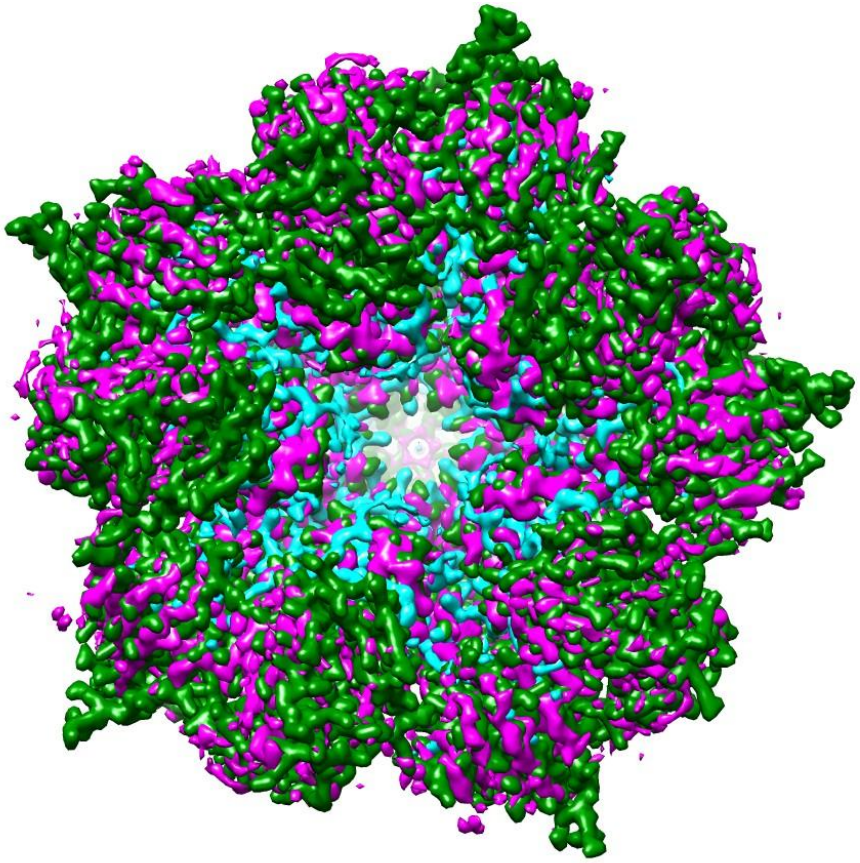
EXAMENSARBETE INOM BIOTEKNIK,
AVANCERAD NIVÅ, 30 HP
STOCKHOLM, *SVERIGE 2021*

Structural basis of modulation by pH and calcium in a ligand-gated ion channel

OLIVIA ANDÉN

KTH
SKOLAN FÖR KEMI, BIOTEKNOLOGI OCH HÄLSA

Structural basis of modulation by pH and calcium in a ligand-gated ion channel



Olivia Andén

Degree project in biotechnology, second cycle, 30 credits

KTH Royal Institute of Technology

Stockholm, Sweden 2021

Abstract

Pentameric ligand-gated ion channels (pLGICs) are crucial for the conversion of chemical to electrical signaling in the nervous system of mammals. Dysfunction in these channels has been found to be connected to several diseases including epilepsy, schizophrenia, Alzheimer's, and autism, making them the target of a wide variety of therapeutic agents. However, studying eukaryotic channels is challenging so the discovery of prokaryotic homologs that are much easier to study has thus greatly helped in the understanding of the structure and function in this family of proteins. In this project, a prokaryotic pLGIC called DeCLIC was produced and purified from *Escherichia coli.* Structural determination of the channel was pursued using cryo-electron microscopy at a low pH and in the presence of calcium. An electron density at 3.4 Å resolution was achieved and compared to previously determined structures at different conditions in an attempt to determine the structural modulation of calcium and pH. Results show multiple differences in channel conformation in the presence and absence of calcium as well as in different pH conditions. Furthermore, analysis of the determined electron density suggests a possible intermediate state at low pH in the presence of calcium.

Sammanfattning

Pentameriska ligandstyrda jonkanaler (pLGICs) är avgörande för omvandlingen av kemisk till elektrisk signalöverföring i djurs nervsystem. Dysfunktion i dessa kanaler har visat sig vara kopplad till flera sjukdomar inklusive epilepsi, schizofreni, Alzheimers och autism, vilket gör dem till en måltavla för en mängd olika läkemedel. Att studera eukaryota kanaler är dock mycket utmanande, så upptäckten av prokaryota homologer, som är mycket lättare att studera, har därmed bidragit mycket till förståelsen för struktur och funktion hos proteiner i denna familj. I detta projekt producerades och renades en prokaryotisk pLGIC kallad DeCLIC från *Escherichia coli*. Strukturell bestämning av kanalen genomfördes med användning av kryo-elektronmikroskopi vid lågt pH och i närvaro av kalcium. En elektrontäthet med 3.4 Å upplösning uppnåddes och jämfördes med tidigare bestämda strukturer vid olika förhållanden i ett försök att bestämma hur proteinets struktur moduleras av kalcium och pH. Resultaten visar flera skillnader i kanalens konformation i närvaro och frånvaro av kalcium såväl som vid olika pH-värden. Dessutom antyder analys av den bestämda elektrontätheten ett möjligt intermediärt tillstånd vid lågt pH i närvaro av kalcium.

Keywords

Pentameric ligand-gated ion channels, bacterial homolog, protein expression, protein purification, cryo-electron microscopy, data processing, protein structure.

Table of contents

Abstract **1**

Sammanfattning **1**

Keywords **1**

Table of contents **2**

1 Introduction **3**

1.1 Pentameric ligand-gated ion channels 3

1.1.1 Prokaryotic pLGICs *4*

1.1.2 Eukaryotic pLGICs *5*

1.1.3 Pharmacology *5*

1.2 DeCLIC 6

1.3 Cryo-electron microscopy 7

1.4 Aim of the project 9

2 Methods and materials **9**

2.1 Protein production 9

2.2 Purification 9

2.3 Thrombin activity assay 10

2.4 Cryo-electron microscopy 10

2.4.1 Sample preparation and data collection *10*

2.4.2 Data processing and data analysis *11*

3 Results **11**

3.1 Purification 11

3.2 Thrombin activity assay 13

3.3 Cryo-electron microscopy 14

4 Discussion **18**

5 Future perspectives **23**

6 Acknowledgments **24**

7 References **24**

Appendix **27**

1 Introduction

1.1 Pentameric ligand-gated ion channels

Pentameric ligand-gated ion channels (pLGICs) are membrane proteins found in a wide variety of both eukaryotic and prokaryotic organisms [1-3]. These channels facilitate selective transduction of ions across the membrane [1]. In mammals, they play an important role in the conversion of chemical to electrical signaling in the synapses of the nervous system [2,3]. Changes in the function of pLGICs have been found to be connected to several diseases including epilepsy, schizophrenia, Alzheimer's, and autism. Thus, channels in this family have been extensively studied in the pursuit of treatments of such conditions and are the target for a wide variety of therapeutic agents [3,4]. The basic structure of pLGICs (figure 1) is five homologous subunits consisting of three domains: an extracellular ligand-binding domain (ECD), a transmembrane domain (TMD), and, in eukaryotes, a cytoplasmic domain [2]. The ECD typically consists of an N-terminal α-helix followed by ten β-strands in a β-sandwich fold. The agonist binds to the loops of two adjacent ECD subunits causing a conformational change in the inner β-sheet which opens the channel [1,2]. The TMD consists of four α-helices typically denoted M1-M4, with M2 being positioned on the inside of the pore, creating a ring of α-helices together with M2 from the other subunits, and the other three being positioned around M2 and embedded in the lipid bilayer [1]. Residues in the M2-helices form rings of charged/polar regions in the pore which facilitates ion selectivity, as well as rings of hydrophobic residues creating a "hydrophobic gate" preventing smaller ions from passing through the pore in the closed conformation [2].

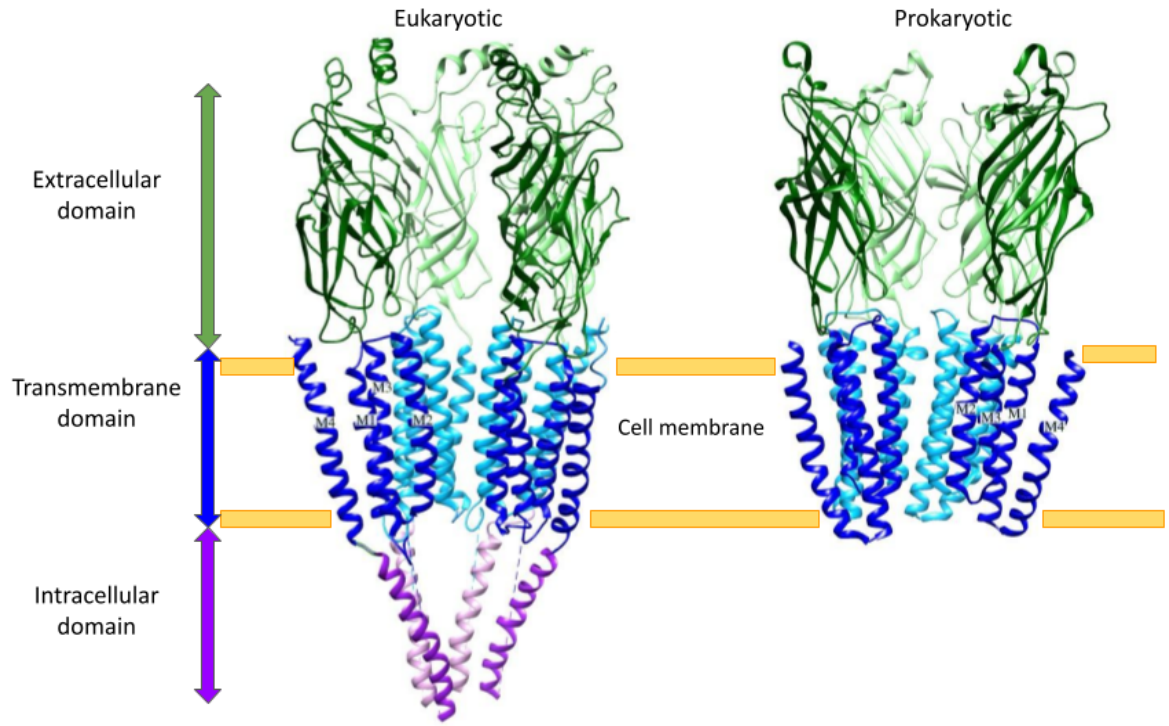


Figure 1: Basic structures of pLGICS. Eukaryotic nAChR (PDB ID: 2BG9) and prokaryotic GLIC (PDB ID: 6HZW) colored by domain as indicated by arrows. One of the five subunits has been hidden for clarity.

