PhD Thesis
Sarah Spruce Preisler

Recombinant production and purification of ion channels for biomimetic membranes and cryo-EM studies

Supervisor: Per Amstrup Pedersen
This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen
Submitted on: 30th of September 2019
THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Name of department: Department of Biology, The Faculty of Science, University of Copenhagen

Author(s): Sarah Spruce Preisler

Title: Recombinant production and purification of ion channels for biomimetic membranes and cryo-EM studies

Supervisor: Per Amstrup Pedersen

Submitted on: 30th of September 2019
# Table of contents

PREFACE ............................................................................................................................ 5

ACKNOWLEDGEMENTS ................................................................................................. 7

SUMMARY ......................................................................................................................... 9

SUMMARY IN DANISH ..................................................................................................... 10

PUBLICATIONS INCLUDED IN THIS THESIS ................................................................. 11

LIST OF ABBREVIATIONS .............................................................................................. 12

CHAPTER 1 – INTRODUCTION ....................................................................................... 13

An overlooked approach to renewable energy production .................................................. 13

Biological membranes and membrane proteins .................................................................. 15

Biomimetic membrane can improve reverse electro-dialysis ............................................. 16

Ion Channels .................................................................................................................... 17

KcsA ..................................................................................................................................... 17

NavAb ............................................................................................................................... 18

ClC-ec1, ClC-0 and ClC-1 ................................................................................................. 19

Objectives ......................................................................................................................... 22

CHAPTER 2 – MANUSCRIPT I ......................................................................................... 23

Synopsis ................................................................................................................................ 23

*Saccharomyces cerevisiae* as a superior host for production of prokaryotic membrane proteins 24

CHAPTER 3 – MANUSCRIPT II ........................................................................................ 57

Synopsis ................................................................................................................................ 57

Homologous recombination in yeast: A tailor made tool for manipulating expression plasmids 58

CHAPTER 4 – MANUSCRIPT III ....................................................................................... 80
Synopsis ........................................................................................................................................................................... 80

Design and purification of constitutively open potassium, sodium and chloride channels - application for creating ion selective membranes .................................................................................................................................. 81

CHAPTER 5 – JOURNAL ARTICLE I................................................................................................................................. 109

Synopsis ........................................................................................................................................................................... 109

Structure of the human ClC-1 chloride channel ........................................................................................................ 110

CHAPTER 6 – DISCUSSION AND OUTLOOK .............................................................................................................. 111

CHAPTER 7- CONCLUSION ........................................................................................................................................... 117

REFERENCES.............................................................................................................................................................. 118

SUPPLEMENTARY DATA CLC-1 AND CLC-0 ........................................................................................................... 121
Preface

This dissertation titled “Recombinant production and purification of ion channels for biomimetic membranes and cryo-EM studies” was written for the final evaluation for obtaining the degree as Doctor of Philosophy (PhD) from the Faculty of Science, University of Copenhagen, Denmark. This thesis is the result of my 3 years of research under the supervision of Prof. Per Amstrup Pedersen at the Department of Biology, section for Cell Biology and Physiology. It includes collaboration with three external labs.

- The lab of Professor Dan A. Klærke at the Faculty of Health and Medical Sciences, University of Copenhagen, where the single channel currents were recorded.
- The lab of Associate Professor Pontus Gourdon at the Faculty of Health and Medical Sciences, University of Copenhagen, where the cryo-EM samples were prepared and tested.
- Lastly I worked in the biotech company Aquaporin A/S to test the concept of using the ion channels to make ion selective membranes.

The research project was a part of MEMENTO: Membrane energy technology operations, funded by the Danish Innovation Foundation.

This thesis aims to explore the idea of using ion channels to design new biomimetic membranes selective to monovalent cations and anions for the purpose of generating sustainable energy from the mixing of saltwater and freshwater. To test this idea, different ion channels were chosen based on their stability or high ion flux for creating biomimetic membranes. Mutations of the ion channels were selected with the intend to remove any gating mechanisms and making them constitutively open channels, based on previously published data on these ion channels. It turned out that the bacterial chloride channel is a Cl⁻/H⁺ antiporter and therefore two Cl⁻ channels were also chosen as targets, in case it would not be possible to convert the antiporter into a channel. This led to the published cryo-Em structure of the human chloride channel CIC-1.

