Structure of the lens MP20 mediated adhesive junction

William J. Nicolas[1,2*], Anna Shiriaeva[1*], Michael W. Martynowycz[1*], Angus C Grey[3], Yasmeen Ruma[1,2], Paul J Donaldson[3] and Tamir Gonen[1,2,4$]

[1]Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

[2]Howard Hughes Medical Institute, University of California, Los Angeles, CA 90095

[3]Department of Physiology, School of Medical Sciences, University of Auckland, Auckland, NZ

[4]Department of Physiology, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

[*] Denotes equal contribution

[$]Correspondence: tgonen@g.ucla.edu

Keywords

Lens membrane protein, junction, MicroED, lipidic cubic phase, ion-beam milling, cryoEM

Abstract

Human lens fiber membrane intrinsic protein MP20 is the second most abundant membrane protein of the human eye lens. Despite decades of effort its structure and function remained elusive. Here, we determined the MicroED structure of full-length human MP20 in lipidic-cubic phase to a resolution of 3.5 Å. MP20 forms tetramers each of which contain 4 transmembrane α-helices that are packed against one another forming a helical bundle. Both the N- and C- termini of MP20 are cytoplasmic. We found that each MP20 tetramer formed adhesive interactions with an opposing tetramer in a head-to-head fashion. These interactions were mediated by the extracellular loops of the protein. The dimensions of the MP20 adhesive junctions are consistent with the 11 nm thin lens junctions. Investigation of MP20 localization in human lenses indicated that in young fiber cells MP20 was stored intracellularly in vesicles and upon fiber cell maturation MP20 inserted into the plasma membrane and restricted the extracellular space. Together these results suggest that MP20 forms lens thin junctions in vivo confirming its role as a structural protein in the human eye lens, essential for its optical transparency.

Introduction

In the lens, terminally differentiated fiber cells are derived from equatorial epithelial cells. These cells undergo a process of differentiation in which they dramatically elongate, lose their cellular organelles and nuclei and become internalized by the continual addition of new layers of differentiating fiber cells at the lens periphery[1]. Hence, the lens exhibits a highly ordered spatial gradient of cell age and differentiation, which makes it useful for studying how protein expression patterns change in aging. One such protein is MP20, a lens-specific membrane protein that is a member of the PMP22/EMP/MP20 subfamily of tetraspanins[2–6] and a distant member of the occludin family of junction forming proteins[7,8]. While MP20 is the second most abundant membrane protein in lens fiber cells[9] and several mutations in MP20 have been shown to result in cataract[10–18], the structure and function of MP20 are unknown. Previous studies have suggested that MP20 may be involved in signaling[19,20], due to its ability to interact with calmodulin, together with the identification of several phosphorylation sites on its cytoplasmic carboxy terminus[20,21]. In peripheral fiber cells, MP20 is found in cytoplasmic vesicles.[4,5] In mature fiber cells, MP20 becomes entirely membrane-associated[22,23] and in the rat lens this membrane insertion of MP20 correlates with a dramatic decrease in the extracellular spaces[22]. Furthermore, MP20 has also been shown to interact with galectin-3, a prominent cell adhesion modulator[24,25]. Together these suggest that MP20 may form adhesive junctions in vivo although both its structure and function of MP20 remain elusive.

Past efforts at determining the structure of MP20 yielded low resolution images that could not identify the MP20 function[26]. The predicted MP20 mass is 22kDa, but the literature often refers to the protein anywhere between 16-22kDa based on its appearance on SDS PAGE[24,27]. Either way, the MP20 size is way below the limit for structural characterization using single particle cryo-electron microscopy (cryo-EM)[27]. For this reason past efforts focused on negative stain electron microscopy and electron crystallography of two-dimensional crystals[26]. These studies showed that MP20 can oligomerize in membranes and form tetrameric assemblies, but its role remained unclear as the attainable resolution was limited and the crystals were too small and sparse for the available technologies. Efforts using X-ray crystallography likewise failed because MP20 crystals were far too small even by today's standards. Such vanishingly small crystals are amenable to structure determination by another more recent cryo-EM method known as Microcrystal Electron Diffraction, or MicroED[28,29]. With MicroED, crystals that are a billionth the size needed for X-ray crystallography can be interrogated by electron diffraction under cryogenic conditions. The crystal is continuously rotated in the electron beam while diffraction is recorded as a movie using a fast camera. This method has recently been enhanced by ion-beam milling with gallium[30,31] or plasma[32] to prepare optimally thick samples- either by making large crystals smaller, or by removing excess surrounding material around the microcrystals. This method has successfully elucidated structures such as the A_{2A} adenosine[32,33] and vasopressin 1B[34] receptors by initially tagging proteins with fluorophores, crystallizing them in a lipidic cubic phase (LCP), and using cryogenic fluorescence microscopy for targeting, followed by plasma beam milling and MicroED analysis.

Here we have expressed human MP20 heterologously, purified the protein to homogeneity and successfully grew very small crystals of the protein embedded in LCP. By combining fluorescence cryo-plasma Focused Ion Beam (cryo-pFIB) milling and MicroED, the structure of human MP20 was determined at a resolution of 3.5Å. The structure revealed the overall protein fold, and identified the role of MP20 in forming cell-to-cell adhesive junctions. Interrogation of the protein in native tissues suggested that in human lenses MP20 forms the 11nm "thin" junctions, a function that helps support lens architecture and structure.

Results & Discussion

Lens MP20 is a small membrane protein that on SDS PAGE appears as 18kDa[2]. We expressed human MP20 in SF9 insect cells and purified the protein by immobilized metal affinity chromatography (IMAC) as described in materials and methods. While the protein runs as a single band on reducing SDS PAGE consistent with ~20kDa, the FPLC trace suggests that it purifies as an oligomer of much larger size (Figures 1a andS1). Indeed, past reports have indicated that MP20 could form higher oligomers[26]. Considering the micelle size of the detergent that we used for purification, the FPLC trace likely corresponds to 8 x MP20 monomers. The purified protein was fluorescently labelled and mixed with lipids for crystallization in LCP[32–35]. We knew that lipids were imperative for this process based on past cryo-EM and electron crystallography studies we conducted on MP20 that was purified from native sources[26]. Some conditions indicated the presence of puncta when screened under cross-polarized light, suggesting that microcrystals of MP20 formed (Figure 1b, red arrowheads). The sample was deposited onto a cryo-EM grid as described before[32,34] and nanocrystals were initially identified using cryogenic fluorescence microscopy. Grids were thoroughly screened by cryo-fluorescence to identify the best grids and the best grid regions with the most promising crystals (Figure 1c). The best grids were then transferred into the cryo-pFIB/SEM with an integrated fluorescence microscope inside the chamber (iFLM) that allowed to recall the previously identified areas of interest (Figure 1d, red arrowhead). Following targeting and correlative light-EM methods that were described before[32,34] we were able to prepare crystalline lamella of MP20 crystals (Figure 1e-g) for MicroED analyses.

The grids were then transferred to a cryo-TEM for MicroED analyses. Using a highly attenuated electron source, an energy filter, and a sensitive direct electron detector we were able to collect 3.5 Å resolution MicroED data (Figure 1i, Video S1). While the reflections appeared strong at the beginning of the data collection, the intensity quickly dropped, likely due to the accumulation of radiation damage[36]. Indexing the data indicated that the crystal symmetry was $P\,4\,2\,2$ with unit cell dimensions of a=b= 56.18 Å and c= 142.50 Å. Typically a single nanocrystal would provide sufficient information to cover approximately 60° of reciprocal space so several data sets originating from 6 different nanocrystals had to be merged to increase completeness to ~90% (Table 1). Phasing of the MP20 data was accomplished using a poly-alanine model of an MP20 monomer generated by ColabFold[37], a cloud-based implementation of alphafold2[38], since no other experimental MP20 homologous structure was available at the protein data bank. Phasing results indicated a true space group of $P\,4\,2_1\,2$ with a single copy in the asymmetric unit. This creates an arrangement of four monomers around a central axis with an additional four monomers coordinating above it, creating an overall octamer (Figure S2).

MP20 has the overall topology consisting of four transmembrane helices (Ser3-Ala22, Ala59-Phe86, Ser100-Phe126, and Trp138-Cys166) connected by two extracellular loops, ECL1 and ECL2 (Met23-Ile58 and Leu127-Ser137), and one intracellular loop (Ala87-Phe99), with both its N and C termini localized intracellularly (Figure 2a, Video S2). This topology appears to be common in tetraspanin and the broader claudin family[39]. A long ECL1 domain connecting helices 1 and 2 is folded in the extracellular side of the membrane in an anti-parallel fashion presenting two subdomains that we name ECL1a and ECL1b. The four alpha-helices form a tight bundle, in contradiction with past reports suggesting that MP20 might form a channel[2,40]. A space filling model of MP20 indicated that no cavities or channel/path exists in a single MP20 monomer (Figure 2b and c). While it is possible that, at this resolution, we are not resolving the channel path, we note that for another lens membrane protein, namely aquaporin-0 (AQP0), even at a similar resolution a clear channel was observed[41]. These analyses and observations suggest to us that an MP20 monomer in this case does not form a channel, although it might form a proton channel which would not require a soluble path but rather a proton wire[42]. However, we could not locate a clear path for a continuous proton wire to form from the cytoplasmic to the extracellular side of MP20. Past electron crystallography analyses of reconstituted 2D crystals of MP20 indicated that the protein might assemble into tetramers[26]. Indeed, we can find an MP20 tetramer in our crystals. A close look of MP20 tetramers likewise does not identify any selectivity or path for substrates. This is again consistent with MP20 not forming a channel in this conformation.

When observing the full unit cell in our crystals it became clear that MP20 forms adhesive junctions. One MP20 tetramer was found interacting with an opposing MP20 tetramer from the opposing bilayer (Figure 3a and Figure S2). The interaction is mainly mediated through the loop regions of each MP20 monomer (Figure S3). ECL1a, ECL1b and ECL2 form an adhesive interaction resembling a handshake with the ECL1a and ECL1b and ECL2 from the opposing monomer. The interactions appear to be mainly electrostatic packing as several charged and hydroxylated residues can be seen forming adhesive interactions (Figure 3b, c, S3, Videos S3-S6). The ECL1b has cysteine residues Cys46 and Cys51 that are in close proximity and can potentially form a disulfide bonds in vivo. These cysteine residues are conserved within the PMP22/EMP/MP20 indicating that they may be functionally important (Figure S3 and Video S3, yellow dashed pseudobonds). In ECL1a residue Tyr52 is located at the interface, and in close proximity with Asn49 of the ECL1b of another MP20 molecule in the apposing bilayer. Lys50 in the same loop is directed towards a neighboring molecule in the same lipidic bilayer (Figure S3C, green dashed pseudobonds). ECL1b has polar residues Ser34, Ser36 that are directed towards the interface with the Thr55 in ECL2 of the opposing molecule. The smaller ECL2 loop, connecting transmembrane helices 3 and 4, has two consecutive arginine residues Arg129 and Arg130 involved in the contact with the ECL2 loop from another MP20 molecules. Phe131 in ECL2 is forming hydrophobic interactions with Tyr31 in ECL1b, Leu127 in ECL2, and Leu33 in the opposing molecule. The total thickness of the MP20 mediated membrane junction is approximately 11nm (Figure 3a), which is consistent with the lens "thin" 11nm junctions that have been described for decades[43].