1.1.1 Prokaryotic pLGICs

The family of pLGICs was originally known as Cys-loop receptors, referring to the characteristic cysteines flanking the β6-β7 loop in all animal pLGICs. However, with the discovery of prokaryotic homologs without these cysteines, Cys-loop receptors became a subfamily under the superfamily pLGIC. [5] Prokaryotic pLGICs adopt a very similar structure as in eukaryotes, with the exception that they generally lack the intracellular domain and can contain additional extracellular domains at the N-terminus (NTDs) [3]. Although several structures of prokaryotic pLGICs have been solved, very little is known about the physiological role of these channels. It is believed that one possible function could be that the channels are involved in quorum sensing, which is a mechanism used for communication between bacterial cells through the release of signal molecules [1]. Relative to eukaryotic pLGIGs, the distribution of pLGICs in prokaryotes is sporadic and can differ greatly between closely related taxa while, simultaneously, similarities can exist in distantly related taxa. This could be an indication of a high degree of horizontal gene transfer and gene loss. This great diversity compared to eukaryotic channels also suggests that the prokaryotic pLGICs predate eukaryotic and the diversification largely happened before the emergence of eukaryotic pLGICs. Additionally, some prokaryotic pLGICs are much more similar to eukaryotic channels than other prokaryotes, suggesting these may have a common ancestor. [6]

1.1.2 Eukaryotic pLGICs

Unlike prokaryotic pLGICs are eukaryotic members of this family not always homo-oligomers with five identical subunits, but are rather pseudosymmetrical hetero-oligomers. These hetero-oligomers are assembled from several subtypes of pentameric subunits encoded by over 40 genes (in humans) in a very strict way to form distinct subfamilies. Assemblies within a subfamily, however, can be quite complex and can even differ in different tissue types or different regions of the brain. This heteromeric assembly of subunits also creates an asymmetry in ligand binding sites, where the orthosteric sites may exist in one subunit interface, while an allosteric modulatory site may exist in a different, homologous, interface of different subunit types in the same receptor. This is an important aspect to take into consideration in drug development since binding to these different sites can have very different effects depending on whether the drug acts as an agonist, antagonist, or allosteric modulator. Additionally, since most pharmacologically relevant binding sites exist in the subunit interface, this great divergence in subunit assembly presents a wide variety of possible drug targets but also makes it difficult to pinpoint the exact target for the desired effect [1].

1.1.3 Pharmacology

In humans, some of the most prominent proteins in the pLGIC family are the receptors for acetylcholine (nAChR), serotonin ($5HT_3R$), glycine (GlyR), and gamma-aminobutyric acid ($GABA_AR$) [4]. The nAChRs are responsible for postsynaptic activation of both nerve and muscle cells and are thus involved in physiological activities such as cognitive function, stress, anxiety, and pain processing [7,8]. The $5HT_3Rs$ are presynaptic receptors responsible for the release of several neurotransmitters involved in the reward system, anxiety control, vomiting reflex, autonomic and gastrointestinal function, drug addiction, schizophrenia, and regulation of immune response. Antagonists of the $5HT_3$ receptors are commonly used to relieve nausea and manage irritable bowel syndrome [8]. The GlyRs are both post and presynaptic with their most critical role being the postsynaptic receptors that mediate inhibitory neurotransmission in the spinal cord and brainstem. These receptors have been found to be connected to mainly chronic inflammatory pain, but also hyperekplexia, epilepsy, and autism [9]. The $GABA_AR$ exhibits both pre- and postsynaptic inhibition and has been found to be involved in immune cell and stem cell proliferation as well as diseases such as diabetes, inflammatory disease, and drug addiction. Common drugs that target these receptors are used for muscle relaxation, anxiety, epilepsy, sleep, and general anesthesia [10,11].

The discovery of prokaryotic orthologs to these mammalian channels has presented an opportunity to more easily study the structure and function of this family of proteins [3], as

well as providing possible model systems for drug development. The homology of these model systems is not perfect. In fact, a comparison of amino acid sequences shows only 18-20% homology between eukaryotic and prokaryotic pLGICs. However, multiple sequence alignment shows several conserved residues and motifs that are important for the function of the protein. Additionally, secondary structures and topology are very similar, especially in the transmembrane domain [1]. Prokaryotic channels are also a lot easier to produce in large quantities for structural and functional studies and the lack of heterologous subunits makes them much easier to study due to lower complexity.

1.2 DeCLIC

The ion channel that was analyzed in this project is the prokaryotic channel DeCLIC which is derived from a *Desulfofustis* deltaproteobacterium. DeCLIC, like all pLGICs, has a pentameric structure with subunits consisting of a membrane-embedded TMD and a periplasmic ECD, but with the addition of an NTD that is a large extracellular domain, accounting for almost 50% of the total mass of the protein, fused to the ECD (figure 2). The NTD consists of two lobes, NTD1 and NTD2, each mainly containing β-strands, with the exception that NTD1 contains two α-helices between β4 and β5. The β-strands in both lobes are arranged in a jelly-roll fold. Very little is known about the physiological role of this NTD but it provides additional interactions in the ligand-binding site, suggesting that it may have a regulatory function.

Calcium has been found to be an effector binding to the TMD-ECD interface and causing the channel to adopt a closed state. Ca^{2+} binds near the ECD-TMD interface where it is coordinated by two Glutamates in the β1-β2 loop and the loop-F as well as part of the backbone in the Pro-loop. Upon exclusion of Ca^{2+}, the NTDs and EDC both contract while the helices in the TMD are twisted and translocated outward, opening up all constricting gates in the pore. Electrophysiological studies have also shown the transition from a non-conducting to a conducting state upon the depletion of Ca^{2+} [3].

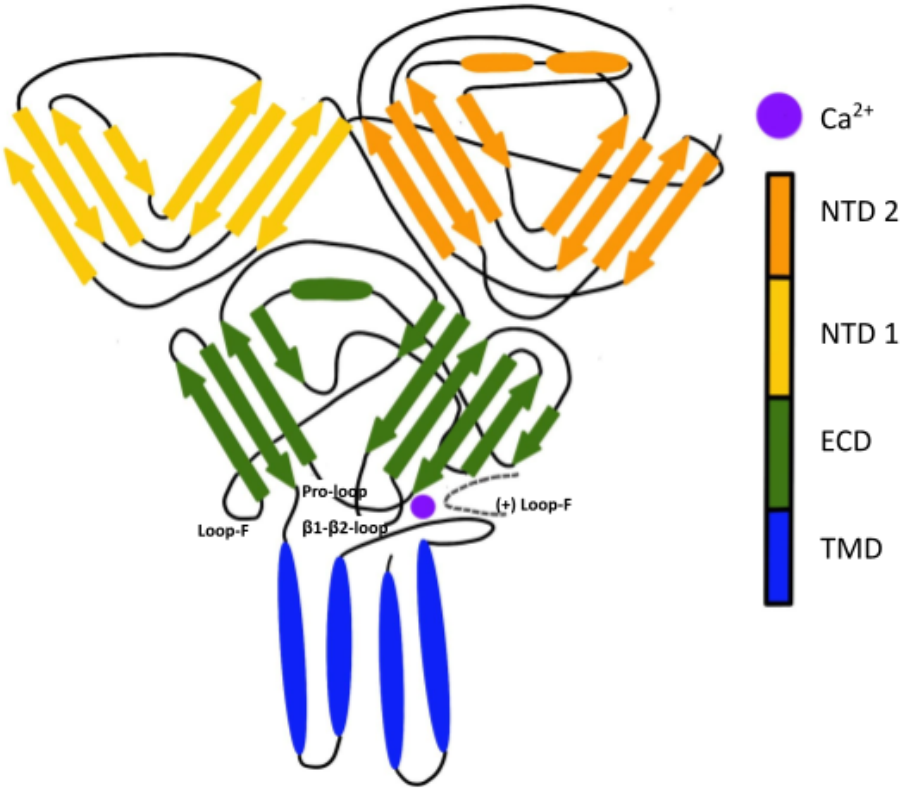


Figure 2: DeCLIC subunit topology adapted from Hu et. al. [3]. Domains have been colored as indicated on the right and loops involved in the calcium-binding site are labeled. Dashed line indicates loop-F of the adjacent subunit.

1.3 Cryo-electron microscopy

Cryo-electron microscopy (cryo-EM) is a microscopic imaging technique using the refraction of an electron beam to produce a 2D image of a specimen (micrographs) that can be reconstructed into a high-resolution 3D structure. [12-16] Development of this technique began in the 1970s [12] and in 2017, the pioneers were awarded the Nobel prize in chemistry [13]. The electron microscope functions similarly to the regular light microscopy where the source of radiation is visible photons passing through a specimen mounted on a glass slide and then refracted through an optical lens to produce an image. In an electron microscope, the radiation source is instead electrons being accelerated through a column in a vacuum with an accelerating voltage of 80-300 kV and refracted using micromagnetic lenses. Since the achievable resolution of a microscope is directly related to the wavelength of the radiation source, and visible light has a wavelength of 4000-7000 Å while a typical electron transmission source used in electron microscopy can produce a wavelength of 0.02 Å, electron microscopy can thus produce images of about 10000x the magnification compared to a light microscope. However, the use of an electron microscope does introduce some

problems, the main ones being beam-induced drift, causing blurring of the particles, and radiation damage [14].

Radiation damage is caused by the high-energy electrons breaking the bonds in the molecules, denaturing the sample and creating free radicals that, in turn, further the damage. One approach to protect the sample from radiation damage is to reduce the electron dose ($e/Å^2$) applied to the sample, however, this will provide very noisy images and thus decrease the signal-to-noise ratio (SNR), greatly decreasing the obtainable resolution. Another, much more effective, approach is to capture the sample in vitrified ice maintained in cryogenic temperatures, giving the method its name (cryo-EM) [14]. The cryogenic temperature protects the sample from damage while the vitrified ice keeps the sample in its biological state. Still, limitation in the possible electron dose limited the application of cryo-EM to highly symmetric and large protein complexes with low resolution. This stayed true until 2013 when a new type of detector camera was introduced, the direct electron detector (DED), that drastically improved SNR and image quality. These detectors are able to record several frames per micrograph (movies) that can be averaged into a single image [15]. Being able to record movies, rather than a single exposure image, means a higher total electron dose can be applied, since early frames correspond to a low dose with minimal damage but a high amount of specimen movement and low SNR, while later frames correspond to a high dose with a higher SNR and less blurring due to movement but with more damage. Averaging these frames will thus provide a high-resolution image with a good SNR and minimal motion [16]. In the years since 2013, the field of cryo-EM has seen an incredible speed in development and today resolutions higher than 4 Å are routinely achieved, with a resolution of as high as 1.2 Å having been reported [17]. This surge in the technology has been denoted the “Resolution revolution” and the introduction of the DED is a major contributor to this development, along with the development of faster and better computing power with the ability to store and process big amounts of data [15].

