large photocurrents and high sensitivity of the wild-type

channel while accelerating closing rate two-fold to 25ms. Further work to evaluate models discussed above with targeted mutations are likely to result in additional variants that maintain favorable properties of ChRmine while narrowing its absorbance spectrum and accelerating closing rate. Such tools could enable new experiments including large-scale or deep-tissue optogenetics and manipulations in challenging models such as primates.

## 3.4 Methods

### Cloning and protein expression

The coding sequence for ChRmine from Rhodomonas lens was cloned into a custom vector based on the pACEBAC1 backbone (MultiBac; Geneva Biotech, Geneva, Switzerland) with an added C-terminal PreScission protease (PPX) cleavage site, linker sequence, superfolder GFP (sfGFP), and 7xHis tag to generate the construct ChRmine-LNS-LEVLFQGP-SRGGSGAAAGSGSGS-sfGFP-GSS-7xHis for expression in insect cells. Primers used to clone ChRmine into the custom vector:

Forward- AATGATACGGCGACCACCGATctcgagACCACCATGGCACAC,

Reverse- tcGAATTCagAGACAGTCTCCGCAG.

Mutations were introduced using inverse PCR. The following primers were used to generate mutants: T119Afwd TGCTGGCCTGCCCAATGCTG, T119Arev GGCAGGCCAGCATGTAGTCGG, I146Mfwd TTTGCCATGCTGATGTCTGGCGTG, I146Mrev GACATCAGCATGGCAAAGATGATGGCGC, G174Sfwd CGCCTGGTATGGCTTTAGCTGTTTCTGGTTTATCTTCGCCTACTCTATCGTGATGAGC, G174Srev CAGAAACAGCTAAAGCC ATACCAGGCGTAGGCGCCATTCCTCAGCCTAG, G174Tfwd GCTTTACCTGTTTCTGGTTTATCTTCGC, G174Trev AAACAGGTAAAGCCATACCAGGC, F178Sfwd TTCTGGTCCATCTTCGCCTACTCTATCGTG, F178Srev GAAGATGGACCAGA AACAGCCAAAGCCATAC, Y116Wfwd GACTGGATGCTGACCTGCCCAATG, Y116Wrev GGTCAGCATCCAGTCGGCATACCT, Y116Mfwd GACATGATGCTGACCTGCCCAATG,

Y116Mrev GGTCAGCA-TCATGTCGGCATACCT, H33Rfwd GGCGCCAGGTGGTCTTGCTTTATCGTG, H33Rrev AGACCACCTGGCGCCGATGGC.

The ChRmine construct was transformed into DH10Bac E. coli to generate a bacmid according to the manufacturer's instructions. Subsequent steps used Spodopetera frugiperda SF9 cells cultured in ESF 921 medium (Expression Systems, Davis, CA). Bacmid was transfected into adherent SF9 cells using the Escort IV (MillaporeSigma, Burlington, MA) to produce P1 virus. P1 virus was used to infect SF9 cells in suspension at 2 million cells/mL at a multiplicity of infection (MOI) 0.1 to generate P2 virus. Infection was monitored by fluorescence and P2 virus was harvested 48–72hours post infection. P3 virus was generated in a similar manner. P3 viral stock was then used to infect Sf9 cells at 2–4 million cells/mL at a MOI 2–5 for large scale protein expression. For ATR-bound ChRmine samples, 5$\mu$M all-trans retinal (ATR) was added to media 48hours post infection and cells continued to

grow for an additional 12–16h. Cells were harvested by centrifugation at 5000×g for 15min, flash-frozen in liquid nitrogen, and stored at -80°C.

## Protein purification

Cells from 0.5-1L of culture ( 7.5–15mL cell pellet) were thawed and resuspended in 100mL of lysis buffer (50mM Tris, 150mM NaCl, 1mM EDTA pH 8). Protease inhibitors (final concentrations: E64 (1µM), Pepstatin A (1µg/mL), Soy Trypsin Inhibitor (10µg/mL), Benzimidine (1mM), Aprotinin (1µg/mL), Leupeptin (1µg/mL), AEBSF (1mM), and PMSF (1mM)) were added immediately before use and Benzonase (4µl) was added after the cell pellet thawed. Cells were lysed by sonication and membranes pelleted by centrifugation at 150,000 x g for 45minutes. The supernatant was discarded and membrane pellets were scooped into a dounce homogenizer containing extraction buffer (50mM Tris, 150mM NaCl, 1mM EDTA, 1% n-Dodecyl-β-D-Maltopyranoside (DDM, Anatrace, Maumee, OH), pH 8). A 10% solution of DDM was dissolved and clarified by bath sonication in 200mM Tris pH 8 prior to addition to buffer to the indicated final concentration. Membrane pellets were homogenized in extraction buffer and this mixture (150mL final volume) was gently stirred at 4°C for 1–2hr. The extraction mixture was centrifuged at 33,000 x g for 45minutes and the supernatant, containing solubilized ChRmine, was bound to 5mL of Sepharose resin coupled to anti-GFP nanobody for 1–2hour at 4°C. The resin was collected in a column and washed with 10mL of buffer 1 (20mM Tris, 150mM NaCl, 1mM EDTA, 0.025% DDM, pH 7.4), 40mL of buffer 2 (20mM Tris, 150mM NaCl, 0.025% DDM, pH 7.5), and 10mL of buffer 1. The resin was then resuspended in 6mL of buffer 1 with 0.5mg of PPX and rocked gently in a concial tube overnight at 4°C. Cleaved ChRmine protein was then eluted with an additional 8mL of buffer 1, spin concentrated to 500µl with a 10kDa cutoff Amicon Ultra spin concentrator (Millipore), and then loaded onto a Superose 6 increase column (GE Healthcare, Chicago, IL) on an NGC system (Bio-Rad, Hercules, CA) equilibrated in buffer 3 (20mM Tris, 150mM NaCl, 0.025% DDM, pH 7.4). Peak fractions containing ChRmine were collected and spin concentrated prior to incorporation into MSP1E3D1 nanodiscs.

## Nanodisc reconstitution

Freshly purified ChRmine in buffer 3 was reconstituted into nanodiscs formed by the scaffold protein MSP1E3D1 and a mixture of lipids (DOPE:POPS:POPC at a 2:1:1 molar ratio, Avanti, Alabaster, Alabama) at a final molar ratio of 1:4:400 (monomer ratio: ChRmine, MSP1E3D1, lipid mixture). Final concentrations were 20uM ChRmine, 80uM MSP1E3D1, 8mM lipid mix, and 5mM DDM in buffer (20mM Tris pH 7.5, 150mM NaCl). Lipids solubilized in buffer (20mM Tris, 150mM NaCl, pH 7.5) and DDM detergent were first added to ChRmine protein and incubated at 4°C for 30minutes. Purified MSP1E3D1 (prepared as described without His-tag cleavage) was then added and the solution was mixed at 4°C for 10min before addition of 100mg of Biobeads SM2 (Bio-Rad). Biobeads were washed into methanol, water, and then buffer and weighed after liquid was removed by a P1000 tip prior

to use. This final mixture was gently tumbled at 4°C for 12h. Supernatant was cleared of beads by letting large beads settle and carefully removing liquid with a pipette. The sample was spun for 10minutes at 21,000 x g before loading onto a Superose 6 increase column in 20mM Tris, 150mM NaCl, pH 7.5. Peak fractions corresponding to ChRmine protein in MSP1E3D1 were collected for grid preparation.

## CryoEM grid preparation

ChRmine in nanosdiscs was concentrated (10kDa cutoff) to 2.5 or 14.4mgml$^{-1}$ for apo and ATR-bound samples, respectively, and cleared by centrifugation at 21,00×g for 15min at 4°C prior to grid preparation. A 3-µl drop of protein was applied to freshly glow discharged Holey Carbon, 300 mesh R 1.2/1.3 gold grids (Quantifoil). Samples were plunge frozen in liquid nitrogen cooled liquid ethane using a FEI Vitrobot Mark IV (Thermo Fisher Scientific) at 4°C, 100% humidity, 1-blot force, 5s wait time, and 3s blot time.