We tested the ability of MP20 to form adhesive interactions by using vesicle reconstitution assays (Figure 3d and e). We formed vesicles made of DMPC alone using dialysis and in a parallel experiment reconstituted MP20 into DMPC vesicles and observed the preparations by negative stain electron microscopy. Vesicles made of lipid alone appeared evenly distributed on the electron microscopy grid (Figure 3d). TheMP20 containing vesicles clumped together rather than being evenly distributed on the grid (Figure 3e). This observation is consistent with the above postulate, suggesting that lens MP20 forms adhesive junctions in membranes.

Next, to assess whether MP20 forms junctions in vivo, we utilized an optimized fixation and sectioning protocol that allows sections through all regions of the human lens to be labelled with MP20 antibodies and imaged using fluorescence[44]. Past studies using this approach have identified MP20 throughout the lens and showed that its subcellular location changed during lens development. Using WGA, we labelled cell membranes in a human lens (Figure 4a). Three areas can be easily identified: a single layer of epithelium, which is followed by the fiber cells that make the bulk of the lens. The fiber cells in the outer cortex are arranged neatly in hexagonal arrays while deeper into the lens toward the core the fiber cells undergo remodeling and maturation and the hexagonal packing is lost. The subcellular localization of MP20 changes during these maturation events (Figure 4b). Initially, MP20 is found intracellularly as puncta inside the fiber cell cytoplasm, not on the plasma membrane. This is consistent with past studies that were done on lenses from rodents[22,25]. Deeper in the lens, MP20 labeling shifts from the intracellular cytoplasmic puncta to the plasma membrane.In the adult nucleus, all MP20 appears to be inserted into the plasma membrane.

We next asked whether adhesive junctions form upon MP20 insertion to the plasma membrane. We incubated the same lenses with Texas red Dextran which is known to penetrate lenses via the extracellular spaces[44]. Indeed this analysis indicated that labeling was observed between lens fiber cells in the outer cortex of the human lenses (Figure 4c). The dye still penetrated deeper into the lenses until it abruptly stopped right after the remodelling zone (Figure 4c, white line). An overlay of the MP20 labeling with the extracellular marker is presented in Figure 4d. This overlay indicates that upon insertion of MP20 into the plasma membrane of the lens fiber cells, the extracellular spaces between cells is dramatically constricted to the point where Texas red dextran could no longer diffuse. This restriction in the extracellular space does not appear to be diffusion limited, since increasing the incubation time did not increase the depth of Texas Red-dextran penetration past the adult nucleus region (Figure 4c, dotted line)[44]. This indicated that dye diffusion into the lens through the extracellular space was not limited by diffusion velocities but by a physical barrier that coincided with the insertion of MP20 into the membrane (Figure 4d). Together these results indicate that MP20 forms adhesive junctions between cells in vivo and that these junctions restrict the extracellular spaces in a fashion reminiscent of tight junctions.

Concluding Remarks

As the MicroED method matures and more technologies are developed and optimized for dealing with difficult samples, novel structures that were beyond the reach of other structural biology methods can now be determined experimentally. In this case, lens MP20 was subjected to intensive structural biology efforts for several decades by several laboratories worldwide employing all known structural biology methods including freeze fracture, thin section electron microscopy, single particle cryo-EM, electron crystallography, X-ray crystallography and even attempts using NMR, and AFM. In all these cases, only anecdotes were obtained and the MP20 enigma remained unsolved for half a century.

The MicroED structure presented here is the first experimentally determined structure of MP20, and more largely of any claudin-like protein involved in a tight junction, proving that in a lipid membrane MP20 can form cell-to-cell adhesive junctions. While each MP20 traversed the membrane 4 times, the protein assembled into tetramers in membranes and its extracellular loops formed adhesive interactions with an apposing MP20 in a head to head fashion. The interaction is best described as a handshake. The length of this junction is consistent with Lens thin 11nm junctions, which have been known to also contain AQP0 junctions[41,43]. However, in the case of MP20 it is still unknown whether the protein can also form a channel under some circumstances.

MP20 is a small membrane protein that belongs to the PMP22/EMP/MP20 family and is the first to be structurally characterized. This family is distantly related to the claudin superfamily of proteins that are known to form tight junctions[45] . While all these proteins contain 4 transmembrane helices, MP20 is the first to be determined in membranes and indeed the first to pack in the crystals as a junction. Small membrane proteins like MP20 are beyond the reach of single particle cryo-EM due to their size which is 3x smaller than the physical limit for cryo-EM (approximately 50 kDa)[46]. Recent work using single particle cryo-EM on claudin-4 inserted into micelles and bound to *Clostridium perfringens* enterotoxin were complexed with sFab to increase their size and facilitate imaging[47,48]. Even with this approach a relatively low-resolution image of the claudin was obtained and no junction observed since the Fab prevented junction formation. This illustrates the importance of solving membrane protein structures in the context of the lipid membrane rather than in detergents.

We also showed that the insertion of MP20 into cell membranes coincides with a constriction of the extracellular space *in vivo* in human lenses. It is presumed that MP20 therefore contributes to the structural stability of lens tissue and maintenance of transparency throughout lifetime. Future studies should delve into the structure of MP20 in its tetrameric form and analyze whether the protein can associate with other proteins and lipids to form channel complexes as has been observed in our previous lower resolution studies[26]. Further, a variety of congenital cataracts are associated with MP20 mutation[49], but it is not clear what is the structural impact of these mutations have on the protein and how they may affect its function.

While predictive methods like Alphafold[50] are increasing in popularity, the study presented here serves an important reminder to always validate models with experiments and, if possible, to experimentally determine structures. While Alphafold provided us with a reliable molecular replacement model of an MP20 monomer for phasing, even Alphafold3[50] which was marketed as a server for predicting protein interactions failed in predicting the oligomeric assembly of lens MP20 (Figure S4). When asked for predictive models for a dimer or for the octamer nonsensical assemblies were predicted with a high degree of certainty. In this case, since MP20 monomer is similar to the previously determined claudin[47] and an experimental structure exists in the protein data bank, the predicted model for the monomer was relatively correct. However, since the adhesive junction was never observed experimentally it is not surprising that the predictive power of Alphafold was limited. This serves as a reminder of the value of an experimental approach to structure determination and the limitation predictive methods still have when dealing with novel assemblies. It is clear that methods like MicroED will play a pivotal role in unraveling the biology behind MP20 and other small integral membrane proteins that are difficult to study with the more widely spread X-ray crystallography and cryo-EM techniques.

MATERIALS AND METHODS

Construct design. Several constructs with various positions of tags were screened.
The construct selected for crystallization due to successful expression and sufficient yield contained the following N-terminal tags: Hemagglutinin precursor peptide fragment MKTIIALSYIFCLV, FLAG (FADYKDDDDAK), LQTM, His$_{10,}$ PreScission cleavage site (LEVLFQ) (Figure S2). The construct used for the Fast Protein Liquid Chromatography (FPLC) is as follows: full-length MP20 cloned to the plasmid pcDNA3.1(+) (Genscript) with a C-terminal his tag and flag tag.

Protein expression and purification. MP20 constructs were expressed in sf9 insect cells using the Bac-to-bac baculovirus expression system (Invitrogen). Cells with a density of (2-3) × 10^6 cells mL^{-1} were infected with baculovirus at 27 °C at a multiplicity of infection of 10, harvested by centrifugation 48 hours post-infection and stored at −80 °C until use. Cells were lysed with hypotonic buffer (10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl). The membrane fraction was isolated from 2000 mL of biomass using Dounce homogenization and ultracentrifugation (25 min, 43000 g, 4 °C). The membranes were subsequently washed with high-salt galactose-containing buffer (1M NaCl, 10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 100 mM galactose) and isolated by ultracentrifugation (30 min, 43,000 g, 4 °C). Washed membranes were incubated in the hypotonic buffer in the presence of 2 mg mL^{-1} iodoacetamide for 20 min. MP20 was solubilized from membranes in a volume of 200 mL (4 × 50mL) by addition of 2× solubilization buffer (100 mM HEPES pH 7.5, 250 mM NaCl, 1 % (w/v) n-dodecyl-β-D-maltopyranoside (DDM, Anatrace) and 0.1% (w/v) cholesterol hemisuccinate (CHS, Sigma-Aldrich), 0.1% n-nonyl-β-D-glucopyranoside (NG, Anatrace)) for 2.5 h at 4 °C. Unsolubilized membranes were separated by centrifugation (60000g, 50 min). The supernatant was incubated overnight with 1 mL of Talon (immobilized metal affinity chromatography, IMAC) resin (Takara) in the presence of 10 mM imidazole. The sample was washed on a gravity column (Bio-Rad) with 10 column volumes (cv) of wash buffer #1 (50 mM HEPES pH 7.5, 800 mM NaCl, 25 mM imidazole pH 7.5, 10% (v/v) glycerol, 10 mM $MgCl_2$, 0.05%/0.01% (w/v) DDM/CHS) followed by 5 cv of wash buffer 2 (50 mM HEPES pH 7.5, 800 mM NaCl, 10% (v/v) glycerol, 50 mM imidazole, 0.025%/0.005% (w/v) DDM/CHS). Fluorescent labelling was carried out on column using Janelia Fluor 549 NHS-ester dissolved in DMF (4 mg mL^{-1}) and added in concentration 0.1% (v/v) to the labelling buffer (50 mM Hepes pH 7.5, 800 mM NaCl, 10% (v/v) glycerol, 0.025%/0.005% (w/v) DDM/CHS). Talon with MP20 was incubated with 2 cv of dye-containing labelling buffer for 3h, then washed cv by cv with 10 cv of the labelling buffer without dye and 3 cv of wash buffer #2. The sample was eluted in 500uL fractions using elution buffer (50 mM HEPES pH 7.5, 800 mM NaCl, 250 mM imidazole pH 7.5, 10% glycerol, 0.025%/0.005% (w/v) DDM/CHS). The sample was analyzed by HPLC (Shimadzu) (Figure S1) using a size exclusion column (Nanofilm SEC-250

Sepax). Fractions protein in micelles were concentrated to 20 uL using 100 kDa molecular weight cut-off protein concentrators (Amicon Ultra) at 1000 g for 4-6 h.