Sample to be used for cryo-EM is typically prepared by applying the sample to a gold or copper grid covered in a film of holey carbon that has been rendered hydrophilic using, for example, a glow discharger. The specimen is then captured in vitrified ice, usually by plunge freezing it in liquid ethane or helium. An optimally vitrified specimen is embedded in non-crystalline ice of sufficient thickness to encompass the particles, but thin enough to have all particles clearly visible within the holes of the carbon film [16]. After data acquisition on the microscope, the data can be processed using one of the several processing software packages available such as RELION [18], cryo-SPARC [19], or SPHIRE [20], among others.

1.4 Aim of the project

The aim of this project was to produce and purify DeCLIC, and use it to collect and process a dataset at pH 5 with calcium present. Preliminary results from the group have shown low pH in combination with depletion of calcium to be promoting an apparent open state of the channel, the only condition where this has been captured in a cryo-EM experiment. Therefore, the hypothesis was that data collected at low pH in the presence of calcium may reveal an intermediate state, which could provide an insight into the functional transition of the protein structure. To explore this possibility, the determined calcium-present cryo-EM density was compared to a previously published X-ray crystal structure determined at neutral pH in the presence of calcium [3], and a preliminary unpublished cryo-EM density processed by another member of the group Molecular Biophysics Stockholm, Urška Rovšnik, determined at pH 5 in the absence of calcium. Comparison to the X-ray crystal structure could reveal the effect of pH while comparison to the cryo-EM density could reveal the effects of calcium at a low pH.

2 Methods and materials

2.1 Protein production

All reagents and chemicals were obtained from either Sigma-Aldrich® or Fisher Scientific™. A starter culture was prepared by inoculating C43(DE3) *E. coli* transformed with DeCLIC-MBP in 2xYT media with 100 µg/ml ampicillin overnight at 37°C. The starter culture was then inoculated 1:100 in 2xYT-Amp expression media at 37°C and the cell growth was monitored with optical density at 600 nm. Once the optical density reached 0.6-0.8, protein expression was induced by adding Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, and the cultures were left at room temperature with shaking overnight. The cells were harvested by centrifugation at 5000 rpm for 12 minutes and then resuspended in buffer A (300 mM NaCl and 20 mM Tris•HCl, pH 7.4) supplemented with EDTA-free protease inhibitor cocktail, 1 mg/ml lysozyme, and 40 µg/ml DNase I and transferred to 50 ml conical tubes. The cells were lysed by sonication and then ultracentrifuged at 40 000 rpm for one hour. The pellets were solubilized in 1 ml per gram pellet of 4% n-Dodecyl β-D-maltoside (DDM) in buffer A overnight. Unsolubilized material was removed with ultracentrifugation at 28 000 rpm for one hour.

2.2 Purification

Amylose resin was pre-equilibrated with 0.1% DDM in buffer A and then mixed with solubilized supernatant and incubated at 4°C for one hour with gentle mixing. The resin was

then spun down in a centrifuge at 3000 rpm for 15 min and resuspended in an equal volume of 0.1% DDM after removal of the supernatant. The resin was washed with 2 column volumes of 0.1% DDM and 7 column volumes of 0.02% DDM. The protein was eluted in four rounds, first with 2mM Maltose in 0.02% DDM after 15 min incubation (Elution 1), second with 20mM Maltose incubated overnight (Elution 2), and finally two rounds of 20 mM Maltose with 15 min incubation (Elution 3 and 4). Each elution was then concentrated and purified using size exclusion chromatography. The purity of the peak fractions was confirmed using SDS-PAGE and protein concentration was determined with Bradford assay before peak fractions from all elutions were pooled together.

The maltose-binding protein (MBP) fusion tag was cleaved using thrombin reconstituted in a 10x storage buffer with 200 mM Tris, 1.5 M NaCl, and 25 mM $CaCl_2$. 100 u thrombin per mg of protein were added to the purified protein and left to cleave overnight with gentle mixing. The cleaved protein was then concentrated and the fusion partner was separated using size exclusion chromatography with a buffer at desired pH and with the desired detergent for further experiments. The purity of peak fractions was confirmed using SDS-PAGE and protein concentration was determined with Bradford assay.

2.3 Thrombin activity assay

To test the activity of thrombin at different conditions, a thrombin activity assay was designed. 20 µl digestion reactions were prepared using purified DeCLIC-MBP. The conditions that were tested were two different pHs (pH 7 and pH 5), and two different cofactors, 2.5 mM Ca^{2+} (as in original protocol) and 2 or 10 mM Mg^{2+}. Thus a total of six conditions were tested. To lower the sample to pH 5, citrate buffer at pH 5 was added to a concentration of 100 mM.

2.4 Cryo-electron microscopy

2.4.1 Sample preparation and data collection

For the sample used in Cryo-EM, the last purification step was done in a citrate buffer at pH 5 with 0.02% DDM and 10 mM $CaCl_2$. Peak fractions were pooled and concentrated to ~3.5 mg/ml. The sample was frozen onto Quantifoil 1.3/1.2 Cu 200 mesh grids that have been glow discharged for 2 min at 40 mM current using GloQube. 3 µl of sample was applied to the grids and then blotted for 3 s before being plunge-frozen in liquid ethane that was cooled by liquid nitrogen using Vitrobot Mark IV (FEI). Micrographs were collected on a Titan-Krios (FEI) 300 kV microscope with a K3 detector camera at 105000x magnification. A total of 26 260 movies were collected with 40 frames over 2 sec and a total dose of 43.695 e/A^2. The defocus range was -1.4-3.0 µm.

2.4.2 Data processing and data analysis

The collected data set was processed using RELION 3.1 [21]. Micrographs were motion-corrected using MotionCor2 [22]. After CTF estimation using CTFFIND4 [23], 600 particles were manually picked from micrographs with different defocus values. The manually picked particles were classified into 2D classes and 14 were further used for automatic picking. A total of 5.5 million particles were picked in auto-picking and the particles were extracted and 2D classified. Two classes, one top view and one side view of the protein, were chosen to generate an initial model. 3.3 million particles were chosen to be further processed. The particles were then aligned with 3D auto-refine using the initial model as a reference. After multiple rounds of 2D and 3D classification, the number of particles had been reduced to 150 000 and a final round of 3D auto-refinement followed by mask creation and post-processing to generate the final density as well as estimation of local resolution was performed. The density was assessed and analyzed in Chimera [24], where it was also compared to a previously published X-ray crystal structure determined at neutral pH in the presence of calcium [3], and a preliminary unpublished density processed by another member of the group Molecular Biophysics Stockholm, Urška Rovšnik, determined at pH 5 in the absence of calcium.

3 Results

3.1 Purification

Amylose affinity purification successfully extracted the DeCLIC-MBP fusion partner, as seen in the gel in figure 3A. A strong band near 115 kDa, which is the expected molecular weight of the DeCLIC-MBP monomer (113 kDa) is visible. Two additional bands could be observed in the elutions, one can be assigned as self-cleaved MBP and the other was identified as Maltoporin, a sugar-binding protein that is endogenously expressed in *E. Coli* [25]. The protein in these additional bands was separated out in the SEC, as evident by the gel in figure 3B and the chromatogram in figure 4A. The chromatogram of post thrombin cleavage (figure 4B), shows two clean peaks of similar height corresponding to successfully cleaved DeCLIC and MBP fusion partner, as confirmed by the gel in figure 3C.

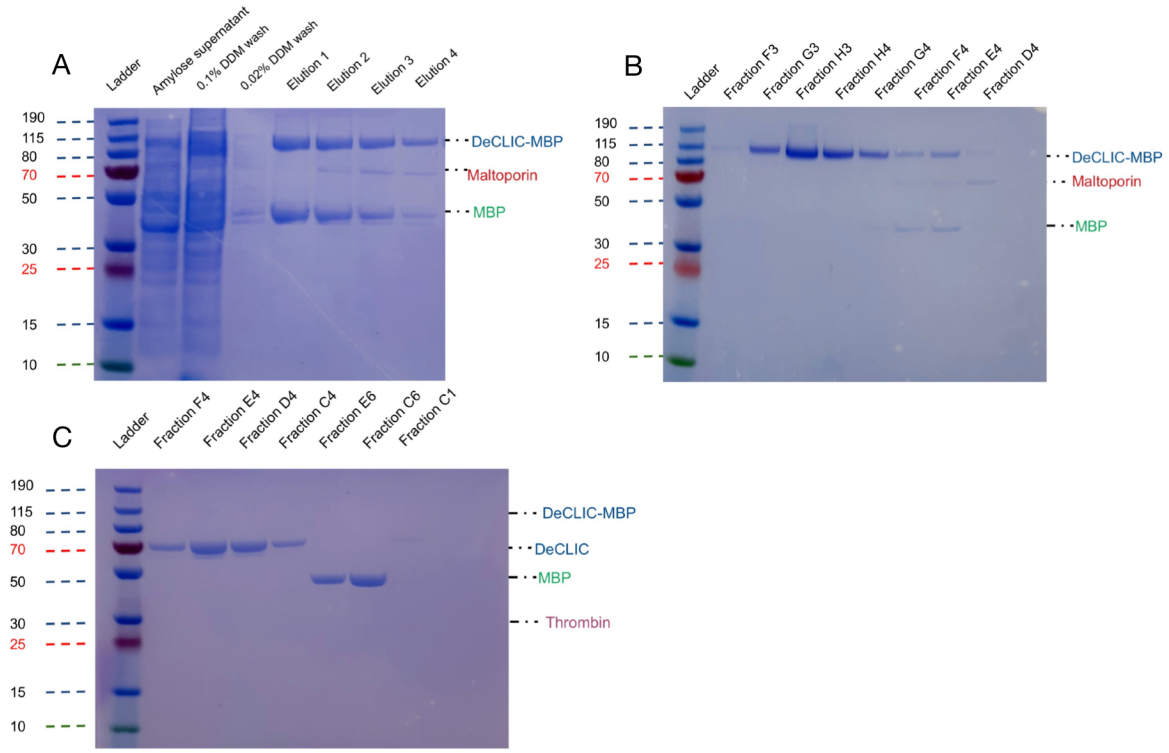


Figure 3: SDS-PAGE gels with samples from the purification process. A) Amylose affinity purification. B) Peak fractions from size exclusion chromatography of one elution. C) Peak fractions from size exclusion chromatography of cleaved product.