The majority of the work was to select the targets and mutations, make expression constructs and test the recombinant production and purification with the aim to scale up for industrial production. This scale up would demand stability of the proteins, low cost of procedure and establishing a robust and simple purification strategy.
In the end of the thesis project I worked to establish the reconstitution into polymersomes for the biomimetic membranes. This procedure turned out to require a lot of optimization for the selected ion channels, and therefore did not yield any reportable results for now.

This dissertation is written as a synopsis and has been divided into 7 chapters. An introduction to the research idea is given in the beginning in Chapter 1. The research is presented in Manuscripts in Chapter 2-5 which constitute the main part of my dissertation. Each manuscript has a synopsis in front, which outlines the background and motivation for the study.

Lastly, Chapter 6 gives a general discussion of the findings in this thesis and an outlook to the great possibilities for further research and novel applications of this innovative idea. Chapter 7 gives the final conclusions of this thesis.
Acknowledgements

“MEMENTO TE MORTALEM ESSE” – Roman proverb

First and foremost I want to thank my supervisor Professor Per Amstrup Pedersen for giving me the opportunity to work on this exciting project. It has been most appreciated that you were always available when I had questions. Thank you for your support and guidance, and I have enjoyed being part of PAPlab through the years.

I wish to thank Pontus Emanuel Gourdon and Kaituo Wang for the successful collaboration on the ClC-1 paper. Also a big thanks to the rest of the membrane structural biology group for letting me feel welcome and for all the help you provided. And a big thanks to Liying Zhang for the good teamwork and good company.

I also want to thank Dan Klærke, Kirstine Calloe and Vibeke Kristensen, who performed the electrophysiological measurements, and for always letting me feel welcome when I drop by.

Thanks to master students Lasse Kjærgaard and Marc Friis for providing much needed and appreciated extra hands and lab-hours for some of the work in this thesis, and helping me to learn how to supervise.

Big thanks to Julie Bomholt and Muntazim Munir Khan, who welcomed me to the lab in Aquaporin A/S where I stayed as part of my exchange

Special thanks to Claus Helix-Nielsen who made this PhD study possible, as part of the MEMENTO project funded by the Danish Innovation Foundation.

Thanks to all the friends and colleagues I have gained over the years in PAPlab. Thanks to Karen Molbæk Hansen and Julie Bomholt for the support and for your friendship. A big thanks to David Sørensen for technical assistance, and for talks on everything from music to life values, and for answering my many questions on house renovations. Thanks to Mads Beich-Frandsen for being an encouraging office-mate and for the scientific discussions. Thanks to Noah Kassem for
the valuable discussions and attending PhD courses together. Thanks to Casper Normann Nurup, Kasper Kvorning and Simon Krabbe for the good company and hours spent together.

At last I want to say thanks to friends and family who have supported me all the way, always grounding me and reminding me what life is really about. I always have your unconditional support and I am truly blessed with a loving and caring family and friends I can always count on.

Summary

Ion channels are responsible for the selective ion permeation across biological membranes, and their openings and closings are regulated by gating mechanisms. Mutating amino acids responsible for gating and by removing excess regulatory domains of selected ion channels, opens the possibility to utilize these ion channels for creating biomimetic membranes with high ion flux and selectivity. These ion selective membranes may then be used to create sustainable energy by reverse electro dialysis (RED), by using a salinity gradient across the membranes to capturing the free energy produced by mixing seawater and river water, and pose huge potential for overcoming the growing pressure on global energy resources.

Membrane proteins are inherently difficult to express and purify in large quantities. With the purpose of using ion channels for this biotechnological application, this thesis reports on establishing; a reliable cloning strategy to easily introduce mutations to the ion channels, find a high yield expression platform for producing these membrane proteins, and establish a robust and scalable purification protocol for obtaining large quantities of stable and pure membrane proteins.