For the FPLC experiments, the MP20 construct was expressed in Expi293 cells (Thermo Fisher scientific) using the method described by Pleiner et al. [51]. The cells were grown in Expi293 expression media (Thermo Fisher scientific) to a density of 2 to 3 million cells /ml and then transfected with the MP-20 containing plasmid purified using a plasmid purification Giga kit (Qiagen). PEI MAX 40k (Polysciences, USA) was used as the transfection reagent. Cells were grown at 37 °C, 8% CO2 with 125 rpm and harvested after 68 to 72 hr followed by lysis in a sonicator with a lysis buffer (20 mM Tris, pH 7.5, 300 mM NaCl, 10 mM MgSO4, 10% glycerol, DNAse, protease inhibitor)[52]. Cell membranes were then separated by ultracentrifugation (60,000 rpm, 1hr, 4°C). MP20 was extracted from the cell membranes by incubating in solubilizing buffer (20 mM Tris, pH 7.5, 300 mM NaCl, 10 mM MgSO4, 10% glycerol, 1% DDM) at 4°C for 1 hr. The supernatant with the protein was then separated by ultracentrifugation (60,000 rpm, 45 min, 4°C) and incubated overnight with Talon metal affinity resin (Takara Bio). The protein was affinity purified from the solubilized fraction in a gravity column (BioRad) using 300 mM imidazole in elution buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 10% glycerol, 0.05% DDM) and further purified by SEC using a Superdex 200 10/300 GL column (GE Healthcare Life Sciences). Protein fractions were then analyzed on a SDS-PAGE gel (BioRad) and detected by western blotting using HRP-conjugated anti-his antibody (Genscript).

Crystallization of MP20 in LCP. Lipid cubic phase (LCP) was prepared by mixing concentrated protein in the ratio 2:3 (v/v) with molten lipid mix (90 % w/w monoolein and 10 % w/w cholesterol) using a syringe coupling system (Hamilton). Protein concentration was 15 mg.mL^{-1}. Crystallization was carried out using NT8 robotic dispensing system (Formulatrix). 40 nL LCP drops were dispensed at 85 % humidity and covered with 400 nL of precipitant solution. The plate was sealed and stored at room temperature. Crystals were detected in conditions with 25 - 100mM ammonium sulfate, 100mM ADA pH 5.6 - 6.0, 27 – 34 % PEG-400. Crystals appeared after 12 hours. Crystals were harvested using 20 - 50 µm nylon loops. Regions of the LCP drops were scooped and gently smeared across pre-clipped Cu200 R2:2 grids (Quantifoil) under a humidifier. Grids were flash-frozen in liquid nitrogen and stored frozen until use.

***Whole-grid fluorescence mapping*.** Shortly after vitrification, grids were batch-screened by cryo-Fluorescence Light Microscopy (cryo-FLM) using the Leica Thunder (Leica Microsystems, model DM6-FS) and the LAS X software (Leica Microsystems, version 3.7.6). For MP20 crystals labelled with Janelia Fluor 549, whole grid 3D atlases were acquired for all frozen grids in fluorescence (Y3 filter cube – 533 - 558 nm excitation / 565 nm dichroic / 570 - 640 nm emission). Several grids were also screened using reflection and/or transmitted light mode to allow for better correlation. 3D atlases were then carefully screened in LAS X after computational cleaning and deconvolution. Regions of interest were labelled. Labels and atlases were then exported for later use in MAPS software (ThermoFisher).

Plasma FIB milling of deeply embedded MP20 microcrystals. The grids selected for milling were inserted in the ThermoFisher Hydra FIB-SEM. The MAPs software (version 3.22) was used to make SEM atlases of the grids. The previously exported LAS X fluorescence atlases were imported into the MAPs software, including the labels and a fluorescence *in-focus* image. Both were realigned to the SEM atlas, allowing to target in X-Y the different crystals labelled during the prior step. The iFLM was then used to confirm presence of crystal and correlate its position as described below.

In-chamber fluorescent monitoring of crystals. The in-chamber Fluorescence Light Microscope (iFLM), from the original design of the PIEscope[53] was used to do the X-Y-Z correlation of the crystals to be milled (see below). It is situated coaxially to the FIB gun. Our iFLM system has an Olympus UPLZAPO 20x objective (0.8 numerical aperture, and working distance of 0.6mm), a Basler acA3088-57 detector with a physical pixel of 2.4 µm and an unbinned pixel size of 120nm. The light source is a 4-LED system from Thor Labs (385, 470,565 and 625 nm) working in conjunction with an epi-illuminator (Thor labs – WFA2001). The filter cube used is a multiband filter-set optimized for DAPI, FITC, TRITC and Cy5 (Semrock LED-DA/FI/TR/Cy5-B-000). It was operated with the Fluorescent Microscope Control v0.7 software (ThermoFisher). Image acquisition settings were the following: Bin 2 (240nm pixel size), 565 nm excitation wavelength at 100%, 100ms exposure time. When Z-stacks were acquired, the Z-step was set to either 0.5 or 1 µm for optimal compromise between fine Z-slicing and time to acquire stack. 3D fluorescent volumes were saved as .tiff stacks.

FIBucial correlation method. Normal-incidence FIB milling was used to create milled holes surrounding the sample (FIBucials). Beam settings were Ar – 5 s unblanking time – 7.6 nA – 30 kV. The area was then screened by iFLM and fluorescence of FIBucials was checked. When enough clearly fluorescent FIBucials were created (~5), a fluorescent stack with Z-step 1 µm was acquired. A normal-incidence and grazing incidence FIB (milling angle) image of the same area were then acquired where the FIBucials are also visible. The stack was then pre-processed using a custom Jupyter notebook. This entailed reslicing of the stack to have cubic pixels, normalization of intensities with the 3D correlation tool box (3DCT – https://github.com/hermankhfung/3DCT and https://github.com/hermankhfung/tools3dct)[54] library. Then, deconvolution was performed with the Red Lion Fish deconvolution (https://github.com/rosalindfranklininstitute/RedLionfish) library. This allowed to increase the signal-to-noise ratio, especially important for correct Z-estimation of the FIBucials and the crystal[32]. Next custom Jupyter Notebook took the pre-processed stack and opened it with Napari software (https://github.com/napari/napari). The x, y, z coordinates of the FIBucials are registered by adding a New Points layer. The Z coordinate is where the FIBucial is seen the sharpest. The coordinates of the crystal are registered the same way and the layer of points is saved. The x and y coordinates of the FIBucials on the FIB/SEM images are registered and saved the same way. This script then converts the Napari-generated coordinates to fit the 3DCT coordinate system and the 3DCT GUI is started with the fluorescent stack and the FIB/SEM image along with their list of coordinates. The software is then used to predict the position of the crystal on the FIB image and compute the error of the prediction. An error of < 10 pixels was sufficient to prevent over milling of crystal.

FIB milling of crystals. All milling was done with the Argon beam at 30kV. In between each milling step, the iFLM was used to assess the presence and integrity of the targeted crystal. The predicted position of the crystal done with the method described above allowed us to make an informed decision as to where to position our initial trenches. Initial trenches were spaced 15 µm apart with the targeted crystal centered in the initial slab. This step was performed at 7.6 nA, a box depth ranging from 20 - 40 µm and a box width of 20 um. Box height depended on the topology of the region. What is important is that the trenches clear out everything above and below the initial slab, meaning un-obstructed holes visible in FIB and SEM. Next step consisted in rough milling down to a 5 µm thick slab at 2 nA. Because the Z-resolution provided by our iFLM ranges between 5 and 7 µm, we considered that our predicted crystal position was no-good below 5um. We therefore proceeded by shaving only the top of the lamella with 10 µm wide and 0.5 to 1um high milling boxes and checking the presence of the crystal with the SEM beam in immersion mode, using the through lens detector at 1.2 kV. When the crystal became visible by contrast difference, the bottom of the lamella was milled until final 200-300 nm thickness. 0.74 nA, 0.2 nA and 60 pA currents were used until lamella was 1um, 500 nm and 200-300 nm thick, respectively. The top of the lamella was then polished very lightly at 20 pA for ~30 s, presence of fluorescence on the lamella is checked one last time before immediate transfer to Krios for MicroED data acquisition.

MicroED data collection. Milled grids were loaded into a cryogenically cooled Titan Krios G3i transmission electron microscope operating at an accelerating voltage of 300 kV. All grid atlases were taken at 155 × magnification using SerialEM[55] to identify milled lamellae on the grid. Each lamella was brought to eucentric height. The crystals were not visible using low or high SA imaging at low dose due to the low contrast between the crystal and the surrounding LCP. The crystals were located by initially correlating the iFLM fluorescence and the SEM low-voltage images with the TEM low SA (3600 ×) images. Optionally, a custom SerialEM script was used to screen sub-areas of the lamellae in 2 µm steps in a grid pattern to find the best diffraction. Continuous rotation MicroED data were collected using an in-house dedicated SerialEM script, on a Falcon 4i direct electron detector in electron counting mode using the electron event representation format (.eer)[56]. A single dataset was collected from each crystal over a wedge of 60 degrees over 420s with a nominal camera length of 2500mm.

Structure determination and refinement of the MP20 structure. Individual movies in EER format were converted to SMV format using MicroED-tools mrc2smv (available from cryoEM.ucla.edu)[57]. EER files contained approximately 130,000 individual event frames after taking dark frames into account. SMV frames were created by summing 308 individual EER frames, or 1s of counting data, correcting for post-counting gain, and binning the data by 2 in both the x and y dimension. In this way, each 420s EER movie resulted in 420 2048 x 2048 SMV formatted images containing 16-bit unsigned integers. MicroED data on SMV frames were processed in XDS[58] where the data were indexed and integrated. Datasets were scaled using XSCALE. The data were indexed in space group P 4 2 2 (#89) with unit cell (a, b, c) (Å) = (56.18, 56.18, 142.50) and (α, β, γ) (°) = (90, 90, 90). The resolution was cut to 3.5 Å, where the $CC_{1/2}$ value fell to a value of approximately 30%. The structure of MP20 was determined by molecular replacement. The search model was an Alphafold[38] structure truncated to a polyalanine chain. The molecular replacement search was accomplished using PHASER[59] searching in all possible space groups, where the final model was placed in space group P4 2_1 2 (#90) with a Translation Function Z-score (TFZ) of 20 and Log Likelihood Gain (LLG) of 700. The model was initially refined in PHENIX.REFINE[60] using electron scattering factors with group B-factors and rigid body refinement. The sidechains were docked into the trace using the known sequence. The updated model was refined iteratively using COOT[61] and PHENIX.REFINE until the model could no longer be improved. The final R-work and R-free were 33.16% and 35.10%, respectively.

Data availability. The coordinates of the MicroED structure of MP20 were deposited to the PDB and the corresponding potential maps deposited to the EMDB under accession codes XXXX and EMD-XXXX, respectively. The MicroED tools mrc2smv software is available freely at cryem.ucla.edu. Additional materials or code are available upon request to the lead contact.