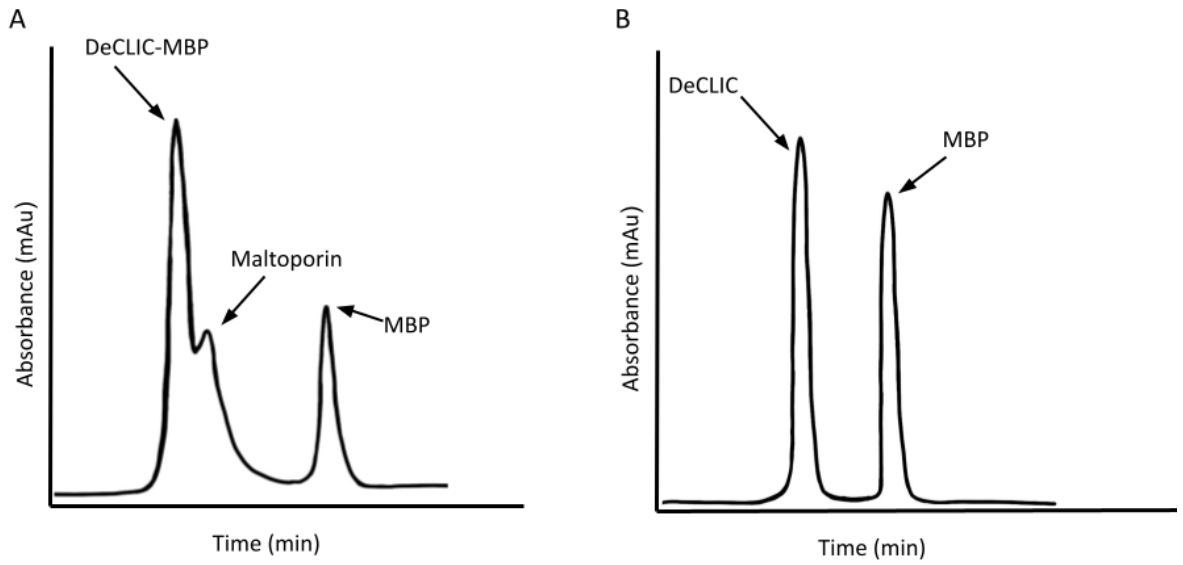


Figure 4: Trace-over of example chromatograms, for original chromatograms see appendix, figure A1. A) The same elution as in figure 3B. B) The same run as in figure 3C.

3.2 Thrombin activity assay

Thrombin cleavage is usually performed at pH 7 with calcium as a cofactor. However, since the channel is known to be regulated by calcium and could be regulated by pH, it would be beneficial to have the option to perform the cleavage in different conditions. Therefore, an activity assay was performed. The results from this assay can be seen in figure 5. This shows that full cleavage was only achieved when using calcium at pH 7. Doing the reaction at pH 5 does not work at all and using the alternative cofactor, magnesium, appears to have worked on about half of the protein with no apparent improvement when using a higher concentration (based on visual assessment of the gel bands).

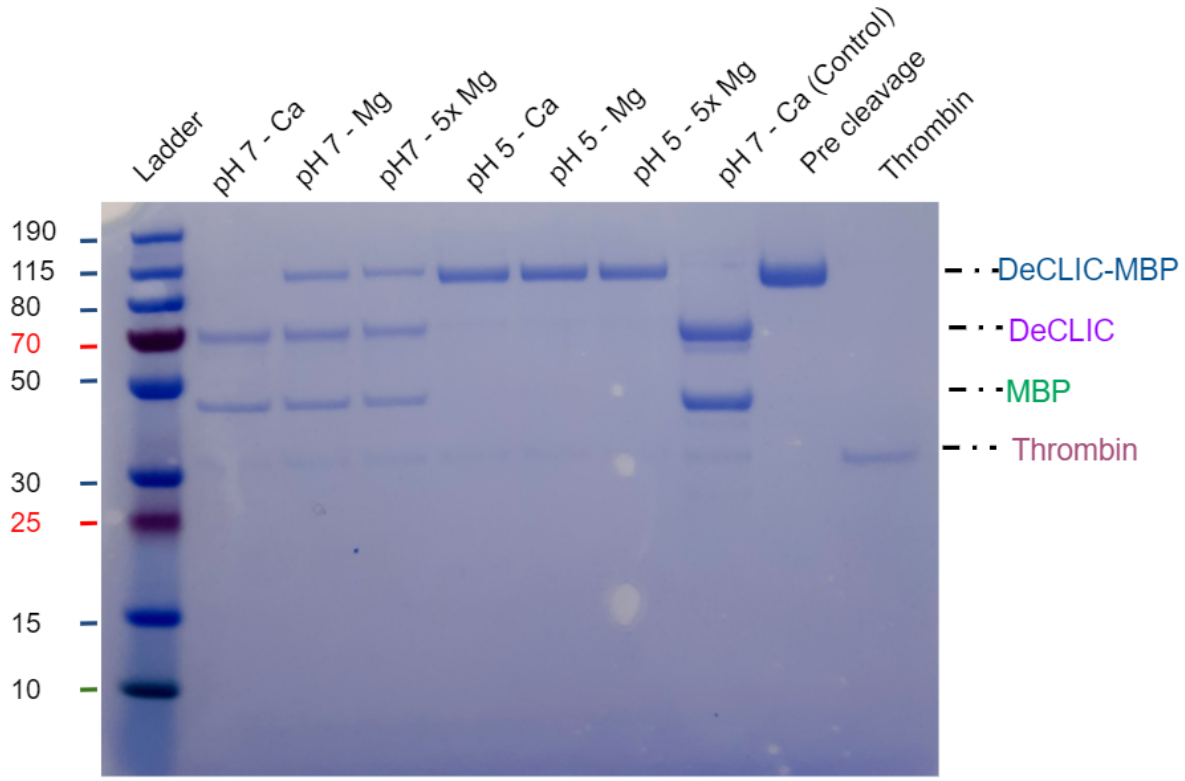


Figure 5: Thrombin activity assay. Lane labels indicate the pH of the digestion reaction and what cofactor was added.

3.3 Cryo-electron microscopy

The data processing pipeline is summarised in figure 6 . Initial 2D classes in panel A confirm that the protein was captured in several different orientations and each domain was clearly visible in the class averages. The distribution of orientations was sufficient to produce an initial model with the expected structure (panel B) that could be used for 3D alignment of the particles and 3D classifications, one of which is shown in panel C. The final density in panel D had a global resolution of 3.4 Å. Important to note, despite the channel structure being homo-pentameric, the data could initially not be processed with C5 symmetry imposed but rather required no symmetry applied (C1) to produce a 3-dimensional density that corresponds to the structure of the protein (see appendix, figure A2). Only when the number of particles had been reduced to 150 000 could C5 symmetry be imposed to produce the final density.

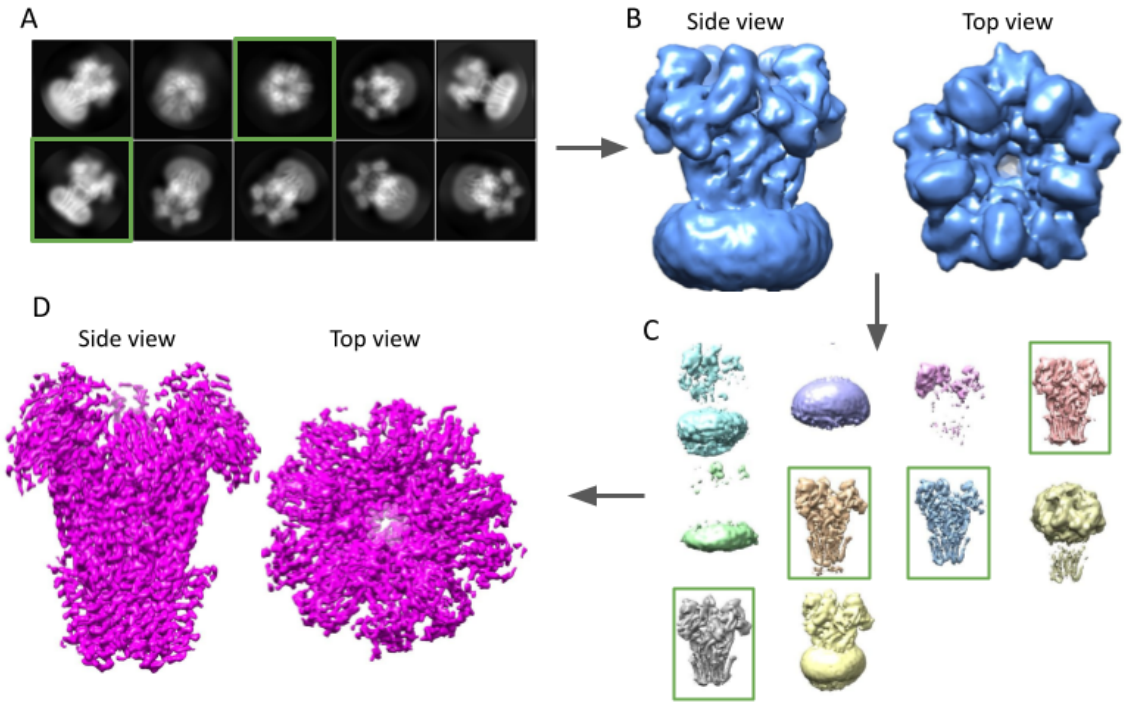


Figure 6: (A) Initial 2D classes selected after classification of auto-picked particles. Classes marked in green were used to generate the initial model. (B) Initial model at 10 Å resolution generated from 2D classes. C) 3D classes from the first round of 3D classification of 3.3 million particles. Classes marked in green were used in subsequent rounds of 2D and 3D classifications. D) Final 3.4 Å density generated from 150 000 particles.

Although the global resolution of the determined electron density was 3.4 Å, the estimation of local resolution reveals the resolution in reality spans from 3.2 to 4.5 Å in different parts of the protein, as seen in figure 7. In the TMD, the resolution is the highest inside the pore, with the M2 helices being resolved to 3.2 Å, while it is decreasing, as we move out towards the M4 where it is about 3.5 Å. The resolution in the ECD stays consistently above 3.5 Å, with the highest resolution being on the inside of the pore and towards the bottom of the domain. The resolution of the NTDs, in contrast, stays consistently below 3.5 Å but follows the same pattern of being lower towards the outermost parts of the protein and higher towards the center.