## CryoEM data acquistion

Grids were clipped and transferred to a 200kV Talos Arctica microscope. Fifty frame videos were recorded on a Gatan K3 Summit direct electron detector in super-resolution counting mode with pixel size of 0.5685Å. The electron dose was 1 e$^-$(Å$^2$s)$^{-1}$ and total dose was 50 e$^-$(Å$^2$)$^{-1}$. Nine movies were collected around a central hole position with image shift and defocus was varied from -0.6 to -2.0µm through SerialEM[73]. Apo-chRmine was collected with a 0.4µm offset from hole center due to higher particle distribution along the carbon edge. See Table 1 for data collection statistics.

## CryoEM data processing

For ATR-bound ChRmine, motion correction and dose weighting were performed on 10,900 movies using RELION 3.1's implementation of MotionCor2 and twice 'binned' to 1.137Å per pixel[76, 74, 119]. Contrast transfer function parameters were fit with CTFFIND 4.1[77]. 8,444,523 particles were template-based autopicked from movies CTF fit to 3.7Å or better, extracted at a 220 pixel box size, and transferred to cryoSPARC v3[75, 120]. After iterative rounds of 2D classification and three class ab initio jobs, two distinct ChRmine particle sets emerged with 192,008 particles in large diameter nanodiscs and 162,477 particles in small diameter nanodiscs. The large and small diameter particle sets were refined separately using non-uniform refinement to a nominal resolution of 3.9Å and 3.1Å, respectively.

The small diameter nanodisc particle set was transferred to RELION3.1 for initial Bayesian polishing followed by CTF refinement and a second round of Bayesian polishing. Particles were then imported into cryosparc and subjected to five additional rounds of 2D classification resulting in a final set of 81,839 particles. An ab initio reconstruction was performed to provide an initial volume and subsequent non-uniform refinement with C3 symmetry imposed resulted in a final map at 2.7Å nominal resolution.

The large nanodisc particle set was further 2D classified for three additional rounds to remove junk, low quality, and damaged particles. This resulted in a final set of 105,606 particles which were subjected to ab initio reconstruction and non-uniform refinement with C3 symmetry imposed to generate a 3.2Å resolution map. These particles were then transferred back to RELION3.1 for two rounds of CTF refinement and Bayesian polishing before being re-imported into cryoSPARC and two-dimensionally classified. Ab initio reconstruction and subsequent non-uniform refinement with C3 symmetry imposed on a final 100,946 particle set resulted in a final map at 3.1Å nominal resolution.

For apo-ChRmine, patch motion correction was performed on 5339 movies using cryoSPARC v2 and later twice 'binned' to 1.137Å per pixel. Contrast transfer function parameters were estimated with patch-based CTF and used to eliminate poor quality data, resulting in 3704 movies. Particles were automatically picked using cryoSPARC's blob picker and the resulting good classes were used for the template-based autopicking of 4,431,227 particles. These particles were extracted at a box size of 220 pixels and two-dimensionally classified yielding 260,305 particles. Subsequently, particles were primarily two-dimensionally classified with the exception of three ab initio jobs using five and four classes. Iterative classification resulted in 41,053 particles. After iterative rounds of two- and three- dimension classification a final set of particles were used to non-uniform refinement to achieve a resolution of 4.1Å.

UCSF pyem tools were used to convert data from cryoSPARC to RELION format.

## Modeling, refinement and analysis

Cryo-EM maps were sharpened in cryoSPARC and used to build models de novo in Coot[79]. Models were refined in Phenix[121] and assessed for proper stereochemistry and geometry using Molprobity[122]. Structural analysis and figure preparation were performed with ChimeraX[123], HOLE[124], DALI[125], PyMOL, JalView[126], Prism 8, and Adobe Photoshop and Illustrator software.

### electrophysiology

Chinese hamster ovary (CHO) cells were transfected with either wild-type or mutant ChRmine fused to surface expression (SE) and endoplasmic-reticulum export (ER) motifs and tagged with the fluorophore mRuby2 fused to the proximal clustering, soma-targeting (ST) domain of the potassium channel Kv2.1(SE-ChRmine-mRuby2-ST-ER). 1-2 days after transfection, photostimulation of cells was performed at 510nm using a Spectra X light engine (Lumencor). Patch pipettes of 4-6 MOhm resistance were pulled from borosilicate glass filaments (Sutter Instruments) and filled with K-gluconate solution (in mM: 110 K-gluconate, 10 HEPES, 1 EGTA, 20 KCl, 2 $MgCl^2$, 2 Na2ATP, 0.25 $Na_3GTP$,10 Phosphocreatine, 295mOsm, pH 7.45). Data was recorded at 20kHz using 700b Multiclamp Axon Amplifier (Molecular Devices). Bath solution (in mM: 119 NaCl, 2.5 KCl, 1.3 $MgSO_4$, 1.3 $NaH2PO_4$, 20 glucose, 26 $NaHCO_3$, 2.5 $CaCl_2$) was maintained at 30-32°C with inline heating. 5-10µM all-trans retinal was added prior to use. All data was acquired and analyzed with custom code written in

Matlab. Currents were measured at a holding potential of -60mV. Decay time constants were measured by fitting the traces from stimuli offset to a single exponential. Light sensitivity is reported as ED50 from a five-parameter logistic regression to plots of photocurrent versus measured light intensity recorded from successive sweeps with varying LED power[127].

## 3.5 Data availability

For ATR-bound ChRmine in small diameter and large diameter nanodiscs the final models are in the PDB under 7SFK and 7SFJ, the final maps are in the Electron Microscopy Data Bank (EMDB) under EMD-25091 and EMD-25079. For apo-ChRmine, the final model is in the PDB under 7SHS, the final map is in the Electron Microscopy Data Bank (EMDB) under EMD-25135. Original micrograph movies and final particle stack is in the Electron Microscopy Public Image Archive (EMPIAR). Source data are provided with this paper.

## 3.6 Acknowledgements

We thank J. Remis, D. Toso, and P. Tobias for microscope and computational support at the Cal-Cryo facility. We thank members of the Brohawn and Adesnik laboratories for discussions and feedback on the manuscript. H.A. and S.G.B. are New York Stem Cell Foundation-Robertson Neuroscience Investigators. This work was funded by an NSF Graduate Research Fellowship and a UC Berkeley Chancellor's Fellowship (K.T); NIH grant UF1NS107574 (H.A.) and the New York Stem Cell Foundation (H.A. and S.G.B.); and NIGMS grant GM123496, a McKnight Foundation Scholar Award, a Sloan Research Fellowship, a Winkler Family Scholar Award (to S.G.B.).

## 3.7 Author information

### Authors and affiliation

Department of Molecular & Cell Biology, University of California Berkeley, Berkeley, CA, 94720, USA Kyle Tucker, Savitha Sridharan, Hillel Adesnik & Stephen G. Brohawn Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, 94720, USA Kyle Tucker, Savitha Sridharan, Hillel Adesnik & Stephen G. Brohawn California Institute for Quantitative Biology (QB3), University of California, Berkeley, CA, 94720, USA Kyle Tucker & Stephen G. Brohawn

### Contributions

K.T. designed and performed all biochemistry and cryo-EM experiments. K.T. performed all cryo-EM data processing. K.T. and S.G.B. modeled and refined the structure. S.S. designed,

performed, and analyzed all electrophysiology experiments. H.A. and S.G.B. conceived of and supervised the project. S.G.B. and K.T. wrote the manuscript with input from all authors.

## Corresponding authors

Correspondence to Hillel Adesnik or Stephen G. Brohawn.