Human lenses. Lenses of a range of ages were obtained from donor eyes from the New Zealand National Eye Bank within 24-48 hours of death. All human lens work was conducted in compliance with the Declaration of Helsinki and was approved by the Northern X Regional ethics Committee (Ref: NTX/07/08/079). Corneas were removed for transplantation and the lens removed and immediately assessed by dark and bright field microscopy and then either fixed immediately for immunohistochemical experiments or first incubated in Artificial Aqueous Humor (125mM NaCl, 0.5mM $MgCl_2$, 4.5mM KCl, 10mM $NaHCO_3$, 2mM $CaCl_2$, 5mM glucose, 20mM sucrose, 10mM HEPES, pH 7.2-7.4, 300±5 mOsmol) that contained 1.25 mg/ml of Texas red-dextran (Molecular Probes, Eugene, OR, USA), for 6 hours or 24 hours before being fixed as described below.

Immunohistochemistry. Whole lenses were fixed for 24 hours in 0.75% paraformaldehyde in PBS, encased in 6% agarose and then either cut in half perpendicular (equatorial) or parallel (axial) to the optic axis using a sharp blade. The halved lenses were fixed for 24 hours in 0.75% paraformaldehyde and then cryoprotected by incubation in 10% sucrose-PBS for 1 hour, 20% sucrose-PBS for 1 hour and then 30% sucrose-PBS overnight at 4^0C. For sectioning, half lenses were mounted in an equatorial or axial orientation on pre-chilled chucks and encased in Tissue-Tek O.C.T compound. Lenses were cryosectioned into 12-14 µm thick sections and transferred onto poly-lysine coated microscope slides (Superfrost Plus; ESCO, Electron Microscopy Sciences, Fort Washington, Pennsylvania. USA). Sections were washed three times in PBS and then incubated in blocking solution (3% BSA, 3% normal goat serum (NGS)) for 1 hour to reduce non-specific labelling. Sections were labelled with a carboxyl tail specific MP20 antibody[40] that was diluted in PBS, followed by secondary fluorescein-conjugated antibody (1:200; Santa Cruz Biotechnology, Santa Cruz, California, USA) for 1 hour each[22]. Control sections omitting primary antibody were also prepared. Cell membranes were also labelled with a fluorescein-conjugated wheat germ agglutinin (WGA-TRITC). The WGA-TRITC- was diluted 1:50 in PBS and the sections were incubated for 1 hour. After extensive washing in PBS, sections were mounted in anti-fading reagent (Citifluor™, AFI, Canterbury, UK) and viewed with a confocal laser scanning microscope (Leica TCS 4D, Heidelberg, Germany). Hoescht stain was used to highlight the nuclei. Images were pseudo-coloured and combined using Adobe Photoshop software.

ACKNOWLEDGEMENTS

This study was supported by the National Institutes of Health P41GM136508. The Gonen laboratory is supported by funds from the Howard Hughes Medical Institute. PJD is funded by New Zealand Health Research Council Programme Grant.

Author Contributions

AS expressed, purified and crystallized human MP20. YR performed FPLC analyses. WJN prepared crystal lamellae, collected and processed the MicroED data and participated in structure analyses. MW processed MicroED data and participated in structure analyses. PJD and ACG performed fluorescence analyses on human lenses. TG and PJD conceived the project and participated in analyses. All authors participated in writing the manuscript and figure preparation.

FIGURE LEGENDS

Figure 1: Expression, purification and crystallization of human lens MP20.

a. FPLC trace of MP20 with accompanying SDS-PAGE gel (right lane) and western blot (left lane). Arrows indicate the MP20 eluted as a uniform peak consistent in size with an MP20 octamer (8 x MP20 plus two detergent micelles). Inset SDS-PAGE of the peak showing purified MP20 and faint bands for its higher molecular weight oligomers (Lane 1) and the corresponding western blot (Lane 2). b. Cross-polarized light image of a typical drop of LCP with small MP20 crystals appearing as tiny puncta (red arrowheads). c. Region of a whole grid fluorescent atlas showing the LCP with JaneliaFluor 550 conjugated crystals (yellow signal). d. Same crystal boxed out in red in panel c. but imaged in the SEM with the iFLM (red arrowhead). FIBucial markers created with the FIB beam are also visible (green dots and arrows). e-g. SEM top-down image (e) and FIB grazing incidence image (f) of the same field of view (FOV) as in (d) with reprojected coordinates of the crystal (red markers and arrowhead) and the FIBucials (green arrowheads). f. shows an inset of the boxed-out region after milling was performed to create a lamella on the targeted crystal. g. Overlay of the SEM view of the final lamella with the fluorescence signal emanating from the crystal on the milled lamella, confirming presence of the crystal. i. An example of a diffraction pattern from the MicroED dataset collected on the crystal lamella shown in panel g. Blue dashed ring represents diffraction limit for this dataset (3.25Å).

Figure 2 : MicroED structure of the lens MP20

a. Structure of lens MP20 in rainbow with the N terminus in blue and C-terminus in red. Both termini are cytoplasmic. The loops are indicated as ECL1a, ECL1b, ECL2 and ICL. b and c. Space filling model (half is clipped in c.) of the MP20 monomer showing that no channel or pathway can be seen in the protein. The black lines indicate the position of the lipid bilayer with the extracellular and intracellular sides as indicated.

Figure 3 : MicroED structure of the lens MP20 junctions

a. MP20 mediated membrane junction. Two MP20 tetramers interact in a head to head fashion (yellow and blue). The adhesive interactions are mediated by the extracellular loops. b. MicroED density map around one yellow and one blue MP20 monomer as they form the adhesive interaction. c. two monomers shown in rainbow with the N terminus in blue and C terminus in red. The black lines indicate the position of the lipid bilayer with the extracellular and intracellular sides as indicated. d and e. Vesicles without and with MP20, respectively. Vesicles without MP20 are evenly distributed on the EM grid while the vesicles with MP20 aggregated. Scale bar = 200nm

Figure 4 : Insertion of MP20 into fiber cell membranes correlates with the constriction of the extracellular space in vivo

Human lenses labelled with a. WGA b. MP20 c. extracellular dye. d. overlay of b and c. Image montage of the outer cortical region of cryosection taken through the equator of a human donor lenses organ cultured in the extracellular space marker Texas Red-dextran for 6 hours. a. WGA-TRITC labelling to highlight the change in membrane morphology as fiber cells differentiate and become internalized into the adult nucleus. b. MP20 labelling showing the shift form the cytoplasm to the membrane. c. Texas Red-dextran labelling of the extracellular space. White line showing where the dye stops and the extracellular spaces become restricted. d. Double labelling of MP20 (green) and Texas Red-dextran (red) showing the formation of the extracellular diffusion barrier corelates with the membrane insertion of MP20 into the membranes of fiber cells in the adult nucleus. e. Schematic diagram (not drawn to scale) summarizing changes in the morphology of differentiating fiber cells in the outer cortex of the human lens. DF = differentiating fiber cells; RZ = remodeling zone; TZ = transition zon

Supplementary Figure 1 | Purification of MP20 protein for crystallization

a. MP20 primary amino acid sequence used for expression and MP20 purification for crystallization. b. High Performance Liquid Chromatography (HPLC) profile (dead volume was cut out). The MP20 absorbance peak can be seen at 27mL elution volume (red dashed line). c. SDS-PAGE of the peak with visible bands at MP20 molecular weight.

Supplementary Figure 1 | Packing of MP20 crystals

a. Side view of MP20 crystal octamer mediated by the loop interactions. b. Top view of the MP20 crystal octamer

Supplementary Figure 3: Adhesive interactions in the head-head MP20 junctions

a. Side view of tetramers inserted in their respective membrane bilayers interacting with an apposing tetramer, forming an octamer. One interaction involving two apposed protomers is highlighted (yellow and blue). Lateral chains involved in the putative interactions are made visible in the loops. Loops and N/C-termini are labelled. b. Magnified view of the red inset in a. with a 90° rotation of the molecule towards the right, showing the interface between the ECLs from the protomers involved in the “hand shake” interaction. ECLs are also labelled. Putative intra, inter-chain electrostatic interactions and di-sulfide bridges are shown in purple, green and yellow dashed lines, respectively. The amino acids involved in those putative interactions are labelled. c. 90° rotation of the molecule in b. downward showing the side-by-side interactions between two protomers of a dimer within one tetramer. The amino acids involved in those putative interactions are labelled and traced in green. The protomers in white belong to the bottom and top tetramer shown in a. and interact with the yellow and blue protomer, respectively.

Supplementary Figure 4: Alphafold3 predictive models for lens MP20. a. Predictive model of the monomer. b.predicted model of a dimer. c. Predicted model of an octamer.

Supplementary Video 1

MicroED dataset diffracting to 3.25A acquired on an MP20 crystalline lamella.

Supplementary Video 2

360° rotational video of the MP20 monomer model in rainbow colors.

Supplementary Video 3

360° rotational video of two MP20 monomers involved in a head-head interaction, in rainbow colors.

Supplementary Video 4

360° rotational video of two MP20 tetramers (yellow and blue) involved in head-head interaction, forming an octamer.

Supplementary Video 5

360° rotational video of an MP20 octamer showing its steric hindrance.

Supplemental Video 6

Detailed rocking view of the head-head interaction with the lateral chains involved in putative contacts mentioned in manuscript made visible. Putative intra, inter-chain electrostatic interactions and di-sulfide bridges are shown by protomer color (yellow or blue), green and red dashed lines.

REFERENCES

1. Audette, D. S., Scheiblin, D. A. & Duncan, M. K. The molecular mechanisms underlying lens fiber elongation. *Experimental Eye Research* **156**, 41–49 (2017).
2. Louis, C. F. *et al*. Identification of an 18,000-dalton protein in mammalian lens fiber cell membranes. *Journal of Biological Chemistry* **264**, 19967–19973 (1989).
3. Taylor, V., Welcher, A. A. & Suter, U. Epithelial membrane protein-1, peripheral myelin protein 22, and lens membrane protein 20 define a novel gene family. *Journal of Biological Chemistry* **270**, 28824–28833 (1995).