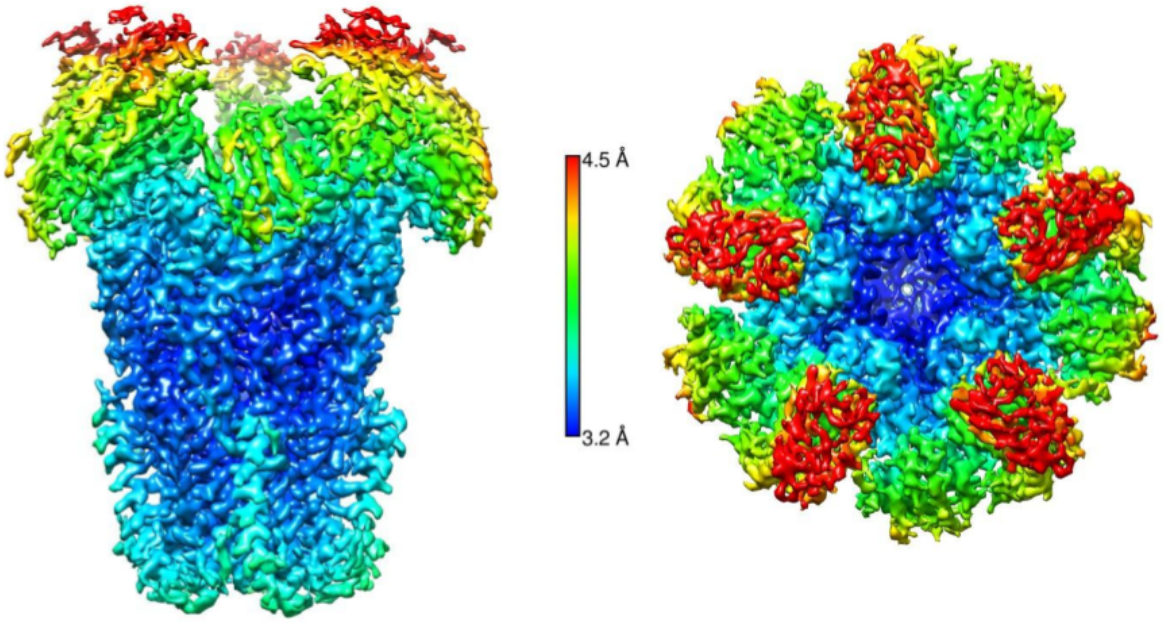


Figure 7: Side (left) and top (right) view of final density colored by local resolution as indicated by the color bar.

Comparing the density to the previously published x-ray crystal structure (PDB ID: 6V4S) of a closed channel determined in the presence of calcium at neutral pH (figure 8A) shows that the model fits well within the density in both the TMD and ECD. In the calcium-binding site (figure 8B) both the Pro-loop and loop-F of the model are completely embedded by the density except for the residues Asp475-Glu480, for which an alternative positioning of the loop can be observed in the density.

The determined density was also compared to a cryo-EM density determined in the absence of calcium at pH 5, resolved to a resolution of 3.7 Å. The protein and data used to create this calcium-free density were produced, collected, and processed by a member of the Molecular Biophysics Stockholm group at SciLifeLab. To align the two densities for comparison they were

both individually fitted to the same 3.5 Å surface representation of the x-ray crystal structure (PDB ID: 6V4S) created using the function molmap in Chimera [24]. Comparison of the superimposition of the two densities (figure 8C) shows that the ECD in the density with calcium-present is expanded about 2.5 Å (measured at the center of the density of corresponding β-strands) relative to the calcium-free density. This difference in the ECD appears to be consistent in the whole domain except for in the TMD-ECD interface, where the densities are more closely aligned, and in the calcium-binding site (figure 8D). In the calcium-binding site, the Pro-loops of the two densities align similarly to other regions in that area, while residues Asp475-Glu480 in Loop-F deviates from the pattern of difference right before and after these residues by as much as 6 Å. Comparison of the NTDs show a similar effect as in most of the ECD where the NTDs in the calcium-present density are extended out and also slightly down.

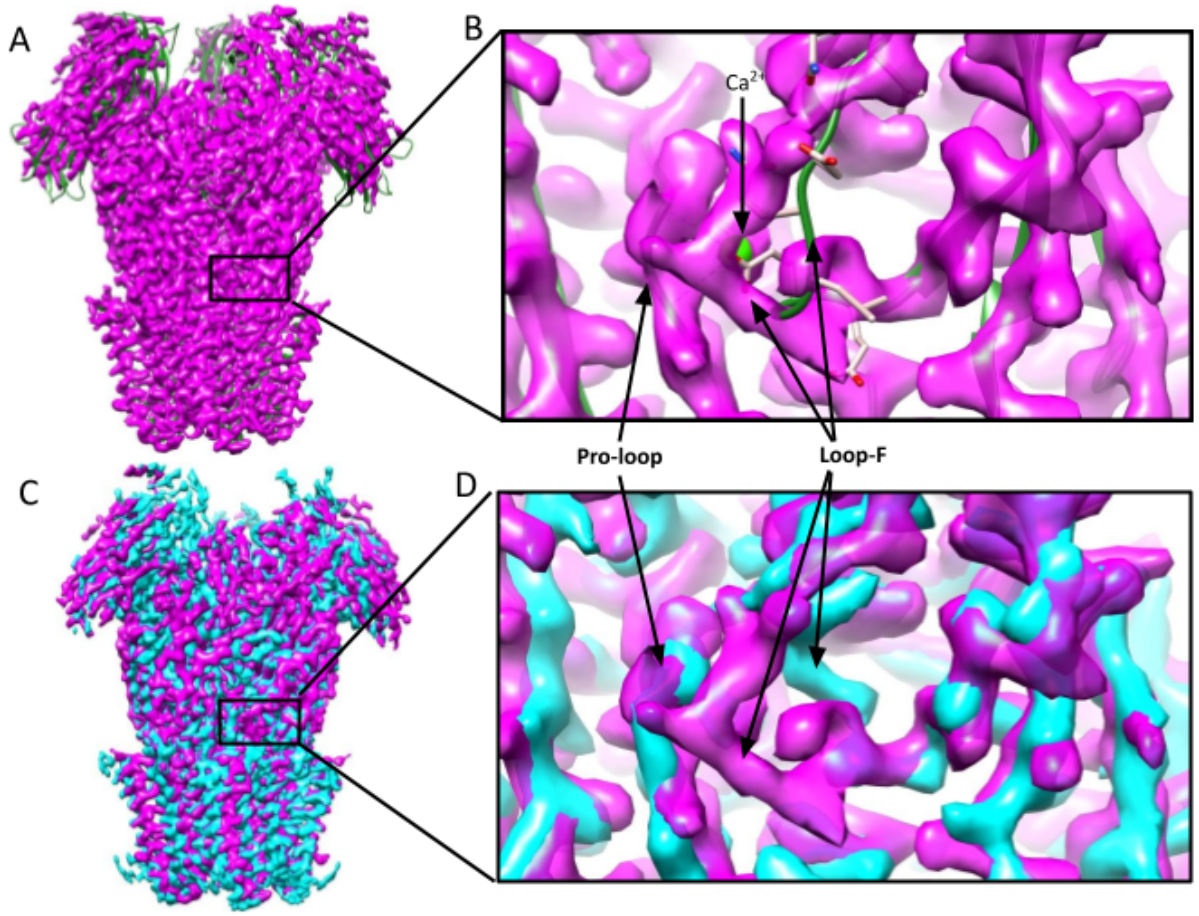


Figure 8: Comparison to different conditions. A-B) X-ray crystal structure (PDB ID: 6V4S) fitted to the calcium-present cryo-EM density (magenta), full protein (A) and zoom in to the calcium-binding site (B). C-D) Superimposition of calcium-present cryo-EM density (magenta) and calcium-free cryo-EM density (cyan). Full protein (C) and zoom in to the calcium-binding site (D). Credit information: the

calcium-free cryo-EM data (cyan) was produced and processed by a member of the Molecular Biophysics Stockholm group at SciLifeLab.

Two main constriction points were identified in the TMD pore. The first (figure 9A) correlates to Phe561 at the 16´ position and has a diameter of 8.6 Å. The second (figure 9B) correlates to Leu554 at the 9´ position and has a diameter of 7.6 Å. In both points, the constriction diameter in the calcium-free cryo-EM density is significantly larger, as is true for the whole pore. In the surface representation of the x-ray crystal structure (PDB ID: 6V4S), however, the constrictions at 16' and 9' are both about 4 Å smaller in the x-ray crystal structure than in the calcium-present cryo-EM density.

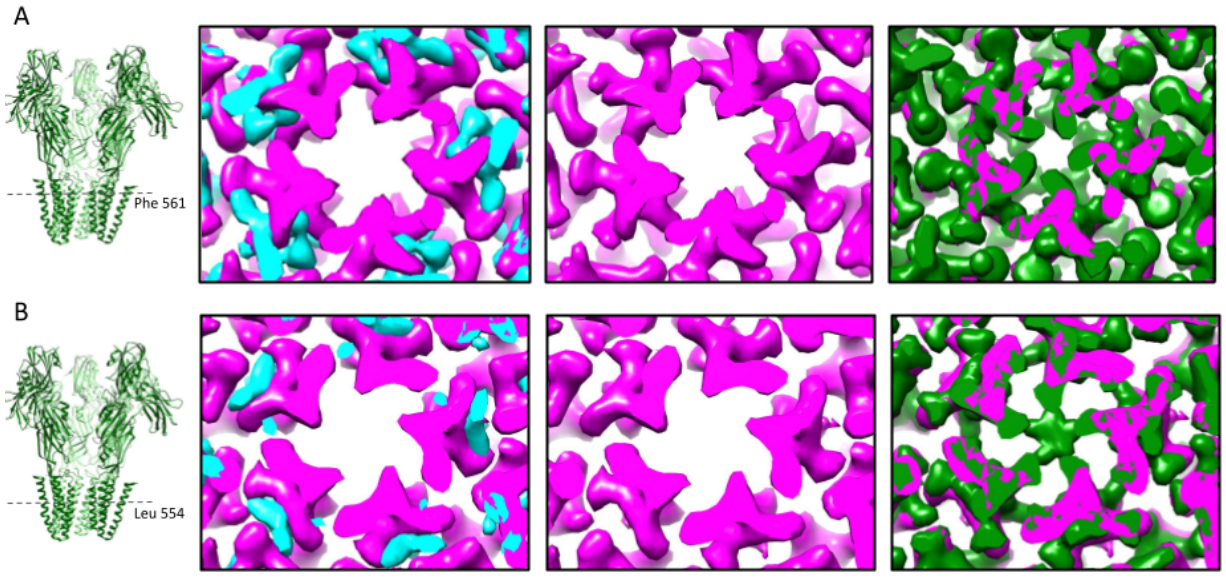


Figure 9: Calcium-present cryo-EM density (magenta) superimposed with calcium-free cryo-EM density (cyan) or surface representation (green) of the x-ray crystal structure (PDB ID: 6V4S) showing the most constricting points in the TMD. Sliced position is indicated on the far right by a dashed line on a ribbon representation of the closed channel (PDB ID: 6V4S) viewed from the membrane. A) Sliced through Phe561. B) Sliced through Leu554. Credit information: The calcium-free cryo-EM data (cyan) was produced and processed by a member of the Molecular Biophysics Stockholm group at SciLifeLab.

In the ECD pore, the situation is reversed (figure 10). Here, the largest constriction can be observed in the calcium-free cryo-EM density (8.5 Å) followed by the surface representation of the x-ray crystal structure (12.5 Å) with the calcium-present cryo-EM density being the least constrictive (15 Å). In all cases the most constricting point corresponds to the residue Trp407.