## 3.8 Supplementary Data and Tables

**Table 1 - Cryo-EM data collection, refinement, and validation statistics**

| | ChRmine small nanodisc (PDB 7SFK) (EMDB 25091) | ChRmine large nanodisc (PDB 7SFJ) (EMDB 25079) | apo-ChRmine nanodisc (PDB 7SHS) (EMDB 25135) |
|---|---|---|---|
| **Data collection and processing** | | | |
| Magnification | 36,000 x | 36,000 x | 36,000 x |
| Voltage (kV) | 200 | 200 | 200 |
| Electron exposure ($e^{-}$/Å$^2$) | 50 | 50 | 50 |
| Defocus range (μm) | -0.6 to -2.0 | -0.6 to -2.0 | -0.6 to -2.0 |
| Super resolution pixel size (Å) | 0.5685 | 0.5685 | 0.5685 |
| Binned pixel size (Å) | 1.137 | 1.137 | 1.137 |
| Symmetry imposed | C3 | C3 | C1 |
| Initial particle images (no.) | 8,444,523 | 8,444,523 | 7,054,805 |
| Final particle images (no.) | 81,839 | 100,946 | 41,053 |
| Map resolution (Å) | 2.7 | 3.1 | 4.1 |
| FSC threshold | 0.143 | 0.143 | 0.143 |
| **Refinement** | | | |
| Initial model used (PDB code) | *de novo* | *de novo* | 7SFK |
| Model resolution (Å) | 2.7 | 3.0 | 4.1 |
| FSC threshold | 0.143 | 0.143 | 0.143 |
| Model composition | | | |
| Nonhydrogen atoms | 7419 | 7377 | 6873 |
| Protein residues | 852 | 852 | 849 |
| Ligands | 24 | 24 | 0 |
| Waters | 57 | 3 | 0 |
| Mean *B* factors (Å$^2$) | | | |
| Protein | 14.21 | 59.86 | 149.44 |
| Ligand | 20.21 | 58.10 | - |
| Water | 16.05 | 51.81 | - |
| R.m.s. deviations | | | |
| Bond lengths (Å) | 0.003 | 0.003 | 0.003 |
| Bond angles (°) | 0.536 | 0.537 | 0.740 |
| CC mask | 0.85 | 0.79 | 0.70 |
| **Validation** | | | |
| MolProbity score | 1.15 | 1.27 | 1.33 |
| Clashscore | 3.61 | 3.87 | 5.95 |
| Poor rotamers (%) | 0.43 | 0.85 | 0 |
| Ramachandran plot | | | |
| Favored (%) | 98.22 | 97.51 | 98.22 |
| Allowed (%) | 1.78 | 2.49 | 1.78 |
| Disallowed (%) | 0 | 0 | 0 |

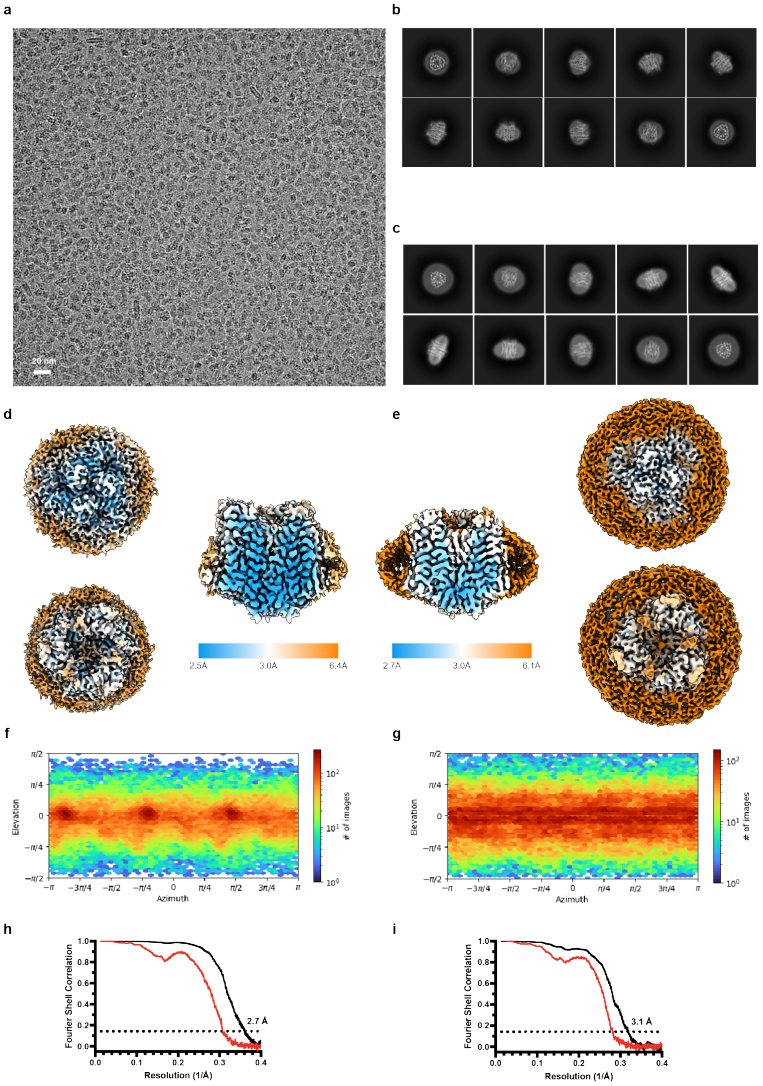


Figure 3.6: **Cryo-EM data and validation for retinal-bound ChRmine structures**

**a,** A representative micrograph and (b,c) 2D class averages for ChRmine bound to all-trans retinal. (b) Small diameter and (c) large diameter nanodiscs were identified in the same dataset and processed separately. (d,e) local resolution drawn on final maps of (d) small diameter and (e) large diameter nanodiscs. (f,g) Angular distribution of particles in the final refinements of ChRmine in (f) small diameter and (g) large diameter nanodiscs. (h,i) Fourier Shell Correlation (FSC) between (black) masked and (red) unmasked half maps from refinement used for calculating resolution at FSC=0.143 from ChRmine in (h) small diameter and (i) large diameter nanodiscs. Displayed micrograph is representative of 10,900 movies collected (Table 1).

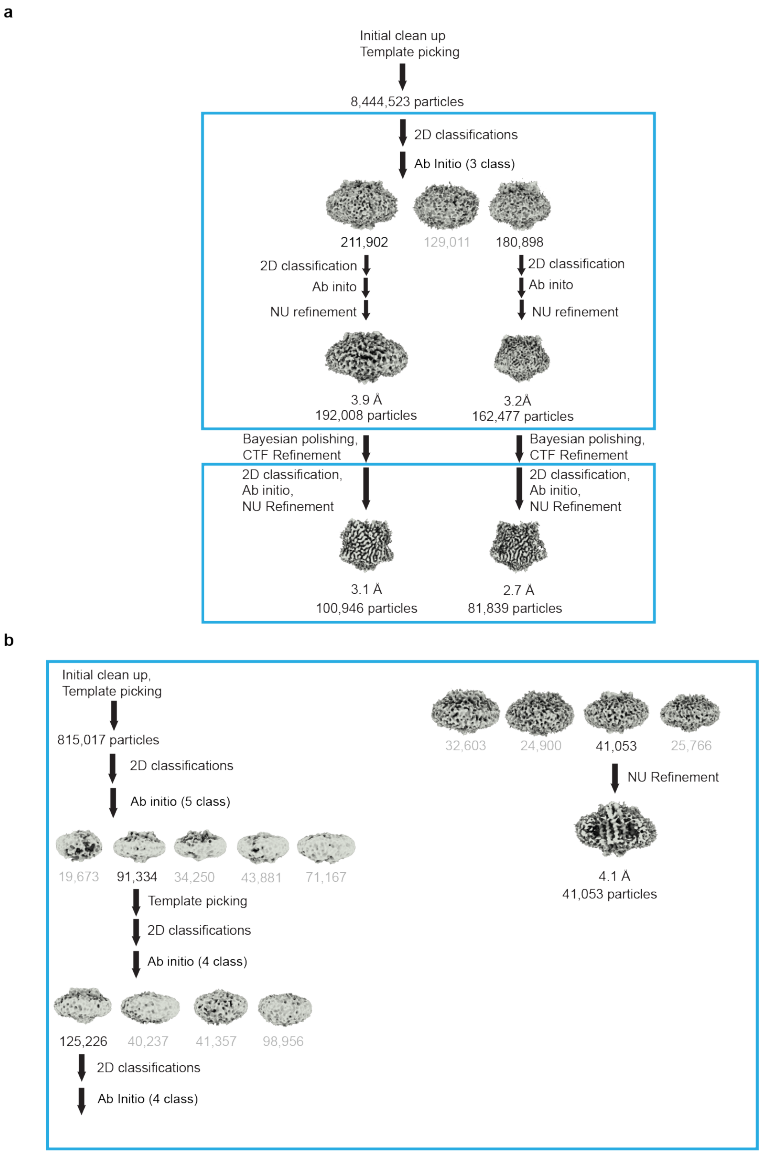


Figure 3.7: **Cryo-EM processing pipelines**

**a,b** Cryo-EM data processing steps for (a) retinal-bound ChRmine and (b) apo-ChRmine in lipid nanodiscs. See Methods for details.