4. Welcher, A. A., Suter, U., De Leon, M., Snipes, G. J. & Shooter, E. M. A myelin protein is encoded by the homologue of a growth arrest-specific gene. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7195–7199 (1991).
5. Lobsiger, C. S. *et al.* Identification and characterization of a cDNA and the structural gene encoding the mouse epithelial membrane protein-1. *Genomics* **36**, 379–387 (1996).
6. Taylor, V. & Suter, U. Epithelial membrane protein-2 and epithelial membrane protein-3: two novel members of the peripheral myelin protein 22 gene family. *Gene* **175**, 115–120 (1996).
7. Asano, A., Asano, K., Sasaki, H., Furuse, M. & Tsukita, S. Claudins in Caenorhabditis elegans: their distribution and barrier function in the epithelium. *Curr Biol* **13**, 1042–1046 (2003).
8. Simske, J. S. Claudins reign: The claudin/EMP/PMP22/γ channel protein family in C. elegans. *Tissue Barriers* **1**, e25502 (2013).
9. Alcala, J., Lieska, N. & Maisel, H. Protein composition of bovine lens cortical fiber cell membranes. *Experimental eye research* **21**, 581–595 (1975).
10. Pras, E. *et al.* A missense mutation in the LIM2 gene is associated with autosomal recessive presenile cataract in an inbred Iraqi Jewish family. *The American Journal of Human Genetics* **70**, 1363–1367 (2002).
11. Ponnam, S. P. G., Ramesha, K., Tejwani, S., Matalia, J. & Kannabiran, C. A missense mutation in LIM2 causes autosomal recessive congenital cataract. *Molecular vision* **14**, 1204 (2008).
12. Irum, B. *et al.* Mutation in LIM2 is responsible for autosomal recessive congenital cataracts. *PloS one* **11**, e0162620 (2016).
13. Pei, R. *et al.* A novel mutation of LIM2 causes autosomal dominant membranous cataract in a Chinese family. *International Journal of Ophthalmology* **13**, 1512 (2020).
14. Berry, V. *et al.* A novel missense mutation in *LIM2* causing isolated autosomal dominant congenital cataract. *Ophthalmic Genetics* **41**, 131–134 (2020).
15. Berry, V. *et al.* A recurrent variant in *LIM2* causes an isolated congenital sutural/lamellar cataract in a Japanese family. *Ophthalmic Genetics* **43**, 622–626 (2022).
16. Zhou, Z. *et al.* Genetic variations in GJA3, GJA8, LIM2, and age-related cataract in the Chinese population: a mutation screening study. *Molecular vision* **17**, 621 (2011).

17. Fernández-Alcalde, C. *et al.* Molecular and genetic mechanism of non-syndromic congenital cataracts. mutation screening in Spanish families. *Genes* **12**, 580 (2021).
18. Wang, X. *et al.* Elongated axial length and myopia-related fundus changes associated with the Arg130Cys mutation in the LIM2 gene in four Chinese families with congenital cataracts. *Annals of Translational Medicine* **9**, (2021).
19. Louis, C. F., Johnson, R., Johnson, K. & Turnquist, J. Characterization of the bovine lens plasma membrane substrates for cAMP-dependent protein kinase. *European Journal of Biochemistry* **150**, 279–286 (1985).
20. Louis, C. F., Hogan, P., Visco, L. & Strasburg, G. Identity of the calmodulin-binding proteins in bovine lens plasma membranes. *Experimental eye research* **50**, 495–503 (1990).
21. Ervin, L. A., Ball, L. E., Crouch, R. K. & Schey, K. L. Phosphorylation and glycosylation of bovine lens MP20. *Investigative ophthalmology & visual science* **46**, 627–635 (2005).
22. Grey, A. C., Jacobs, M. D., Gonen, T., Kistler, J. & Donaldson, P. J. Insertion of MP20 into lens fibre cell plasma membranes correlates with the formation of an extracellular diffusion barrier. *Experimental eye research* **77**, 567–574 (2003).
23. Shi, Y. *et al.* The stratified syncytium of the vertebrate lens. *Journal of cell science* **122**, 1607–1615 (2009).
24. Gonen, T., Donaldson, P. & Kistler, J. Galectin-3 is associated with the plasma membrane of lens fiber cells. *Investigative ophthalmology & visual science* **41**, 199–203 (2000).
25. Gonen, T., Grey, A. C., Jacobs, M. D., Donaldson, P. J. & Kistler, J. MP20, the second most abundant lens membrane protein and member of the tetraspanin superfamily, joins the list of ligands of galectin-3. *BMC Cell Biol* **2**, 17 (2001).
26. Gonen, T. *et al.* Polymorphic assemblies and crystalline arrays of lens tetraspanin MP20. *Journal of molecular biology* **376**, 380–392 (2008).
27. Glaeser, R. M. How good can cryo-EM become? *Nature methods* **13**, 28–32 (2016).
28. Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *elife* **2**, e01345 (2013).
29. Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat Methods* **11**, 927–930 (2014).

30. Martynowycz, M. W., Zhao, W., Hattne, J., Jensen, G. J. & Gonen, T. Collection of Continuous Rotation MicroED Data from Ion Beam-Milled Crystals of Any Size. *Structure* **27**, 545-548.e2 (2019).
31. Martynowycz, M. W., Zhao, W., Hattne, J., Jensen, G. J. & Gonen, T. Qualitative Analyses of Polishing and Precoating FIB Milled Crystals for MicroED. *Structure* **27**, 1594-1600.e2 (2019).
32. Martynowycz, M. W. *et al.* A robust approach for MicroED sample preparation of lipidic cubic phase embedded membrane protein crystals. *Nat Commun* **14**, 1086 (2023).
33. Martynowycz, M. W. *et al.* MicroED structure of the human adenosine receptor determined from a single nanocrystal in LCP. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2106041118 (2021).
34. Shiriaeva, A., Martynowycz, M. W., Nicolas, W. J., Cherezov, V. & Gonen, T. MicroED structure of the human vasopressin 1B receptor. *bioRxiv* (2023).
35. Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences* **93**, 14532–14535 (1996).
36. Hattne, J. *et al.* Analysis of Global and Site-Specific Radiation Damage in Cryo-EM. *Structure* **26**, 759-766.e4 (2018).
37. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679–682 (2022).
38. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
39. Krause, G. *et al.* Structure and function of claudins. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1778**, 631–645 (2008).
40. Arneson, M. L. & Louis, C. F. Structural Arrangement of Lens Fiber Cell Plasma Membrane Protein MP20. *Experimental Eye Research* **66**, 495–509 (1998).
41. Gonen, T., Sliz, P., Kistler, J., Cheng, Y. & Walz, T. Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature* **429**, 193–197 (2004).
42. Agmon, N. The grotthuss mechanism. *Chemical Physics Letters* **244**, 456–462 (1995).
43. Costello, M. J., McIntosh, T. J. & Robertson, J. D. Distribution of gap junctions and square array junctions in the mammalian lens. *Investigative Ophthalmology & Visual Science* **30**, 975–989 (1989).

44. Lim, J. C., Walker, K. L., Sherwin, T., Schey, K. L. & Donaldson, P. J. Confocal Microscopy Reveals Zones of Membrane Remodeling in the Outer Cortex of the Human Lens. *Investigative Ophthalmology & Visual Science* **50**, 4304–4310 (2009).
45. McGuinness, S., Sajjadi, S., Weber, C. R. & Khalili-Araghi, F. Computational Models of Claudin Assembly in Tight Junctions and Strand Properties. *International Journal of Molecular Sciences* **25**, 3364 (2024).
46. Lander, G. C. & Glaeser, R. M. Conquer by cryo-EM without physically dividing. *Biochem Soc Trans* **49**, 2287–2298 (2021).
47. Erramilli, S. K., Dominik, P. K., Ogbu, C. P., Kossiakoff, A. A. & Vecchio, A. J. Cryo-EM structures of a synthetic antibody against 22 kDa claudin-4 reveal its complex with Clostridium perfringens enterotoxin. 2023.06.12.544689 Preprint at https://doi.org/10.1101/2023.06.12.544689 (2023).
48. Orlando, B. J. *et al.* Development, structure, and mechanism of synthetic antibodies that target claudin and *Clostridium perfringens* enterotoxin complexes. *Journal of Biological Chemistry* **298**, 102357 (2022).
49. Berry, V. *et al.* A recurrent variant in LIM2 causes an isolated congenital sutural/lamellar cataract in a Japanese family. *Ophthalmic Genet* **43**, 622–626.
50. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 1–3 (2024) doi:10.1038/s41586-024-07487-w.
51. Pleiner, T. *et al.* WNK1 is an assembly factor for the human ER membrane protein complex. *Molecular Cell* **81**, 2693-2704.e12 (2021).
52. Krasnoselska, G. O. *et al.* Transient Transfection and Expression of Eukaryotic Membrane Proteins in Expi293F Cells and Their Screening on a Small Scale: Application for Structural Studies. *Methods Mol Biol* **2305**, 105–128 (2021).
53. Gorelick, S. *et al.* PIE-scope, integrated cryo-correlative light and FIB/SEM microscopy. *Elife* **8**, e45919 (2019).
54. Arnold, J. *et al.* Site-Specific Cryo-focused Ion Beam Sample Preparation Guided by 3D Correlative Microscopy. *Biophysical Journal* **110**, 860–869 (2016).
55. De La Cruz, M. J., Martynowycz, M. W., Hattne, J. & Gonen, T. MicroED data collection with SerialEM. *Ultramicroscopy* **201**, 77–80 (2019).

56. Martynowycz, M. W., Clabbers, M. T. B., Hattne, J. & Gonen, T. Ab initio phasing macromolecular structures using electron-counted MicroED data. *Nat Methods* **19**, 724–729 (2022).
57. Hattne, J., Clabbers, M. T. B., Martynowycz, M. W. & Gonen, T. Electron counting with direct electron detectors in MicroED. *Structure* **31**, 1504-1509.e1 (2023).
58. Kabsch, W. xds. *Acta Crystallographica Section D: Biological Crystallography* **66**, 125–132 (2010).
59. McCoy, A. J. *et al. Phaser* crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
60. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213–221 (2010).
61. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica section D: biological crystallography* **60**, 2126–2132 (2004).