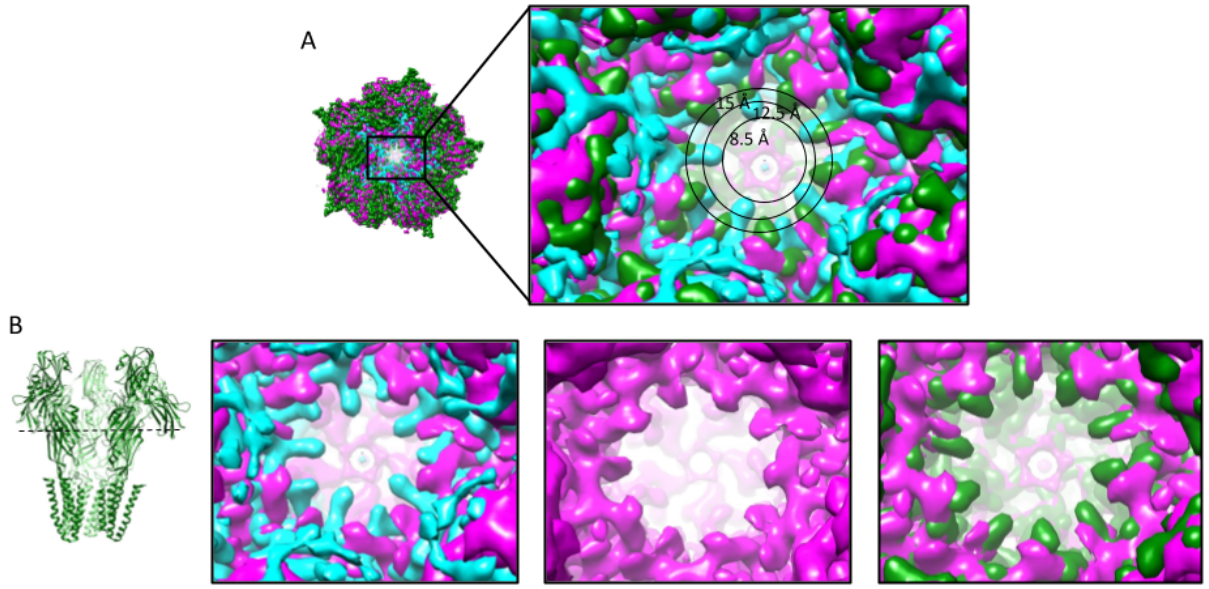


Figure 10: ECD pore viewed from the top down. A) Calcium-present cryo-EM density (magenta) superimposed with calcium-free cryo-EM density (cyan) and surface representation (green) of the x-ray crystal structure (PDB ID: 6V4S) showing the most constricting points in the ECD with pore diameter indicated with circles. B) Same position as in A with one or both comparison densities hidden. Position is indicated on the far right by the dashed line on a ribbon representation of the closed channel (PDB ID: 6V4S) viewed from the membrane. Credit information: The calcium-free cryo-EM data (cyan) was produced and processed by a member of the Molecular Biophysics Stockholm group at SciLifeLab.

4 Discussion

Most methods used to analyze the structure and function of proteins require a relatively high amount of protein, especially if one wants to do multiple experiments using the same sample. Therefore, it is important to achieve as high a yield as possible during purification. In the purification method used for this project, the obtained protein concentration was typically 0.8-1 mg/ml, which meant that the sample had to be concentrated to the 3-4 mg/ml needed for cryo-EM, simultaneously concentrating the detergent which may not be desirable. One factor that could be limiting the yield is the contaminant protein maltoporin. Firstly, it may be limiting the amount of DeCLIC-MBP able to bind to the amylose resin by filling up the available binding spots. This could explain why some DeCLIC-MBP appears to still be in the amylose supernatant, as evident by the thick band around 115 kDa in the first two lanes in figure 3A. Although it is not certain that these bands represent DeCLIC-MBP and the amount of amylose resin may be more than sufficient to allow the binding of both DeCLIC-MBP and Maltoporin, it would still be interesting to see whether an increase in amylose resin used would eliminate these bands. Secondly, maltoporin has a similar runtime on the SEC as

DeCLIC-MBP, causing the peaks to partially overlap and thus making it necessary to discard several fractions containing relatively high amounts of DeCLIC-MBP because they also contain maltoporin.

Considering preliminary results from the group showing that low pH and exclusion of calcium promote an apparent open state of the channel, an attempt to do the purification entirely at pH 5 without adding calcium, by using magnesium as a cofactor in thrombin cleavage instead, was initially done with the hypothesis that it would help stabilize the protein in a single conformation and the addition of calcium only during sample preparation for cryo-EM could promote an intermediate state. This turned out to not be possible due to two main reasons; the low pH appeared to cause aggregation during the overnight elution of the amylose affinity column, as well as interfered with thrombin's ability to digest the protein (see appendix, figure A3). The latter observation prompted the thrombin activity assay which clearly showed that digestion at pH 5 does not work. This is probably connected to the catalytic pH range for thrombin being in the range of pH 5-10, with an optimum at 8.3 [26]. This assay also showed that magnesium is not as effective cofactor as calcium. It did however cleave at least some of the protein, meaning that it is possible to avoid the addition of calcium but at the cost of a significantly lower yield.

The data collected using cryo-EM resulted in an electron density with a 3.4 Å resolution. This is the global resolution of the density as a whole. However, the resolution can vary across different parts of the protein. The estimation of local resolution can thus tell a more complete story. Inspection of the local resolution of the density (figure 7) shows that the NTDs are resolved to the lowest resolution, which is to be expected since these are very flexible and can thus be captured in many different positions in different molecules. The second-lowest resolution can be observed in the M4 helices, especially near the top and bottom of the TMD. This could be explained by the detergent micelle being denser in these regions and effectively "hiding" the atoms of the protein. This may also explain why the center of the pore and lower part of the ECD is the best resolved, there is no micelle or NTDs obstructing the view in these regions.

One obstacle encountered during the processing of the data was the initial problems with processing the data in C5. Processing in C5 effectively multiplies the number of input particles by five, which helps boost the obtainable resolution significantly. However, this data set could initially only be processed in C1 (no imposed symmetry) which resulted in a 4 Å density generated from one million particles. This set of one million particles had been generated from multiple rounds of 2D and 3D classifications, where in each round very few particles ended up in bad classes, suggesting that the remaining particles are quite homogenous. Still, C5 could not be imposed. This could be related to the large number of particles still included,

since C5 symmetry will also multiply the effect of the bad particles still present in this set of one million despite it appearing homogenous. Only when selecting only the best-resolved class after 3D classification, which was only marginally better than the other classes, could C5 symmetry be imposed, presumably due to the great reduction in the number of particles. However, data set of similar quantity has previously been processed in C5 without a problem from the beginning, so there appears to be something different with this data set. It is possible that the protein indeed is in some intermediate state creating an asymmetric distribution of confirmations across subunits. However, if this asymmetry is prominent enough for RELION to not be able to impose C5 symmetry, it should also be prominent enough to be observed in the reconstructions produced in C1, which all appear symmetric.

Comparing the calcium-present cryo-EM density produced in this project to the previously determined (unpublished) calcium-free cryo-EM density reveals several differences. Both the NTDs and the ECD are more extended in the calcium-present density while the TMD is more contracted and slightly twisted. These conformational differences are consistent with what was reported by Hu et al. [3], regarding the open versus closed conformation of the protein. Comparing the calcium-present cryo-EM density to the previously published x-ray structure of the closed channel (PDB ID: 6V4S) shows a good fit of the model inside the density, suggesting that the channel is in a closed-like conformation at pH 5 in the presence of calcium. There are, however, several differences that may point to a different state. The most noticeable difference in comparison to the model is the position of the NTDs, which is the only part of the density that does not line up with the model at all. This can not be viewed as a significant difference, however, due to the low resolution in this region and the high flexibility of the NTDs. A less noticeable but potentially more significant difference is an apparent extension in the ECD of the calcium-present cryo-EM density. Looking at the overlay of the density and the model shows that the model appears to consistently sit on the inner part of the density rather than in the middle, especially in the beta-strands. Additionally, in comparison to the surface representation of the x-ray crystal structure, the EDC shows a very similar pattern of difference as in the comparison of the calcium-present and calcium-free cryo-EM densities, with the corresponding sequences consistently being further out in the calcium-present cryo-EM density than in the surface representation of the x-ray crystal structure, but with a much smaller distance. This observation is also confirmed when looking at the most constricting point in the ECD pore, where the calcium-present cryo-EM density has the smallest constriction, followed by the surface representation of the x-ray crystal structure, and the calcium-free cryo-EM density being the most constricting. Since an extended ECD is associated with the closed state, this would indicate a slightly more closed state of the protein in the calcium-present cryo-EM density compared to both other conditions.

Interestingly, this difference in the ECD is not consistent with the differences observed in the TMD pore, where the calcium-present cryo-EM density appears to be slightly more open than the x-ray crystal structure, suggesting that the low pH promotes a more open pore. This is primarily based on the observation of the pore diameter in the most constricting points, both of which are located near the position of the two hydrophobic residues identified as the main hydrophobic gates [3], Phenylalanine at the 16´ position and Leucine at the 9´ position. Comparison of the sidechain densities for these two positions shows that in both positions the surface representation of the x-ray crystal structure is extending quite far beyond the calcium-present cryo-EM density, making the pore about 4 Å smaller in diameter. It is important to note, however, that the surface representation created for the x-ray crystal structure is not a perfect representation of how the electron density used to build this model looked. Thus, the comparison of these two densities is not entirely accurate, and to be confident in this observation a comparison would have to be between built models rather than densities. However, the same type of analysis was performed on the 4 Å resolution cryo-EM density mentioned above, and this difference in the TMD was increased when moving to a higher resolution, evidence that is reinforcing the idea that this is a true alternative state, although not definitive.

In the calcium-binding site, there is a clear difference in the position of the same part of the loop-F when comparing the calcium-present cryo-EM density to both the x-ray crystal structure as well as the calcium-free cryo-EM density. Compared to the calcium-free cryo-EM density, loop-F has moved closer to the Cys-loop in the calcium-present cryo-EM density, presumably due to the presence of calcium. However, comparison to the x-ray crystal structure shows a rather different mode of calcium interaction at low pH. Looking at the calcium present cryo-EM density, the position of loop-F and its side chains appear to be as indicated in figure 11. According to this interpretation of the density, Glu479 (which is the residue in loop-F that interacts with the calcium in the x-ray crystal structure) is not in a position where it appears to interact with the calcium ion, possibly due to protonation of the residue. Since the binding of calcium has been found to affect the conformation of the whole protein, this alternative position of loop-F could be an explanation for the differences observed between the two calcium-present conditions.