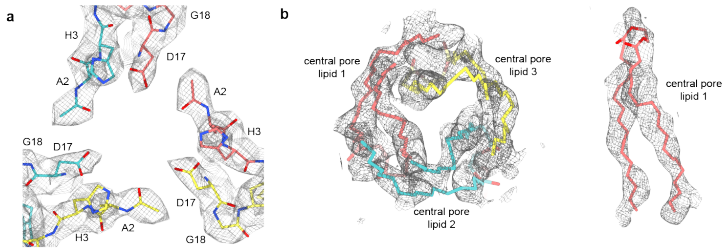


Figure 3.8: **Model and cryo-EM density for retinal-bound ChRmine**

**a** View from the extracellular side of the N-terminal region of ChRmine that forms the mouth of the central pore. The position of acetylated A2 is indicated. (b) View from the (left) intracellular side and (right) the membrane plane of central pore lipid in ChRmine. The model is shown within the final cryo-EM density.

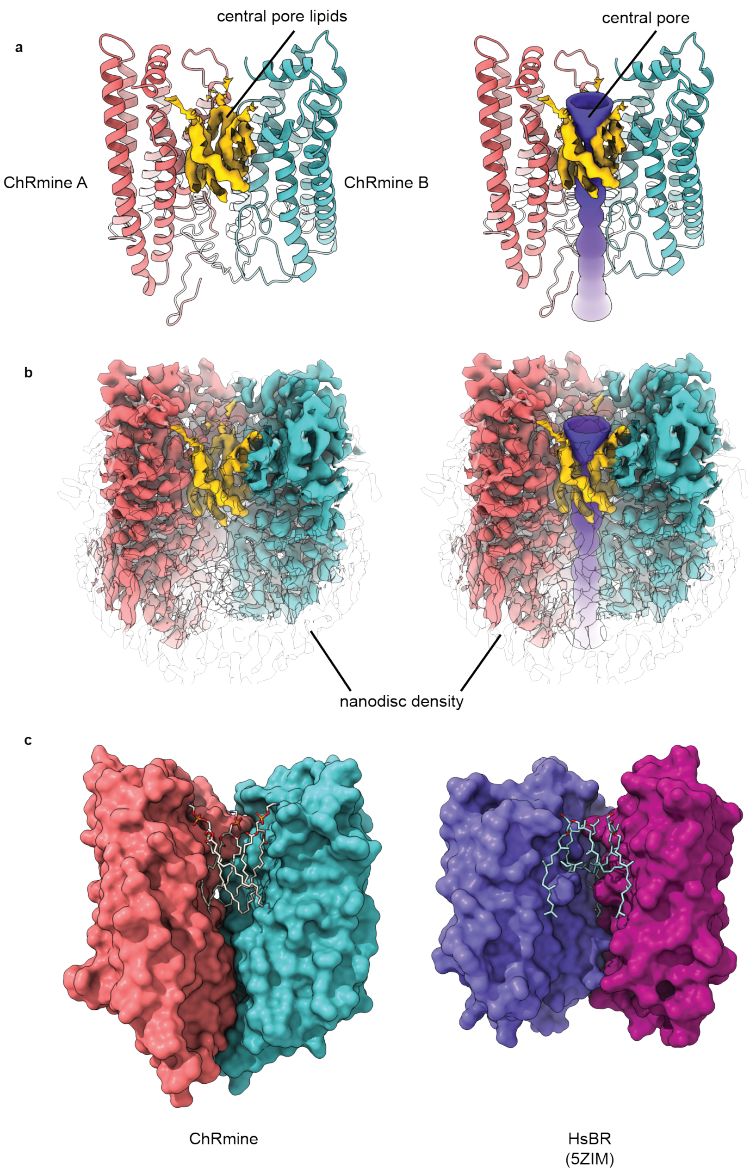


Figure 3.9: **Central pore lipids are occluded within the ChRmine trimer**

**a, left** Tilted view from the intracellular side of two ChRmine subunit models (red and teal) and pore lipid cryo-EM density (orange). (a, right) central pore profile is additionally shown and colored purple. (b) Same as (a) except with cryo-EM density for ChRmine subunits and lipid nanodisc shown. (c) Surface view of two subunits from (left) ChRmine and (right) HsBR with front subunit of each trimer removed. Bound lipids in the central pores are drawn as sticks. Bound lipids block ion conduction through the intersubunit pore and are occluded from bulk lipids by intracellular extensions of ChRmine that form a protein "fence" between the central pore and surrounding membrane.

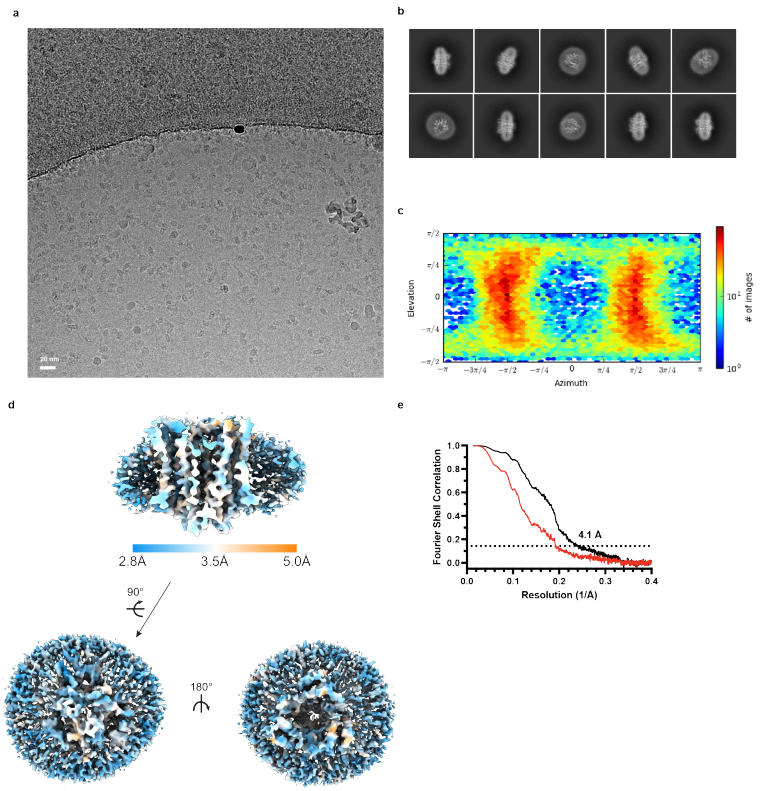


Figure 3.10: **Cryo-EM data and validation for apo ChRmine structures**

**a,** A representative micrograph and (b) 2D class averages for apo-ChRmine in lipid nanodiscs. (c) Angular distribution of particles in the final refinement of apo-ChRmine. (d) Local resolution drawn on final apo-ChRmine map. (e) Fourier Shell Correlation (FSC) between (black) masked and (red) unmasked half maps from refinement used for calculating resolution at FSC=0.143. Displayed micrograph is representative of the 5,339 movies collected (Table 1).