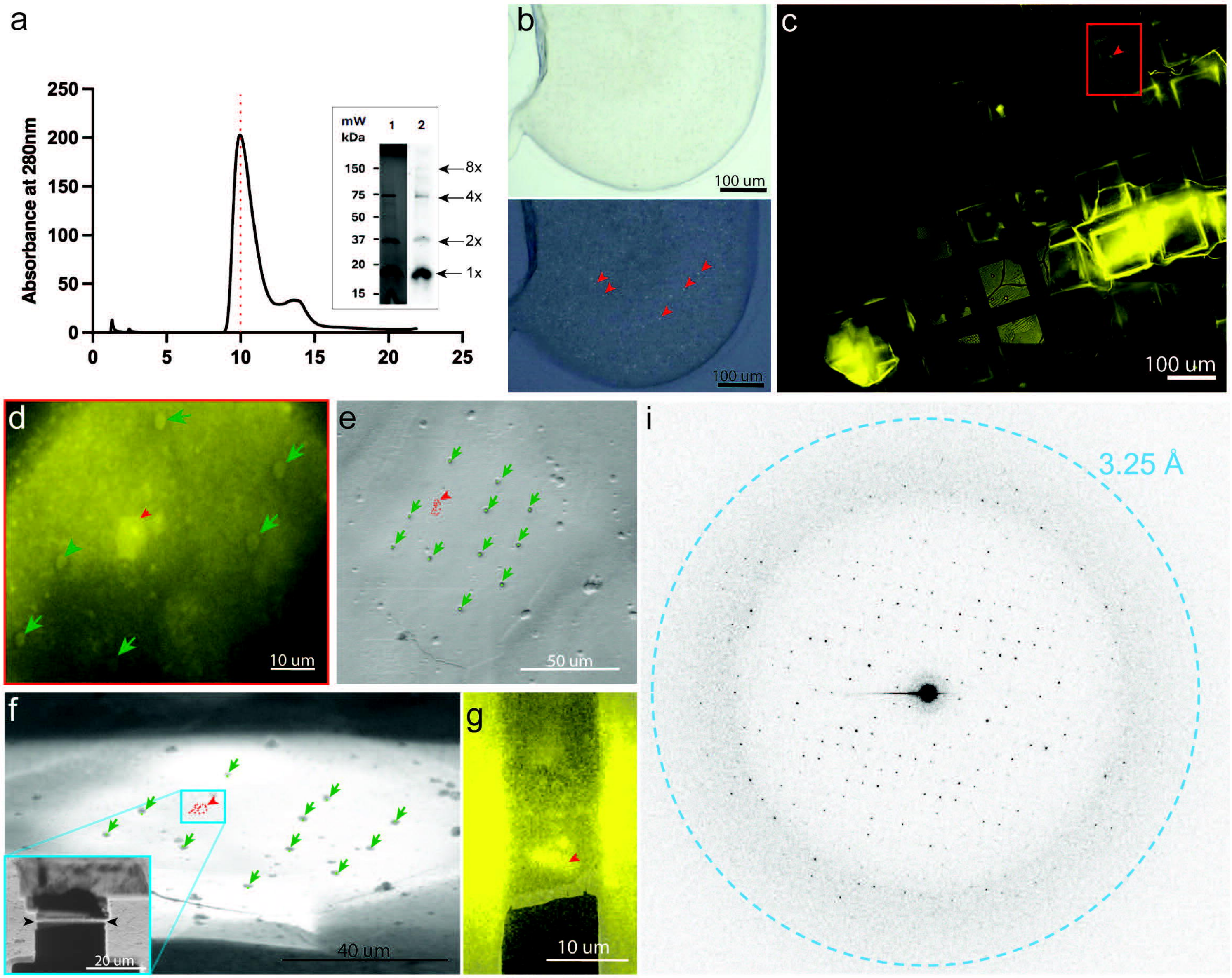

a
Absorbance at 280nm
250
200
150
100
50
0
0
5
10
15
20
25
mW
kDa
1
2
150
75
50
37
20
15
8x
4x
2x
1x
b
100 um
100 um
c
100 um
d
10 um
e
50 um
f
20 um
40 um
g
10 um
i
3.25 Å

Figure 2
MicroED Structure of Lens MP20

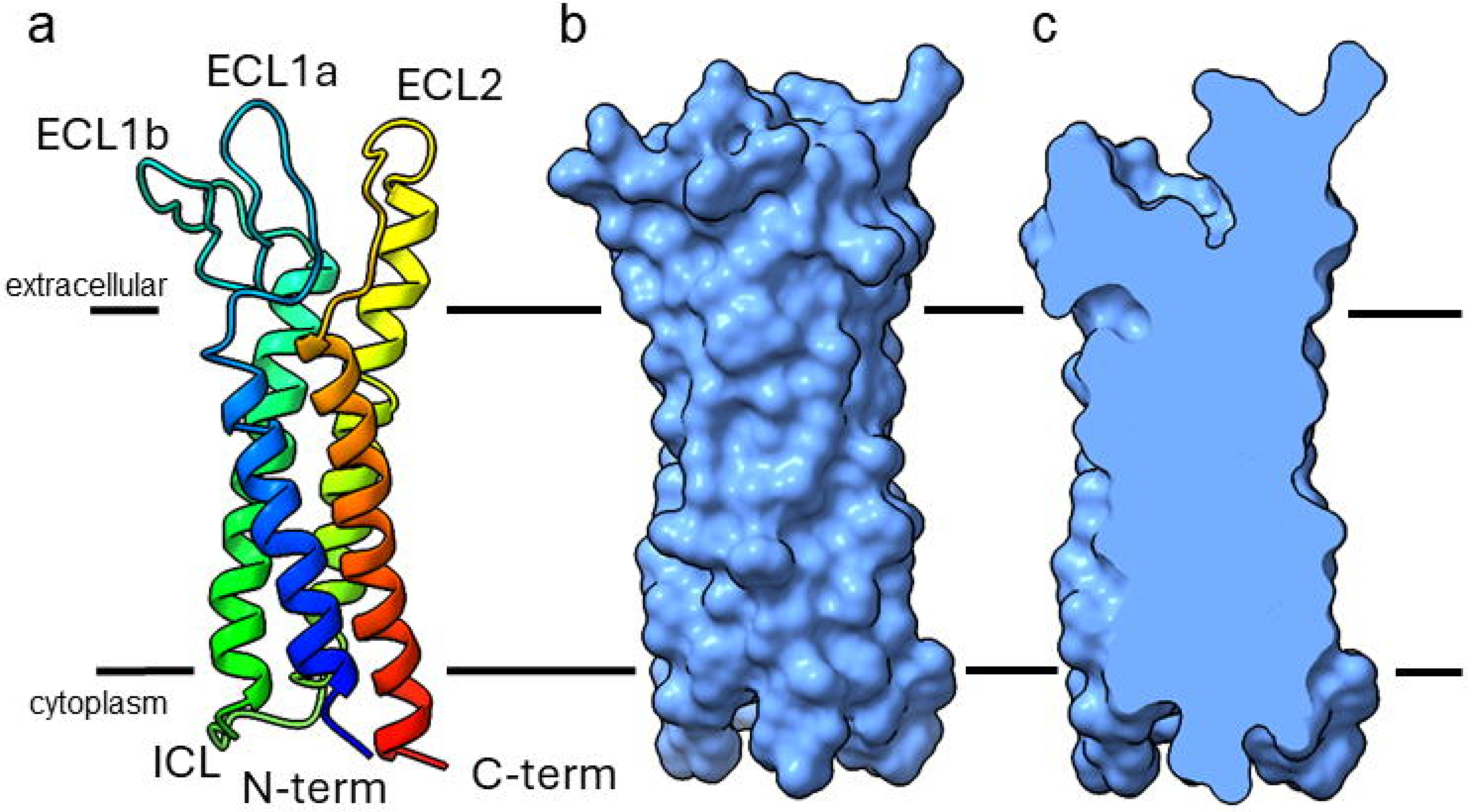
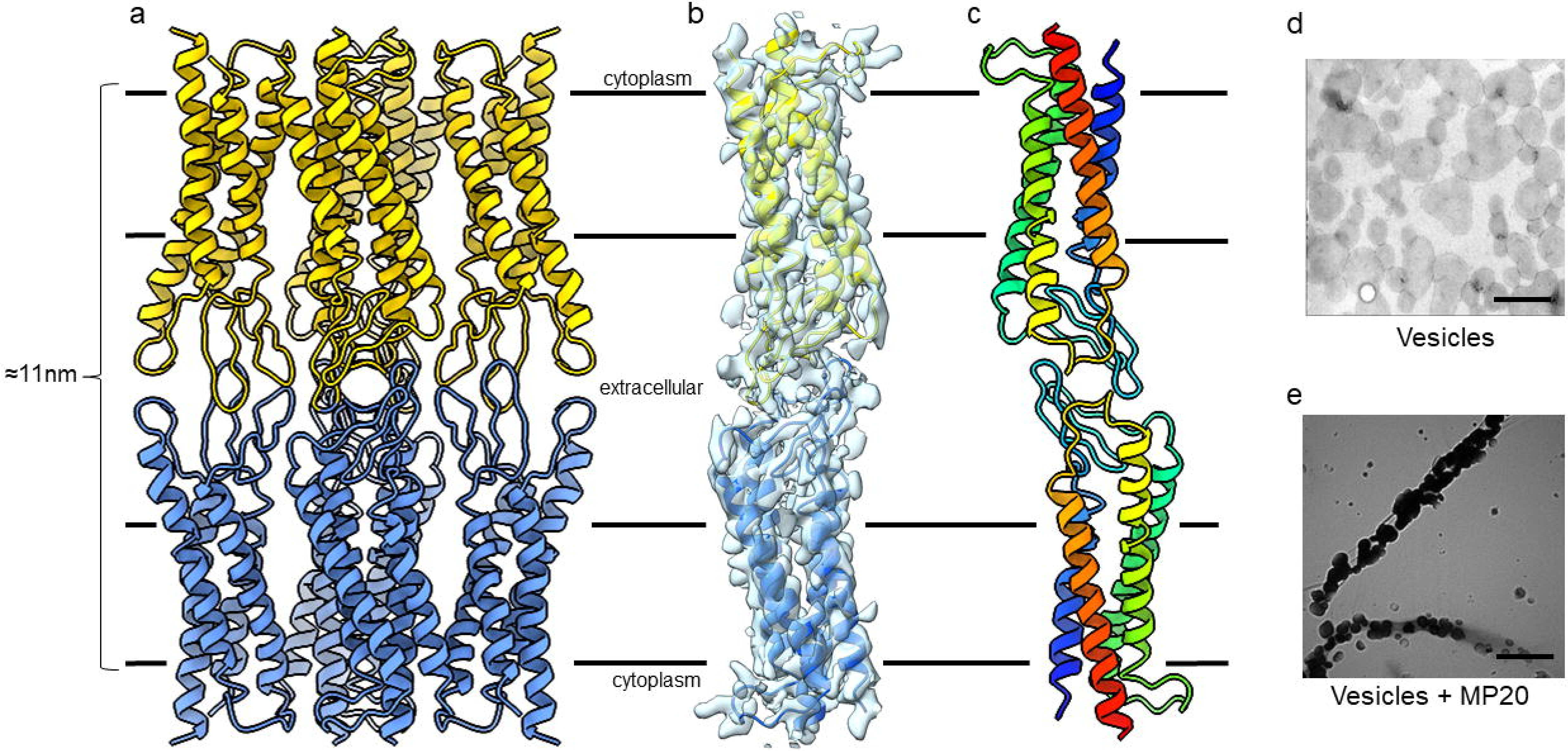