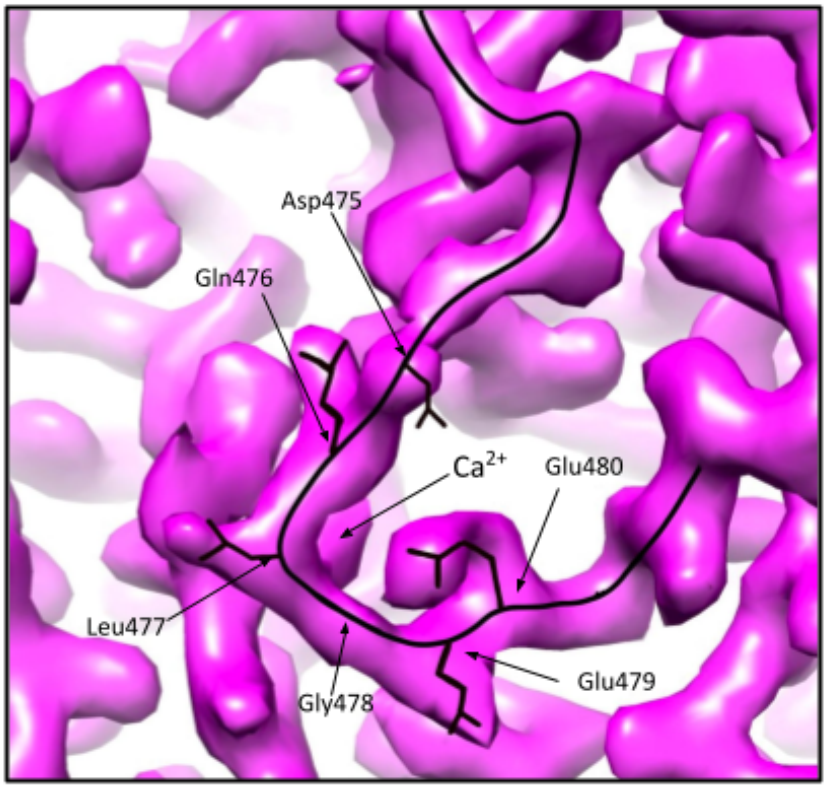


Figure 11: Calcium binding site from the same perspective as in figure 8B. Estimation of loop-F conformation has been drawn in (black) with approximate position of side chains of residues Asp475-Glu480.

Assuming the above discussed differences are true, the three conditions appear to promote three different states. Low pH without calcium allows the channel to adopt an apparent open conformation, while neutral pH and presence of calcium promotes a closed conformation. However, compared to neutral pH, the combination of calcium and low pH appear to cause the TMD to adopt a more open conformation, while the ECD stays in a more expanded confirmation, rather than contract as expected when the channel is opened. Thus it is possible that the channel is in an intermediate state at low pH in the presence of calcium where the channel has started to open up in the TMD but not yet contracted in the ECD or, alternatively, the ECD has already expanded but the TMD has not yet fully contracted. Taking all of these observations together, a possible mechanistic model could be drawn as in figure 12. In this model, the opening of the channel starts by the expansion of the TMD followed by the contraction of the ECD, and when the channel is closed, the expansion of the ECD is completed before the TMD is fully closed. Observations in the calcium-binding site suggest that the positioning of loop-F could be a major contributor to the overall state of the channel. This is all assuming the differences observed between the calcium-present cryo-EM density

and the surface representation of the x-ray crystal structure are true and not an artifact of the differences in the methodology for obtaining the two densities.

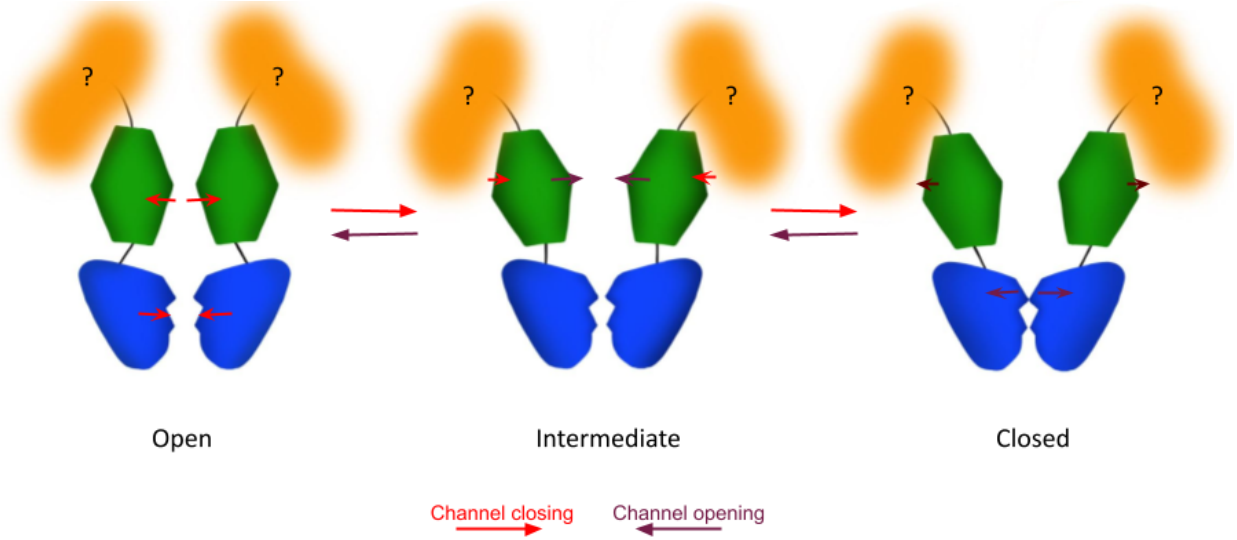


Figure 12: Possible mechanistic model for channel closing and opening based on observations. Domains are colored as in figure 2 and the starting of main movements of the domains are indicated by arrows.

In conclusion, a clear difference in channel conformation can be observed in the presence and absence of calcium, where the presence of calcium causes the TMD pore to contract and the ECD to expand in a way that is consistent with the previously reported effect of calcium. The evidence for the effect of pH is not as definitive as the presence of calcium but it appears to have an opposite effect in the ECD and TMD, promoting a more open-like confirmation in the TMD while promoting a more closed-like confirmation in the ECD, possibly depicting an intermediate state. Further analysis of a built model rather than a density is however necessary for the evidence to be definitive enough to confidently come to this conclusion.

5 Future perspectives

The obtainable information about the structural states of a protein from an electron density without a model is limited, especially in the confidence level of the observations made about structural properties of the protein. However, if a model based on the current density were built in the future, a much more conclusive analysis of the conformational states could be performed. Such a modal could also be used in molecular dynamics simulations to learn more about the transitional changes imposed by calcium and changes in pH. Another future perspective of this project would be to study lipid interactions with the channel to explore possible lipid modulations. Throughout the project, efforts have been made to study these

lipid interactions using mass spectrometry by screening several different types of detergents in an effort to find one that can provide the information needed for this type of study. These efforts have not yet been successful and the work will have to continue beyond the scope of this thesis. Learning more about how the channel is modulated would not only help the understanding of the channel's function and physiological role in itself, but also help provide more information about the family of pLGICs as a whole as well as help drug developers decide if DeCLIC or any of its close relatives would be a good model system for drug development.

6 Acknowledgments

The execution of this project would not have been possible without the invaluable help and guidance of my supervisors Erik Lindahl and Rebecca Howard, as well as my mentor Urška Rovšnik. Special thanks are extended to Urška Rovšnik for lending me her calcium-free density to use for the comparisons made in this thesis. I would also like to thank everyone in the group Molecular Biophysics Stockholm and everyone else on the γ3 floor at SciLifeLab for a very welcoming and helpful environment, as well as the Cryo-EM Swedish National Facility Stockholm.

7 References

[1] P.-J. Corringer, F. Poitevin, M. Prevost, L. Sauguet, M. Delarue, and J.-P. Changeux (2012) *Structure and Pharmacology of Pentameric Receptor Channels: From Bacteria to Brain.* Structure **20**(6), pp. 941-956

[2] C. daCosta and J. Baenziger (2013) *Gating of Pentameric Ligand-Gated Ion Channels: Structural Insights and Ambiguities*. Structure **21**(8), pp. 1271-1283

[3] H. Hu, R. Howard, U. Bastolla, E. Lindahl, and M. Delarue (2020) S*tructural basis for allosteric transitions of a multidomain pentameric ligand-gated ion channel.* PNAS **117**(24), pp. 1337-13446

[4] Á. Nemecz, M. Prevost, A. Menny, and P.-J. Corringer (2016) *Emerging Molecular Mechanisms of Signal Transduction in Pentameric Ligand-Gated Ion Channels.* Neuron **90**(3), pp. 452-470

[5]M. Jaiteh, A. Taly, and J. Hénin (2016) *Evolution of Pentameric Ligand-Gated Ion Channels: Pro-Loop Receptors.* PLoS ONE **11**(3)

[6] A. Tasneem, L. M. Iyer, E. Jakobsson, and L. Aravin (2005) *Identification of the prokaryotic ligand-gated ion channels and their implications for the mechanisms and origins of animal Cys-loop ion channels.* Genome biology **6**(1):R4

[7] A. Miyazawa, Y. Fujiyoshi, and N Unwin (2003) *Structure and gating mechanism of the acetylcholine receptor pore* Nature **423**, pp. 949–955

[8] Z.-s. Wu, H. Cheng, Y. Jiang, K. Melcher, and E. Xu (2015) *Ion channels gated by acetylcholine and serotonin: structures, biology, and drug discovery.* Acta Pharmacology Sin **36**, pp. 895-907

[9] J. W. Lynch, Y. Zhang, S. Talwar, A. Estrada-Mondragon (2017) *Chapter Eight - Glycine Receptor Drug Discovery* Editor(s): Dominic P. Geraghty, Lachlan D. Rash, Advances in Pharmacology, Academic Press Volume 79, pp. 225-253

[10] R. W.Olsen (2018) *GABAA receptor: Positive and negative allosteric modulators* Neuropharmacology **136**(A), pp. 10-22

[11] A. Bryson, R. J. Hatch, B.-J. Zandt, C. Rossert, S. F. Berkovic, C. A. Reid, D. B. Grayden, S. L. Hill, and S. Petrou (2020) *GABA-mediated tonic inhibition differentially modulates gain in functional subtypes of cortical interneurons* PNAS **117**(6), pp. 3192-3202

[12] E. Nogales and S. H.W. Scheres (2015) *Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity.* Mol. Cell. **58**(4), pp. 677-689

[13] The Royal Academy of Science (2017) *The Nobel Price in Chemistry 2017* Press release, Oct 4th, 2017

[14] J. L. S.Milne, M. J. Borgnia, A. Bartesaghi, E. E. Tran, L. A. Earl, D. M. Schauder, J. Lengyel, J. Pierson, A. Patwardhan, S. Subramaniam (2013) *Cryo-electron microscopy - a primer for the non-microscopist* FEBS Journals **280,** pp. 28-45

[15] T. Bendory, A. Bartesaghi and A. Singer (2020) *Single-Particle Cryo-Electron Microscopy: Mathematical Theory, Computational Challenges, and Opportunities*. IEEE Signal Processing Magazine **37**(2), pp. 58-76