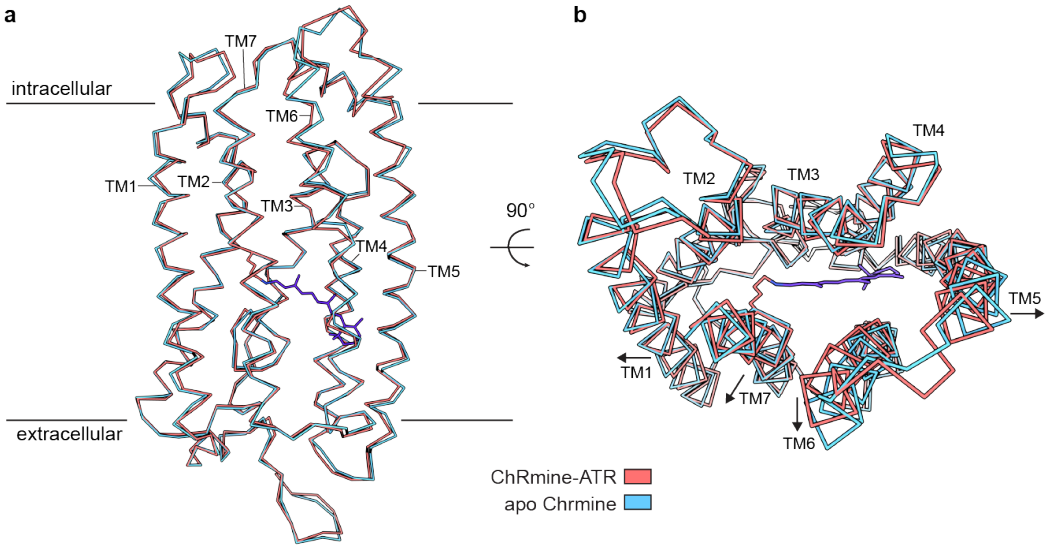


Figure 3.11: **Comparison of retinal-bound and apo ChRmine structures**

**a,** Overlay of retinal-bound (red) and apo (blue) ChRmine from the membrane plane and (b) from the intracellular side. Positions and relative movements of transmembrane helices are indicated.

# Chapter 4

# Conclusions

The visualization of membrane protein channels such as the TOM complex and ChRmine has provided deep insight into how these machines work, how membranes and lipids interact with channels and how one might exploit their structural information for engineering of new properties and fundamental biological insights. For decades, the types of questions asked about protein channels were limited by a lack of high-resolution structural information. However, tremendous advances in cryo-EM and biochemistry have enabled us to capture snapshots of these proteins at near atomic resolution. This has opened the door to asking more complex and detailed questions and even identifying new proteins using structural data.

Decades ago, biochemists using cleverly designed substrates and strategies were able to identify the major preprotein gate for the mitochondria which they termed the TOM complex. Albeit limited, they were able to use crude tools such as protease protection assays, affinity tags and radiolabeling to define the topology of TOM including the identification of its diverse substrates. Eventually some of the fundamental questions surrounding the complex became technically too challenging to address. These included: what was the three-dimensional organization of the complex, and how did this organization/structure facilitate translocation of such diverse substrates (for example, presequence-containing, β-strand-containing, helical, hydrophobic, and cysteine-rich substrates), all destined for different compartments of the mitochondria? While these questions persist, a huge step forward was made once we were able to visualize the TOM complex at high resolution. We now have mechanisms of import grounded in data resolved at the scale of amino acid positions in three-dimensional space. The work in the second chapter provides answers to many questions, including what drives the entrance of presequences into the pore of the TOM complex and once in the pore, what drives presequences toward the inter-membrane space? While we were unable to resolve the interactions of presequences with the pore, we have developed a platform for future work to interrogate these interactions. Many strategies are now available such as creative tethering designs to capture a snapshot of a substrate engaged complex. Given the diversity of substrates there is quite a lot of work to be done to resolve each of these dynamic interactions with the TOM complex. One could imagine an efficient design that is modular and allows for the capture of substrate progression through translocation.

The second chapter addresses another question surrounding the TOM complex and its organization. Due to the lack of high resolution information, many have speculated about the various oligomeric states of the complex. In our system we primarily isolated two major populations of the TOM complex, a dimer and a dimer of dimers. While the single dimer was observed only at a low resolution in previous studies[48], the tetramer was novel state and provided insight into the propensity of the TOM complex oligomerize in this way. We found the TOM complex is sensitive to the removal of lipid interactions that can lead to a collapse into single dimers, which has been further corroborated by recent work[128]. Additionally, TOM dimers can laterally expand at two faces of the core complex, suggesting even higher order oligomer could form. This may have implications for translocation. Future work using cryo-ET and structure-guided biochemistry might be able to address the role of higher order oligomerization. These structures raise additional interesting avenues for future research including whether regions of high protein import and transport efficiency across the outer mitochondrial membrane exist and are associated with higher order oligomerization of TOM complexes[129].

The pump-like channelrhodopsin, ChRmine is less-well studied when compared to the TOM complex, as it was discovered only recently in 2019. However, given its growing use in optogenetics, understanding how ChRmine works is likely to be broadly important because it will allow scientists to rationally tune this tool for their needs. Prior to obtaining the structure, several questions surrounded this channel. These included: what is responsible for ChRmine's high photocurrent? What explains its evolutionary relationship to ion-pump rhodopsins and cation-conducting channelrhodopsins (CCRs)? How does its structure encode ChRmine's unique properties? How might one design optimized versions of ChRmine to enable versatility? In the third chapter, some of these questions are tackled and further developed.

Upon the unveiling of ChRmine's architecture several mysteries were brought further into the light. Initially, the trimer provided a visual representation of ChRmine relatedness to bacteriorhodopsin, a quintessential member of the ion-pump rhodopsin family. It is also clear when compared to CCRs that there are several features of ChRmine that might explain its observed high photocurrent including its increased size, depth, and branching of conduction paths compared to CCRs. Or it might be the number of conduction paths per complex (three versus two). However, these proposals warrant tempering for reasons explained below.

While the trimeric assembly and the shape of the conduction paths in ChRmine were interesting, the architecture of its retinal binding pocket earned the most attention. It's in this pocket one would expect to glean the most insight into the mechanics of ChRmine. This is due to the historical precedent of mutations in these cavities which have dramatically altered the properties of rhodopsins. ChRmine's response to stimulus is initiated by retinal light absorption, leaving all other events secondary or downstream to the photophysics of retinal. The resolved view of this region revealed a relatively spacious pocket and a hybrid layout between ion-pumps and CCRs with respect to the network of counterions surrounding the Schiff base. These two findings likely contribute to ChRmine's unique kinetics and spectra. Leveraging the structure, we were able to test how potentially disrupting bonds,

altering space in the pocket and adapting previously discovered mutations onto ChRmine impact its characteristics. These tests uncovered two variants of ChRmine with increased or decreased rates of closing, which will be most useful for driving neurons at higher frequencies and for further structural work, respectively. Still, the exact mechanism and molecular interactions that explain the phenotype of these mutations is not known. Further work is required to elucidate both how ChRmine cycles through states of activation and inactivation, and how these variants diverge from this process.

Between the two channels described in this thesis, ChRmine and TOM differ in key ways. Based on the work in chapter 2, TOM appears to be a static pore that creates a continuity between the two opposing aqueous fractions across the outer mitochondrial membrane (cytosol and inter-membrane space). For the simplicity of comparison, one could collapse the many conformations of ChRmine into two major states gated by light: open and closed. Considering the dynamic nature of gating in ChRmine, insights gleaned from the closed channel must be limited by the lack of information about other states. As biochemical and specifically structural information is obtained about these other states, we will begin to paint a picture of how work can be performed on each state to enable state transitions (i.e. from a dark-adapted closed state to a light-activated maximally conducting open state). Despite this limitation, the work discussed in chapter 3 describes a non-trivial step forward in our understanding. We now have a well resolved glimpse at a closed state, with and without retinal bound. When combined with the structure from Kishi and colleagues[118], we have a total of three distinct closed states that reveal more about the conformational landscape of ChRmine than ever before. One major finding is that this channel does not appear to undergo large-scale dynamics in the closed state that can be captured by cryoEM. The lack of large movements between structures suggest ChRmine has a “tight” or “compact” fold when closed, which might be a feature of proteins in this broad rhodopsin family with a seven transmembrane architecture. This leads to speculation that even in an activated state, ChRmine might only undergo subtle movements to facilitate the conduction of cations across the membrane. Future work to determine a structure of activated ChRmine is needed to provide more insight into the mechanisms of this channel.