Figure 3

MicroED structure of lens thin MP20 juncitons

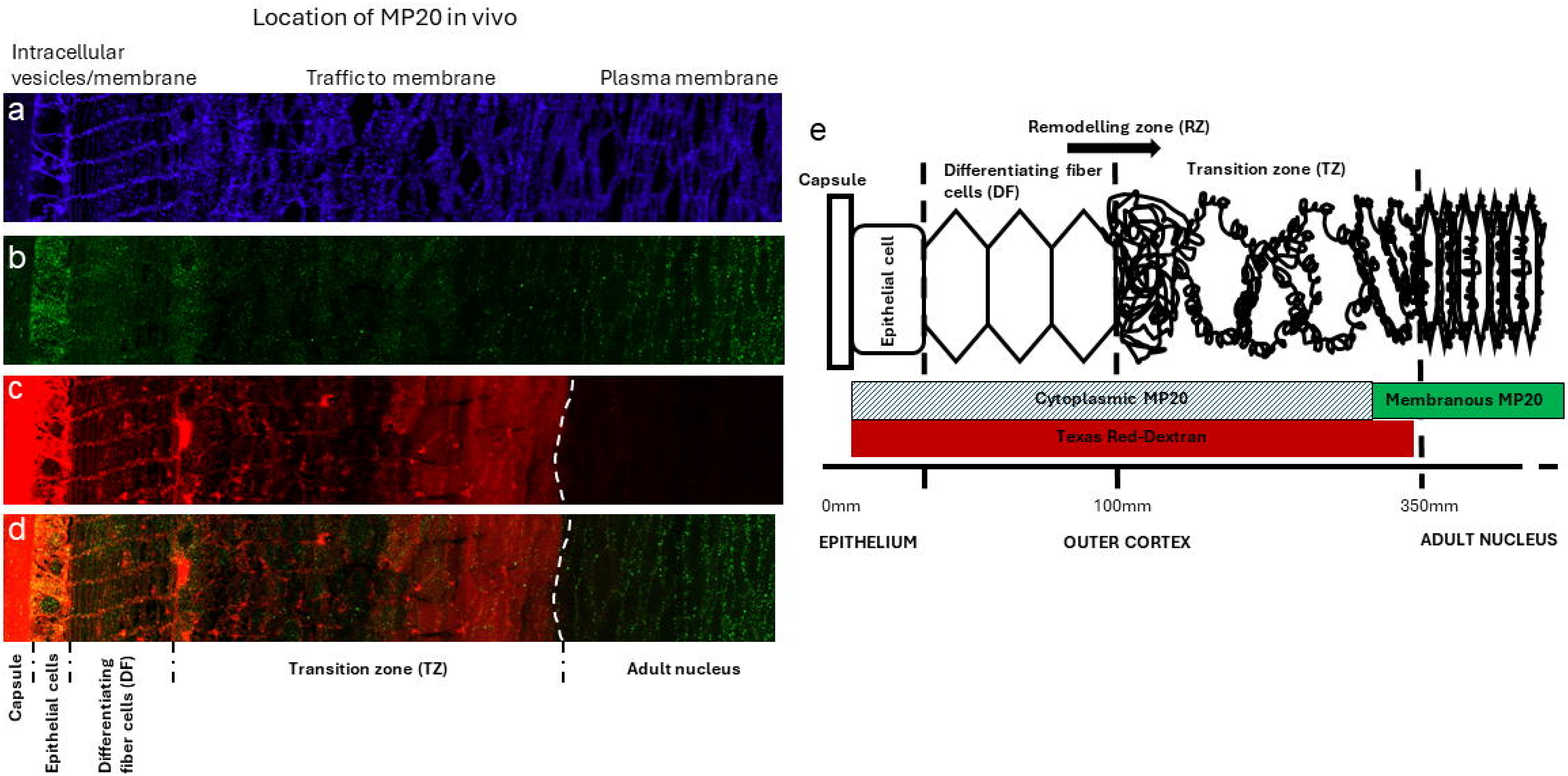


Vesicles

Vesicles + MP20

Figure 4

Insertion of MP20 into fiber cell membranes correlates with constriction of extracellular space



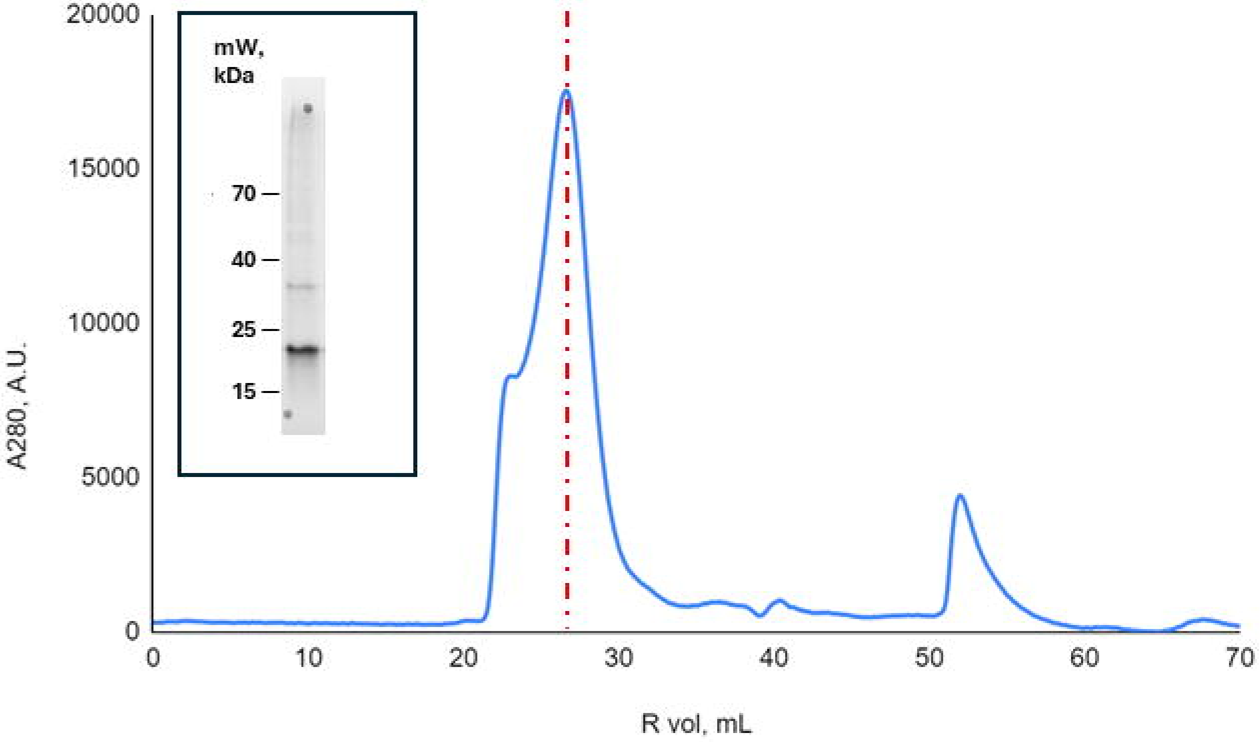
Supplementary Figure 1
Purification of MP20