[16] Y. Cheng, N. Grigorieff, P. A. Penczek, and T. Walz (2015) *A primer to Single-Particle Cryo-Electron Microscopy* Cell **161**(3), pp. 438-449

[17] T. M. de Oliveira, L. van Beek, F. Shilliday, J. É. Debreczeni, and C. Phillips (2020) *Cryo-EM: The Resolution Revolution and Drug Discovery* **26**(1), pp. 17-31

[18] S. H. W. Scheres (2012) *RELION: Implementation of a Bayesian approach to cryo-EM structure determination* Journal of Structural Biology **180**(3), pp. 519-530

[19] A. Punjani, J. L. Rubinstein, D. J. Fleet,and M. A. Brubaker (2017) *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination*. Nat Methods **14**(3), pp. 290-296

[20] T. Moriya., M. Saur, M. Stabrin, F. Merino, H. Voicu, Z. Huang, P. A. Penczek, S. Raunser, and C. Gatsogiannis (2017) *High-resolution Single Particle Analysis from Electron Cryo-microscopy Images Using SPHIRE*. J. Vis. Exp. **123**, e55448

[21] J. Zivanov, T. Nakane, B. Forsberg, D. Kimanius, W. Hagen, E. Lindahl, S. Scheres (2018) *New tools for automated high-resolution cryo-EM structure determination in RELION-3.* eLife **7,** e42166

[22] S. Zheng, E. Palovcak, J.-P. Armache, K. Verba, Y. Cheng, and D. Agard (2018) *MotionCor2 - anisotropic correction of beam-induced motion for improved cryo-electron microscopy* Nature methods, **14**(4), pp. 331–332

[23] A. Rohou and N. Grigorieff (2015). *CTFFIND4: Fast and accurate defocus estimation from electron micrographs.* Journal of structural biology **192**(2), pp. 216–221.

[24] E.Pettersen, T. Goddard, C. Huang, G. Couch, D. Greenblatt, E. Meng, and T. Ferrin (2004) *UCSF Chimera—A visualization system for exploratory research and analysis.* J Comput Chem **25**, pp. 1605–1612

[25] L. Kullman, M. Winterhalter, and S. M. Bezrukov (2002) *Transport of Maltodextrins through Maltoporin: A Single-Channel Study.* Biophysical Journal **82**(2), pp. 803-812

[26] Sigma-Aldrish® (2017) *Thrombin from human plasma (T6884) - Datasheet*

8 Appendix

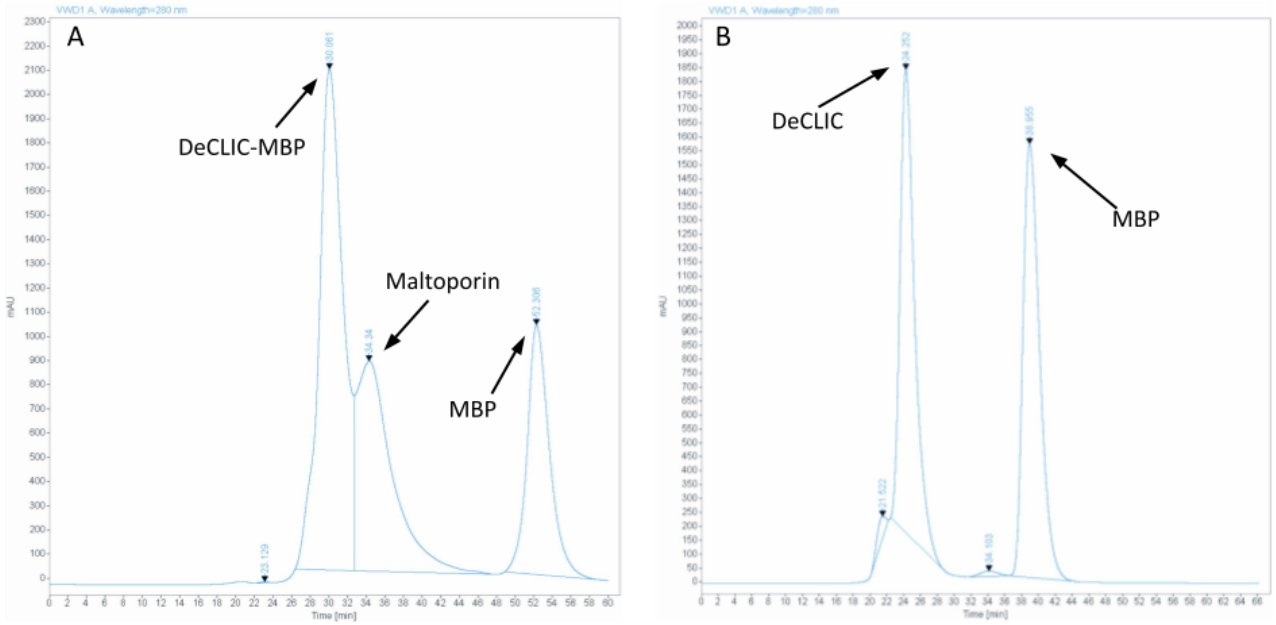


Figure A1: Original chromatograms traced over in figure 4. A) Same as in figure 4 A. B) Same as in figure 4 B.

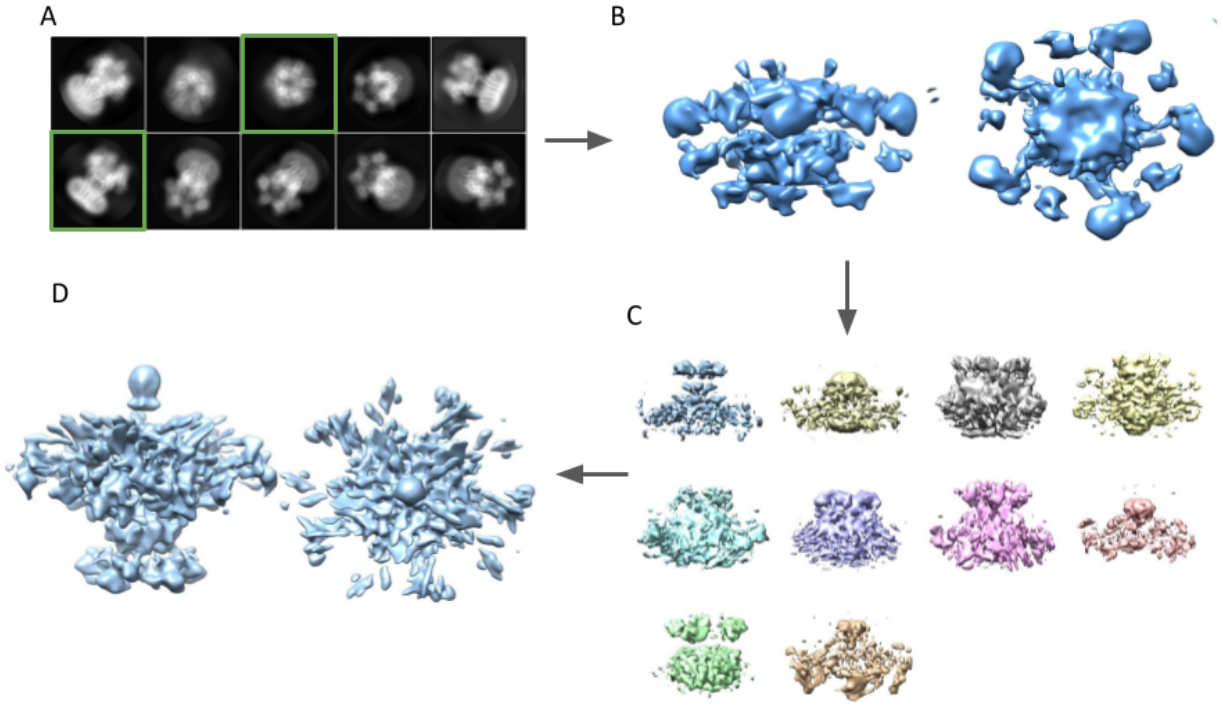


Figure A2: Main outputs of processing pipeline when imposing C5 symmetry. (A) Initial 2D classes selected after classification of auto-picked particles. Classes marked in green were used to generate the initial model. (B) Initial model generated from 2D classes. C) 3D classes generated from 3.3 million particles using the initial model in figure 6B as reference. D) 3D refined density generated from 3.3 million particles using the initial model in figure 6B as reference.

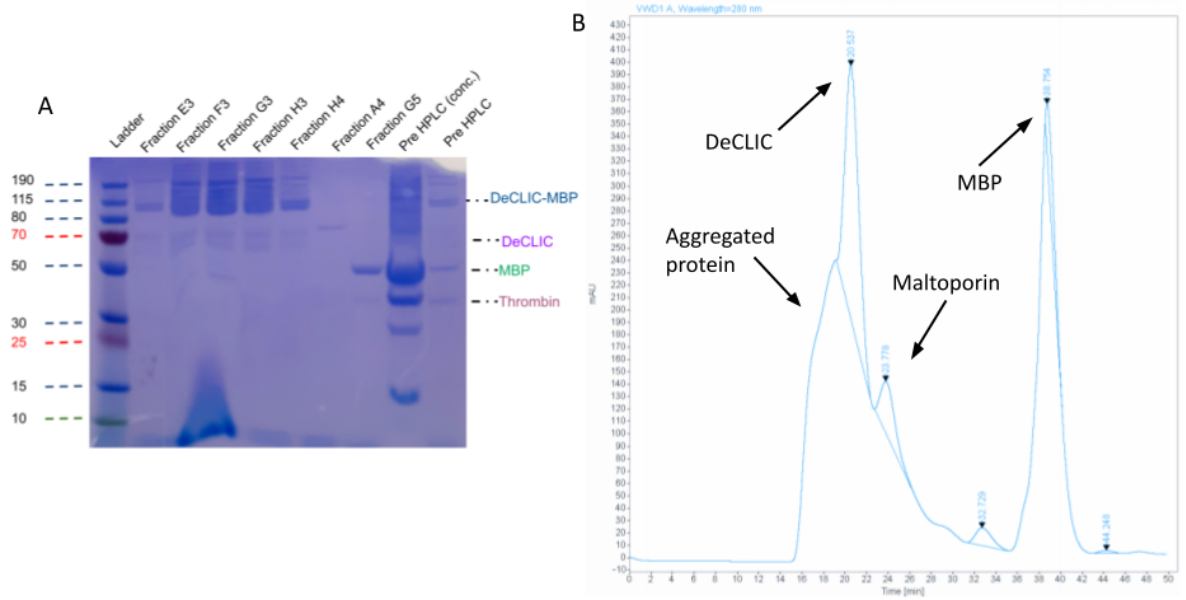


Figure A3: A) SDS-PAGE gel with fractions from SEC after failed Thrombin cleavage at pH 5 showing a thick band at the expected size of DeCLIC-MBP but no cleaved DeCLIC. B) Chromatogram from SEC of one of the elutions done at pH 5.