# Bibliography


[1] S. J. Singer and Garth L. Nicolson. "The Fluid Mosaic Model of the Structure of Cell Membranes". In: *Science* 175 (1972).

[2] ROBERT B. COREY LINUS PAULING and H. R. BRANSON*. "THE STRUCTURE OF PROTEINS: TWO HYDROGEN-BONDED HELICAL CONFIGURATIONS OF THE POL YPEPTIDE CHAIN". In: *PNAS* 37 (1951).

[3] W. T. Astbury and A. Street. "X-Ray Studies of the Structure of Hair, Wool, and Related Fibres. I. General". In: *Philosophical Transactions of the Royal Society of London* 230 (1931).

[4] Perutz M. F. "Structure of Haemoglobin: a three-dimensional fourier synthesis at 5.5-Å. Resolution, obtained by X-ray analysis". In: *Nature* 185 (1960).

[5] Kendrew J.C. "A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis". In: *Nature* 181 (1958).

[6] Kendrew J.C. "Structure of Myoglobin A Three-dimensional Fourier Synthesis at 2Å Resolution". In: *Nature* 185 (1960).

[7] Christoph Fahlke. "Membrane Physiology and Biophysics—What Remains to Be Done?" In: *Frontiers in Physiology* 11 (2020).

[8] David Deamer. "The Role of Lipid Membranes in Life's Origin". In: *Life* 7 (2017).

[9] Anastassia A. Vorobieva. "De novo design of transmembranebbarrels". In: *Science* 371 (2021).

[10] Saotome. "Structure of the mechanically activated ion channel Piezo1". In: *Nature* 554 (2018).

[11] Liao. "Structure of the TRPV1 ion channel determined by electron cryo-microscopy". In: *Nature* 504 (2013).

[12] Scheerer. "Crystal structure of opsin in its G-protein-interacting conformation". In: *Nature* 455 (2008).

[13] H. E. Kato. "Crystal structure of the channelrhodopsin light-gated cation channel". In: *Nature* 482 (2012). DOI: 10.1038/nature10870. URL: https://doi.org/10.1038/nature10870.

[14] Henderson. "Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy". In: *JMB* 213 (1990).

[15] Subramaniam and Henderson. "Molecular mechanism of vectorial proton translocation by bacteriorhodopsin". In: *Nature* 406 (2000).

[16] Doyle. "The Structure of the Potassium Channel: Molecular Basis of K1 Conduction and Selectivity". In: *Science* 280 (1998).

[17] Javier. "Structure determination of GPCRs: cryo-EM compared with X-ray crystallography". In: *Biochemical Society Transactions* 49 (2021).

[18] Jeong. "Structures of the TMC-1 complex illuminate mechanosensory transduction". In: *Nature* 610 (2022).

[19] Xudong Wu. "Cryo-EM structure determination of small proteins by nanobody-binding scaffolds". In: *PNAS* 118 (2021).

[20] Denisov. "Nanodiscs for structural and functional studies of membrane proteins". In: *NSMB* 23 (2016).

[21] Eunyong Park Kyle Tucker. "Cryo-EM structure of the mitochondrial protein-import channel TOM complex at near-atomic resolution". In: *NSMB* 26 (2019).

[22] A. Chacinska et al. "Importing mitochondrial proteins: machineries and mechanisms". In: *Cell* 138 (2009).

[23] W. Neupert and J. M. Herrmann. "Translocation of proteins into mitochondria". In: *Annu. Rev. Biochem.* 76 (2007).

[24] N. Wiedemann and N. Pfanner. "Mitochondrial machineries for protein import and assembly". In: *Annu. Rev. Biochem.* 86 (2017).

[25] M. Kiebler. "Identification of a mitochondrial receptor complex required for recognition and membrane insertion of precursor proteins". In: *Nature* 348 (1990).

[26] K. P. Kunkele. "The preprotein translocation channel of the outer membrane of mitochondria". In: *Cell* 93 (1998).

[27] P. J. Dekker. "Preprotein translocase of the outer mitochondrial membrane: molecular dissection and assembly of the general import pore complex". In: *Mol. Cell Biol.* 18 (1998).

[28] U. Ahting. "The TOM core complex: the general protein import pore of the outer membrane of mitochondria". In: *J. Cell Biol.* 147 (1999).

[29] D. Rapaport. "Dynamics of the TOM complex of mitochondria during binding and translocation of preproteins". In: *Mol. Cell Biol.* 18 (1998).

[30] C. Meisinger. "Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small Tom proteins, and import receptors". In: *Mol. Cell Biol.* 21 (2001).

[31] S. Schmitt. "Role of Tom5 in maintaining the structural stability of the TOM complex of mitochondria". In: *J. Biol. Chem.* 280 (2005).

[32] K. e. r. s. t. i. n. Hill et al. "Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins". In: *Nature* 395 (1998).

[33] U. Ahting. "Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria". In: *J. Cell Biol.* 153 (2001).

[34] T. Shiota. "Molecular architecture of the active mitochondrial protein gate". In: *Science* 349 (2015).

[35] K. Zeth. "Structure and evolution of mitochondrial outer membrane proteins of beta-barrel topology". In: *Biochim. Biophys. Acta* 1797 (2010).

[36] S. W. Lackey. "Evidence supporting the 19 beta-strand model for Tom40 from cysteine scanning and protease site accessibility studies". In: *J. Biol. Chem.* 289 (2014).

[37] L. Bolliger et al. "Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria". In: *EMBO J.* 14 (1995).

[38] K. Dietmeier. "Tom5 functionally links mitochondrial preprotein receptors to the general import pore". In: *Nature* 388 (1997).

[39] Y. Abe. "Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20". In: *Cell* 100 (2000).

[40] K. Yamano. "Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import". In: *J. Biol. Chem.* 283 (2008).

[41] H. Yamamoto. "Roles of Tom70 in import of presequence-containing mitochondrial proteins". In: *J. Biol. Chem.* 284 (2009).

[42] J. Qiu. "Coupling of mitochondrial import and export translocases by receptor-mediated supercomplex formation". In: *Cell* 154 (2013).

[43] E. L. Sherman, N. E. Go, and F. E. Nargang. "Functions of the small proteins in the TOM complex of Neurospora crassa". In: *Mol. Biol. Cell* 16 (2005).

[44] T. Becker. "Biogenesis of mitochondria: dual role of Tom7 in modulating assembly of the preprotein translocase of the outer membrane". In: *J. Mol. Biol.* 405 (2011).

[45] D. Rapaport, W. Neupert, and R. Lill. "Mitochondrial protein import. Tom40 plays a major role in targeting and translocation of preproteins by forming a specific binding site for the presequence". In: *J Biol. Chem.* 272 (1997).

[46] J. Melin. "Presequence recognition by the Tom40 channel contributes to precursor translocation into the mitochondrial matrix". In: *Mol. Cell Biol.* 34 (2014).

[47] F. N. Vogtle. "Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability". In: *Cell* 139 (2009).

[48] T. Bausewein. "Cryo-EM Structure of the TOM core complex from Neurospora crassa". In: *Cell* 170 (2017).

[49] K. Model. "Protein translocase of the outer mitochondrial membrane: role of import receptors in the structural organization of the TOM complex". In: *J. Mol. Biol.* 316 (2002).

[50] K. Model, C. Meisinger, and W. Kuhlbrandt. "Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase". In: *J. Mol. Biol.* 383 (2008).

[51] F. Hauer. "GraDeR: Membrane protein complex preparation for single-particle cryo-EM". In: *Structure* 23 (2015).

[52] R. Ujwal. "The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating". In: *Proc. Natl Acad. Sci. USA* 105 (2008).

[53] R. Allen et al. "A conserved proline residue is present in the transmembrane-spanning domain of Tom7 and other tail-anchored protein subunits of the TOM translocase". In: *FEBS Lett.* 514 (2002).

[54] M. Moczko. "The intermembrane space domain of mitochondrial Tom22 functions as a trans binding site for preproteins with N-terminal targeting sequences". In: *Mol. Cell Biol.* 17 (1997).