a

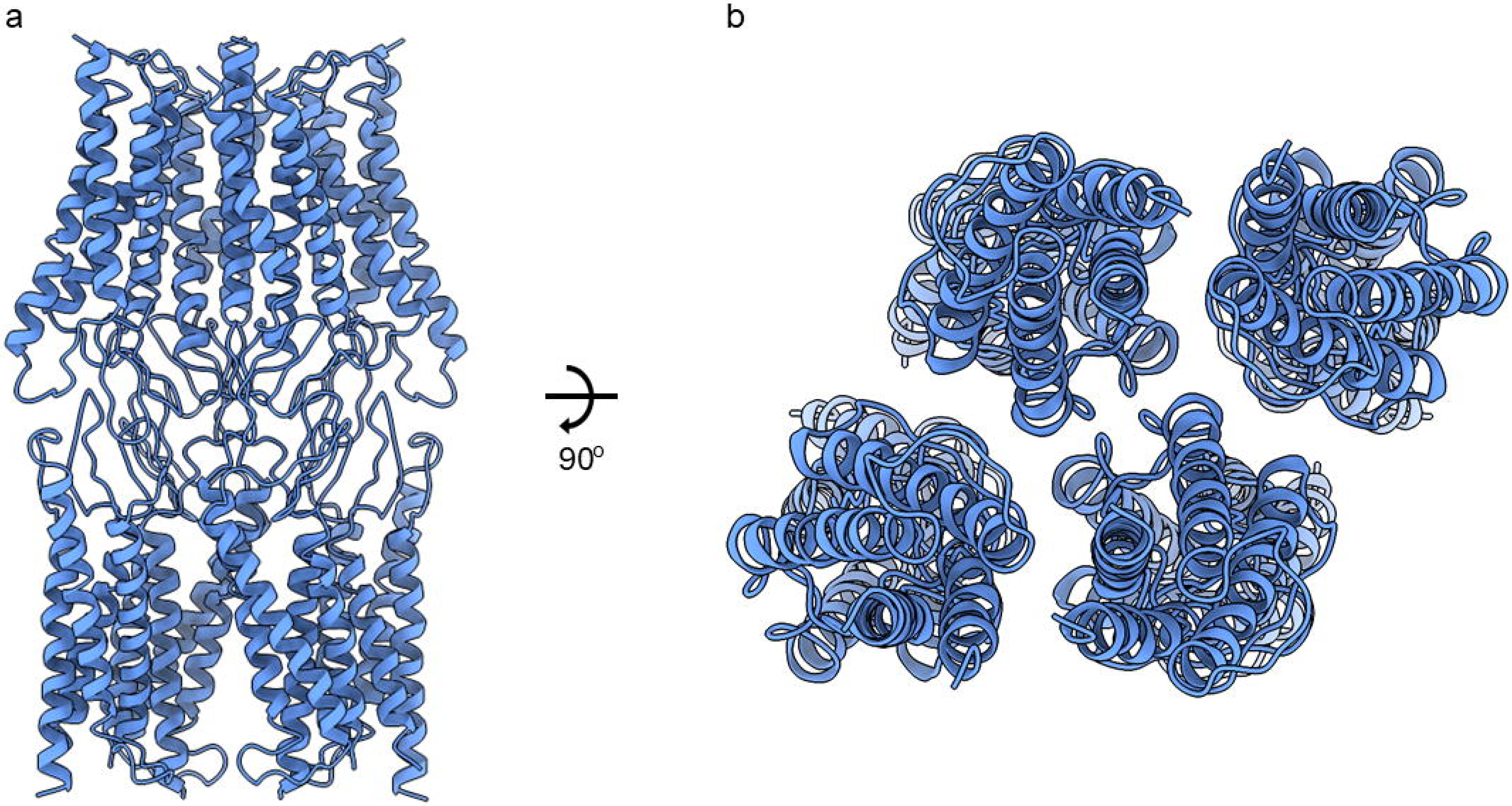
MP20 a.a. sequence used

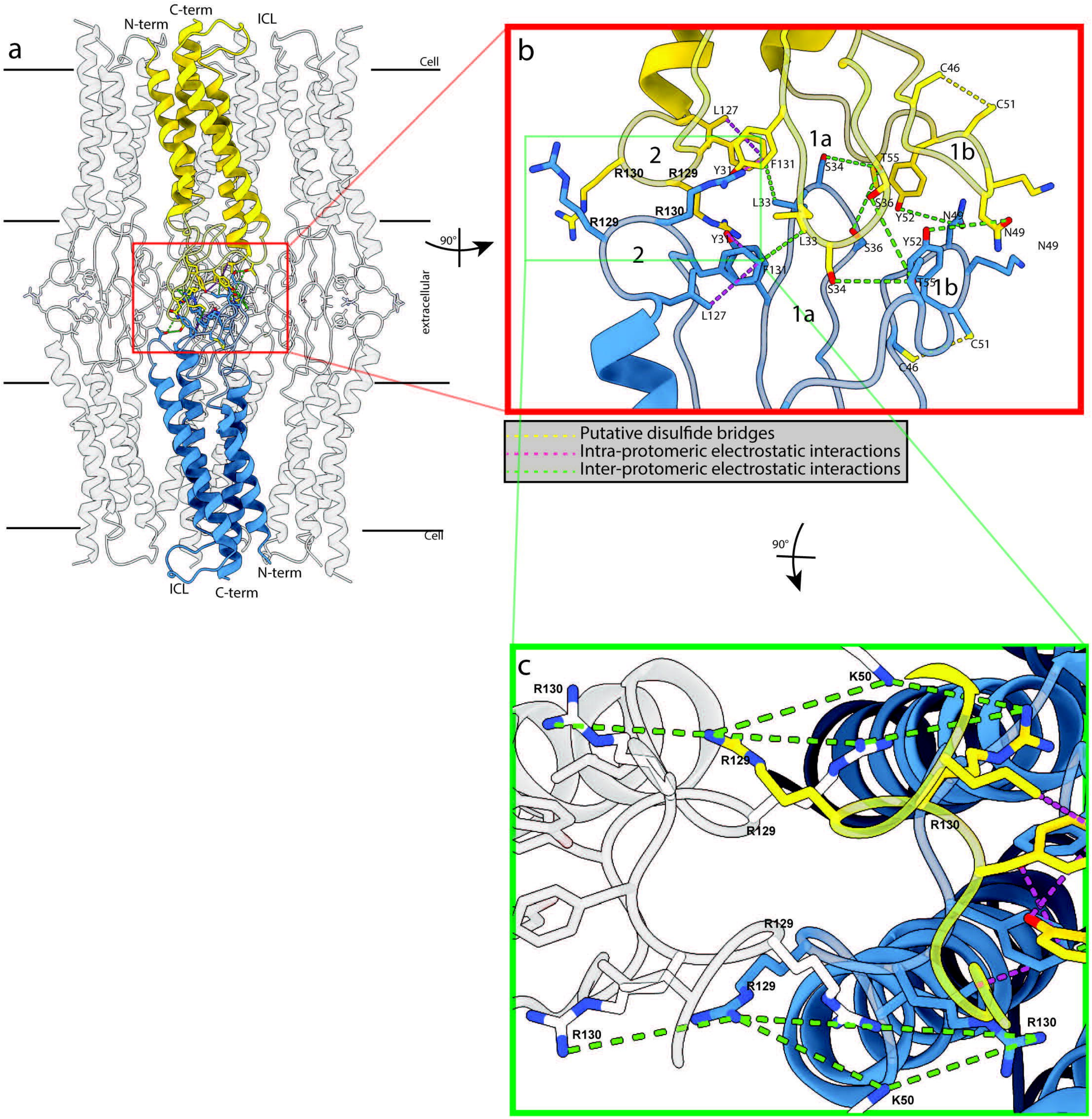
MKTIIALSYIFCLVFADYKDDDDAKLQTMHHHHHHHHHHLEVLFQMYSFMGGGLFCAWVGTILLVVAMATDHWMQYRLSGSFAHQGLWRYCLGNKCYLQTDSIAYWNATR
AFMILSALCAISGIIMGIMAFAHQPTFSRISRPFSAGIMFFSSTLFVVLALAIYTGVTVSFLGRRFGDWRFSWSYILGWVAVLMTFFAGIFYMCAYRVHECRRLSTPR*

b

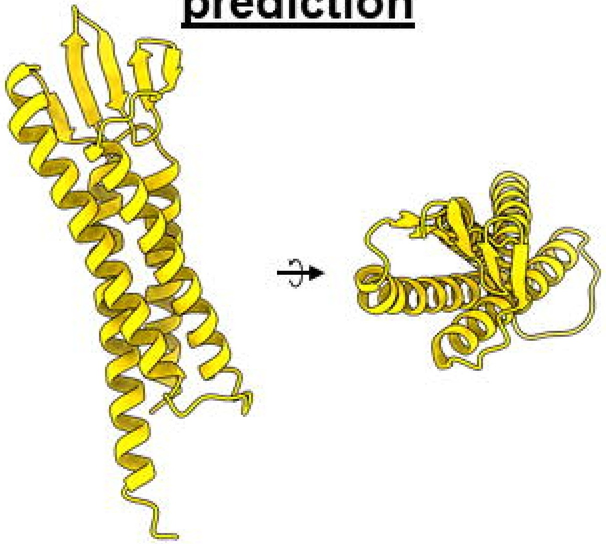


Supplementary Figure 2
Packing of MP20 crystals

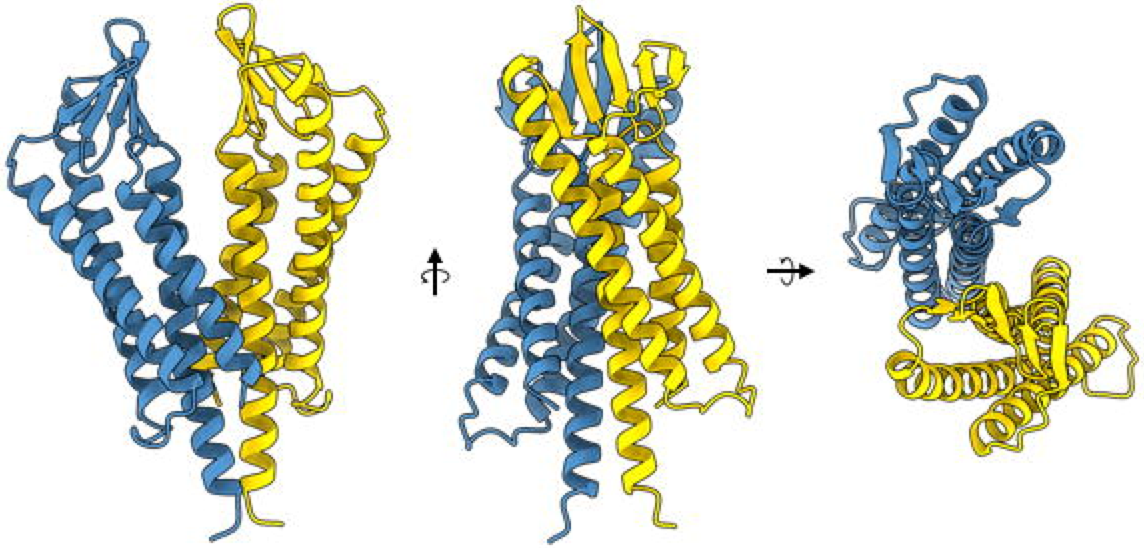



a
N-term
C-term
ICL
Cell
extracellular
Cell
ICL
C-term
N-term
90°
b
L127
2
R130
R129
Y31
F131
1a
S34
T55
1b
C46
C51
L33
S36
Y52
N49
R129
R130
Y31
L33
S36
Y52
N49
N49
2
F131
S34
T55
1b
L127
1a
C51
C46
Putative disulfide bridges
Intra-protomeric electrostatic interactions
Inter-protomeric electrostatic interactions
90°
c
R130
K50
R129
R129
R130
R129
R129
R130
R130
K50

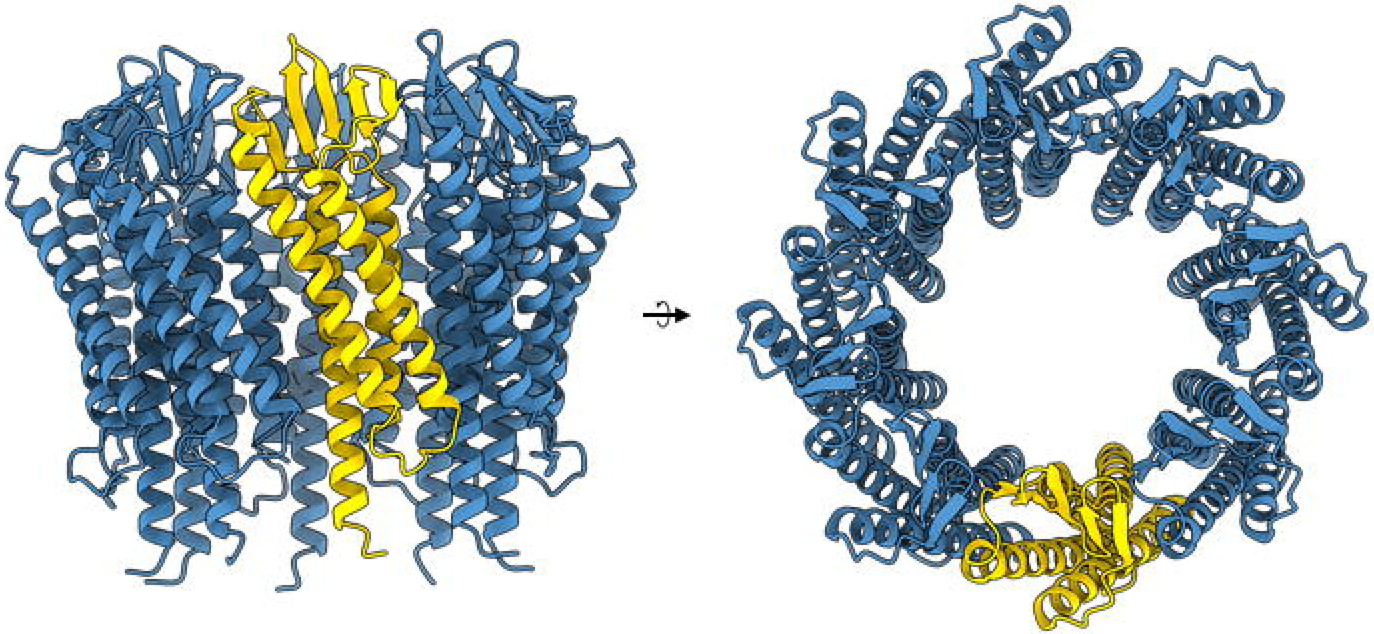
a. Alphafold3 – monomer prediction



b. Alphafold3 – dimer prediction



c. Alphafold3 – octamer prediction



MicroED structure of MP20	
Wavelength (Å)	0.0176
Resolution Range (Å)	52.26 - 3.5 (52.26 - 3.5)
Space Group	$P\ 4\ 2_1\ 2$
Unit Cell (a, b, c) (Å) (α, β, γ)(°)	56.18, 56.18, 142.5 90, 90, 90
Total Reflections (#)	58259 (1931)
Multiplicity	20.7 (11.3)
Completeness (%)	86.6 (78.4)
$< I / \sigma(I) >$	7.19 (2.32)
R-meas	0.39 (1.57)
R-pim	0.087 (0.370)
$CC_{1/2}$ (%)	98.9 (38.0)
R-work (%)	0.3316
R-free (%)	0.3510
Number of atoms (#)	1323
RMS bonds	0.002
RMS angles	0.34
Ramachandran (F/A/O) (%)	83.54 / 14.02 / 2.44
Rotamer Outliers (%)	0.00
Clashscore	5.34
Average B-factor (Å^2)	88.55