[55] R. Albrecht. "The Tim21 binding domain connects the preprotein translocases of both mitochondrial membranes". In: *EMBO Rep.* 7 (2006).

[56] S. Wilpe. "Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase". In: *Nature* 401 (1999).

[57] T. Shiota et al. "In vivo protein-interaction mapping of a mitochondrial translocator protein Tom22 at work". In: *Proc. Natl Acad. Sci. USA* 108 (2011).

[58] J. Brix et al. "Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a presequence-carrying preprotein and a non-cleavable preprotein". In: *J. Biol. Chem.* 274 (1999).

[59] E. L. Sherman et al. "Effect of mutations in Tom40 on stability of the translocase of the outer mitochondrial membrane (TOM) complex, assembly of Tom40, and import of mitochondrial preproteins". In: *J. Biol. Chem.* 281 (2006).

[60] K. Gabriel, B. Egan, and T. Lithgow. "Tom40, the import channel of the mitochondrial outer membrane, plays an active role in sorting imported proteins". In: *EMBO J.* 22 (2003).

[61] A. Honlinger. "Tom7 modulates the dynamics of the mitochondrial outer membrane translocase and plays a pathway-related role in protein import". In: *EMBO J.* 15 (1996).

[62] J. Bakelar, S. K. Buchanan, and N. Noinaj. "The structure of the beta-barrel assembly machinery complex". In: *Science* 351 (2016).

[63] Y. Gu. "Structural basis of outer membrane protein insertion by the BAM complex". In: *Nature* 531 (2016).

[64] A. I. C. Hohr. "Membrane protein insertion through a mitochondrial $\beta$-barrel gate". In: *Science* 359 (2018).

[65] S. Cruz. "Proteomic analysis of the mouse liver mitochondrial inner membrane". In: *J. Biol. Chem.* 278 (2003).

[66] H. Sakaue. "Porin associates with Tom22 to regulate the mitochondrial protein gate assembly". In: *Mol. Cell* 73 (2019).

[67] C. Schulz. "Tim50's presequence receptor domain is essential for signal driven transport across the TIM23 complex". In: *J. Cell Biol.* 195 (2011).

[68] C. C. Williams, C. H. Jan, and J. S. Weissman. "Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling". In: *Science* 346 (2014).

[69] A. B. Harbauer. "Mitochondria. Cell cycle-dependent regulation of mitochondrial preprotein translocase". In: *Science* 346 (2014).

[70] M. E. Lee et al. "A highly characterized yeast toolkit for modular, multipart assembly". In: *ACS Synth. Biol.* 4 (2015).

[71] R. McIsaac. "Fast-acting and nearly gratuitous induction of gene expression and protein depletion in Saccharomyces cerevisiae". In: *Mol. Biol. Cell* 22 (2011).

[72] S. Mnaimneh. "Exploration of essential gene functions via titratable promoter alleles". In: *Cell* 118 (2004).

[73] D. N. Mastronarde. "Automated electron microscope tomography using robust prediction of specimen movements". In: *J. Struct. Biol.* 152 (2005).

[74] J. Zivanov. "New tools for automated high-resolution cryo-EM structure determination in RELION-3". In: *eLife* 7 (2018).

[75] A. Punjani et al. "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination". In: *Nat. Methods* 14 (2017).

[76] S. Q. Zheng. "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy". In: *Nat. Methods* 14 (2017).

[77] A. Rohou and N. Grigorieff. "CTFFIND4: fast and accurate defocus estimation from electron micrographs". In: *J. Struct. Biol.* 192 (2015).

[78] T. C. P. Tegunov. "Real-time cryo-EM data pre-processing with Warp". In: *Nat. Methods* 16 (2019).

[79] P. Emsley et al. "Features and development of Coot". In: *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010).

[80] P. V. Afonine. "Real-space refinement in PHENIX for cryo-EM and crystallography". In: *Acta Crystallogr. D Struct. Biol.* 74 (2018).

[81] T. J. Dolinsky et al. "PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations". In: *Nucleic Acids Res.* 32 (2004).

[82] Hillel Adesnik Kyle Tucker Savitha Sridharan and Stephen G. Brohawn. "Cryo-EM structures of the channelrhodopsin ChRmine in lipid nanodiscs". In: *Nature Communications* 13 (2022).

[83] F. Zhang. "The microbial opsin family of optogenetic tools". In: *Cell* 147 (2011). DOI: 10.1016/j.cell.2011.12.004. URL: https://doi.org/10.1016/j.cell.2011.12.004.

[84] E. G. Govorunova et al. "Microbial rhodopsins: diversity, mechanisms, and optogenetic applications". In: *Annu Rev. Biochem* 86 (2014).

[85] K. Deisseroth and P. Hegemann. "The form and function of channelrhodopsin". In: *Sci. (N. Y., NY)* 357 (2017). DOI: 10.1126/science.aan5544. URL: https://doi.org/10.1126/science.aan5544.

[86] O. P. Ernst. "Microbial and animal rhodopsins: structures, functions, and molecular mechanisms". In: *Chem. Rev.* 114 (2014). DOI: 10.1021/cr4003769. URL: https://doi.org/10.1021/cr4003769.

[87] F. Schneider, C. Grimm, and P. Hegemann. "Biophysics of Channelrhodopsin". In: *Annu Rev. Biophys.* 44 (2015). DOI: 10.1146/annurev-biophys-060414-034014. URL: https://doi.org/10.1146/annurev-biophys-060414-034014.

[88] H. Kandori. "Biophysics of rhodopsins and optogenetics". In: *Biophysical Rev.* 12 (2020). DOI: 10.1007/s12551-020-00645-0. URL: https://doi.org/10.1007/s12551-020-00645-0.

[89] L. Fenno, O. Yizhar, and K. Deisseroth. "The development and application of optogenetics". In: *Annu. Rev. Neurosci.* 34 (2011). DOI: 10.1146/annurev-neuro-061010-113817. URL: https://doi.org/10.1146/annurev-neuro-061010-113817.

[90] G. Nagel. "Channelrhodopsin-1: A Light-Gated Proton Channel in Green Algae". In: *Science* 296 (2002). DOI: 10.1126/science.1072068. URL: https://doi.org/10.1126/science.1072068.

[91] G. Nagel. "Channelrhodopsin-2, a directly light-gated cation-selective membrane channel". In: *Proc. Natl Acad. Sci.* 100 (2003). DOI: 10.1073/pnas.1936192100. URL: https://doi.org/10.1073/pnas.1936192100.

[92] L. A. Gunaydin. "Ultrafast optogenetic control". In: *Nat. Neurosci.* 13 (2010). DOI: 10.1038/nn.2495. URL: https://doi.org/10.1038/nn.2495.

[93] N. C. Klapoetke. "Independent optical excitation of distinct neural populations". In: *Nat. Methods* 11 (2014). DOI: 10.1038/nmeth.2836. URL: https://doi.org/10.1038/nmeth.2836.

[94] A. R. Mardinly. "Precise multimodal optical control of neural ensemble activity". In: *Nat. Neurosci.* 21 (2018). DOI: 10.1038/s41593-018-0139-8. URL: https://doi.org/10.1038/s41593-018-0139-8.

[95] S. Sridharan. "High-performance microbial opsins for spatially and temporally precise perturbations of large neuronal networks". In: *Neuron* (2022). DOI: 10.1016/j.neuron.2022.01.008. URL: https://doi.org/10.1016/j.neuron.2022.01.008.

[96] E. G. Govorunova, O. A. Sineshchekov, and J. L. Spudich. "Structurally Distinct Cation Channelrhodopsins from Cryptophyte Algae". In: *Biophys. J.* 110 (2016). DOI: 10.1016/j.bpj.2016.05.001. URL: https://doi.org/10.1016/j.bpj.2016.05.001.

[97] Y. Yamauchi. "Molecular properties of a DTD channelrhodopsin from Guillardia theta". In: *Biophysics Physicobiology* 14 (2017). DOI: 10.2142/biophysico.14.0_57. URL: https://doi.org/10.2142/biophysico.14.0_57.

[98] O. A. Sineshchekov et al. "Bacteriorhodopsin-like channelrhodopsins: Alternative mechanism for control of cation conductance". In: *Proc. Natl. Acad. Sci.* 114 (2017). DOI: 10.1073/pnas.1710702114. URL: https://doi.org/10.1073/pnas.1710702114.

[99] J. H. Marshel. "Cortical layer–specific critical dynamics triggering perception". In: *Science* 365 (2019). DOI: 10.1126/science.aaw5202. URL: https://doi.org/10.1126/science.aaw5202.

[100] R. Chen. "Deep brain optogenetics without intracranial surgery". In: *Nat. Biotechnol.* 39 (2021). DOI: 10.1038/s41587-020-0679-9. URL: https://doi.org/10.1038/s41587-020-0679-9.

[101] T. Matsubara. "Remote control of neural function by X-ray-induced scintillation". In: *Nat. Commun.* 12 (2021). DOI: 10.1038/s41467-021-24717-1. URL: https://doi.org/10.1038/s41467-021-24717-1.

[102] J. Wietek. "Conversion of Channelrhodopsin into a Light-Gated Chloride Channel". In: *Science* 344 (2014). DOI: 10.1126/science.1249375. URL: https://doi.org/10.1126/science.1249375.

[103] A. Berndt et al. "Structure-guided transformation of channelrhodopsin into a light-activated chloride channel". In: *Science* 344 (2014). DOI: 10.1126/science.1252367. URL: https://doi.org/10.1126/science.1252367.

[104] A. Berndt. "Structural foundations of optogenetics: Determinants of channelrhodopsin ion selectivity". In: *Proc. Natl. Acad. Sci.* 113 (2016). DOI: 10.1073/pnas.1523341113. URL: https://doi.org/10.1073/pnas.1523341113.

[105] M. Rappleye and A. Berndt. "Structural basis for ion selectivity and engineering in channelrhodopsins". In: *Curr. Opin. Struc Biol.* 57 (2019). DOI: 10.1016/j.sbi.2019.04.008. URL: https://doi.org/10.1016/j.sbi.2019.04.008.

[106] P. T. Wingfield. "N-terminal methionine processing". In: *Curr. Protoc. Protein Sci.* 88 (2017). DOI: 10.1002/cpps.29. URL: https://doi.org/10.1002/cpps.29.

[107] H. E. Kato. "Optogenetics, light-sensing proteins and their applications in neuroscience and beyond". In: *Adv. Exp. Med. Biol.* 1293 (2021). DOI: 10.1007/978-981-15-8763-4_3. URL: https://doi.org/10.1007/978-981-15-8763-4_3.

[108] M. Kolbe et al. "Structure of the light-driven chloride pump halorhodopsin at 1.8 Å resolution". In: *Science* 288 (2000). DOI: 10.1126/science.288.5470.1390. URL: https://doi.org/10.1126/science.288.5470.1390.

[109] E. Pebay-Peyroula et al. "X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases". In: *Science* 277 (1997). DOI: 10.1126/science.277.5332.1676. URL: https://doi.org/10.1126/science.277.5332.1676.

[110] O. Volkov. "Structural insights into ion conduction by channelrhodopsin 2". In: *Science* 358 (2017). DOI: 10.1126/science.aan8862. URL: https://doi.org/10.1126/science.aan8862.

[111] Y. S. Kim. "Crystal structure of the natural anion- conducting channelrhodopsin GtACR1". In: *Nature* 561 (2018). DOI: 10.1038/s41586-018-0511-6. URL: https://doi.org/10.1038/s41586-018-0511-6.

[112] K. Oda. "Crystal structure of the red light-activated channelrhodopsin Chrimson". In: *Nat. Commun.* 9 (2018). DOI: 10.1038/s41467-018-06421-9. URL: https://doi.org/10.1038/s41467-018-06421-9.

[113] N. Hasegawa et al. "X-ray structure analysis of bacteriorhodopsin at 1.3 Å resolution". In: *Sci. Rep.-uk* 8 (2018). DOI: 10.1038/s41598-018-31370-0. URL: https://doi.org/10.1038/s41598-018-31370-0.

[114] V. A. Lórenz-Fonfría. "Transient protonation changes in channelrhodopsin-2 and their relevance to channel gating". In: *Proc. Natl. Acad. Sci.* 110 (2013). DOI: 10.1073/pnas.1219502110. URL: https://doi.org/10.1073/pnas.1219502110.

[115] K. Gerwert, G. Souvignier, and B. Hess. "Simultaneous monitoring of light-induced changes in protein side-group protonation, chromophore isomerization, and backbone motion of bacteriorhodopsin by time-resolved Fourier-transform infrared spectroscopy". In: *Proc. Natl. Acad. Sci.* 87 (1990). DOI: 10.1073/pnas.87.24.9774. URL: https://doi.org/10.1073/pnas.87.24.9774.

[116] R. Rajamani, Y. Lin, and J. Gao. "The opsin shift and mechanism of spectral tuning in rhodopsin". In: *J. Comput Chem.* 32 (2011). DOI: 10.1002/jcc.21663. URL: https://doi.org/10.1002/jcc.21663.

[117] H. E. Kato. "Atomistic design of microbial opsin-based blue-shifted optogenetics tools". In: *Nat. Commun.* 6 (2015). DOI: 10.1038/ncomms8177. URL: https://doi.org/10.1038/ncomms8177.

[118] K. E. Kishi. "Structural basis for channel conduction in the pump-like channelrhodopsin ChRmine". In: *Cell* (2022). DOI: 10.1016/j.cell.2022.01.007. URL: https://doi.org/10.1016/j.cell.2022.01.007.

[119] J. Zivanov, T. Nakane, and S. H. W. Scheres. "A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis". In: *IUCrJ* 6 (2019). DOI: 10.1107/S205225251801463X. URL: https://doi.org/10.1107/S205225251801463X.

[120] A. Punjani, H. Zhang, and D. J. Fleet. "Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction". In: *Nat. Methods* 17 (2020).

[121] D. Liebschner. "Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix". In: *Acta Crystallogr. Sect. D., Struct. Biol.* 75 (2019). DOI: 10.1107/S2059798319011471. URL: https://doi.org/10.1107/S2059798319011471.

[122] C. J. Williams. "MolProbity: More and better reference data for improved all-atom structure validation". In: *Protein Sci.: a Publ. Protein Soc.* 27 (2018). DOI: 10.1002/pro.3330. URL: https://doi.org/10.1002/pro.3330.

[123] T. D. Goddard. "UCSF ChimeraX: Meeting modern challenges in visualization and analysis". In: *Protein Sci.: a Publ. Protein Soc.* 27 (2018). DOI: 10.1002/pro.3235. URL: https://doi.org/10.1002/pro.3235.

[124] O. S. Smart et al. "HOLE: a program for the analysis of the pore dimensions of ion channel structural models". In: *J. Mol. Graph.* 14 (1996). DOI: 10.1016/S0263-7855(97)00009-X. URL: https://doi.org/10.1016/S0263-7855(97)00009-X.

[125] L. Holm and C. Sander. "Dali: a network tool for protein structure comparison". In: *Trends Biochemical Sci.* 20 (1995). DOI: 10.1016/S0968-0004(00)89105-7. URL: https://doi.org/10.1016/S0968-0004(00)89105-7.

[126] A. M. Waterhouse et al. "Jalview Version 2-a multiple sequence alignment editor and analysis workbench". In: *Bioinforma. (Oxf., Engl.)* 25 (2009). DOI: 10.1093/bioinformatics/btp033. URL: https://doi.org/10.1093/bioinformatics/btp033.

[127] N. C. Pégard. "Three-dimensional scanless holographic optogenetics with temporal focusing (3D-SHOT)". In: *Nat. Commun.* 8 (2017). DOI: 10.1038/s41467-017-01031-3. URL: https://doi.org/10.1038/s41467-017-01031-3.

[128] Jiayue Su. "Structural basis of Tom20 and Tom22 cytosolic domains as the human TOM complex receptors". In: *PNAS* 119 (2022).

[129] Wenhe Wang. "Atomic structure of human TOM core complex". In: *Cell Discovery* 6 (2020).