Using Saccharomyces cerevisiae as the expression platform it was possible to produce and purify the ion channels to homogeneity and we were able show functionality of ion channels purified in detergent. This platform further enabled the structure determination using cryo-EM of the human chloride channel ClC-1.
Ionkanaler står for den selektive transport af ioner over biologiske membraner, og deres aktivitet er reguleret af åbne-lukke mekanismer. Ved at mutere aminosyrer der står for disse mekanismer og ved at fjerne overflødige regulatoriske domæner af udvalgte ionkanaler, åbnes muligheden for at bruge disse ionkanaler til at skabe biomimetiske membraner der har høj selektivitet og hurtig passage for ioner. Disse ionselective membraner kan bruges til at producere vedvarende energi ved omvendt elektro-dialyse "reverse-electro dialysis" (RED), som bruger en salt gradient henover membranerne til at høste den frie energi der produceres ved når havvand og flodvand blandes, hvilket har stort potential for at overkomme det stigende pres på de globale energieresurser.

Membranproteiner er svære at udtrykke og oprense i store mængder. Med det formål at anvende ionkanalerne i denne bioteknologisk platform, beskriver denne afhandling etableringen af en pålidelig kloningsstrategi til at introducere mutationer i ionkanalerne, finde en ekpressionsplatform med et højt udbytte til produktion af disse membranproteiner og etablering af en robust og skalerbar oprensningsprotokol til opnåelse af store mængder stabile og rene membranproteiner.

Ved anvendelse af *Saccharomyces cerevisiae* som ekpressionsplatform var det muligt at fremstille og oprense ionkanalerne til stor homogenitet, og vi var i stand til at vise functionalitet af ionkanaler oprenset i detergent. Denne platform muliggjorde endvidere strukturbestemmelsen af den humane kloridkanal CIC-1 ved anvendelse af cryo-EM.
Publications included in this thesis

1. *Saccharomyces cerevisiae* as a superior host for production of prokaryotic membrane proteins. 
   **Preisler, SS.**; Friis, M.; Kjærgaard, L.; Calloe, K.; Klærke, DA.; Pedersen, PA. (2019)  
   Manuscript, to be submitted.

   **Preisler, SS.** & Pedersen, PA. (2019)  
   Manuscript, Submitted.

3. Design and purification of constitutively open potassium, sodium and chloride channels - application for creating ion selective membranes. **Preisler, SS.**; Friis, M.; Pedersen, PA. (2019)  
   Manuscript, to be submitted.

4. Structure of the human ClC-1 chloride channel. Wang, K.; **Preisler, SS.**; Zhang, L.; Cui, Y.; 
   Missel, JW.; Grønberg, C.; Gotfryd, K.; Lindahl, E.; Andersson, M.; Calloe, K.; Egea, PF.; 
   Journal article.

Other publications

1. Characterization of Hailey-Hailey Disease-mutants in presence and absence of wild type 
   SPCA1 using *Saccharomyces cerevisiae* as model organism. Muncanovic, D.; Justesen, MH.; 
   Journal article.
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEM</td>
<td>Anionic exchange membranes</td>
</tr>
<tr>
<td>CBS</td>
<td>Systathionine-β-synthase</td>
</tr>
<tr>
<td>CEM</td>
<td>Cationic exchange membranes</td>
</tr>
<tr>
<td>CMC</td>
<td>Critical micellar concentration</td>
</tr>
<tr>
<td>Cryo-EM</td>
<td>Cryogenic electron microscopy</td>
</tr>
<tr>
<td>CTD</td>
<td>C-terminal domain</td>
</tr>
<tr>
<td>DDM</td>
<td>n-Dodecyl β-D-maltopyranoside</td>
</tr>
<tr>
<td>DM</td>
<td>n-Decyl β-D-maltopyranoside</td>
</tr>
<tr>
<td>FSEC</td>
<td>Fluorescence-detection size exclusion chromatography</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>Glu&lt;sub&gt;ex&lt;/sub&gt;</td>
<td>External glutamate</td>
</tr>
<tr>
<td>Glu&lt;sub&gt;in&lt;/sub&gt;</td>
<td>Internal glutamate</td>
</tr>
<tr>
<td>His&lt;sub&gt;10&lt;/sub&gt;-tag</td>
<td>10 x Histidine tag</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized metal affinity chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-d-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>LDAO</td>
<td>Lauryldimethylamine N-oxide</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>MEMENTO</td>
<td>Membrane energy technology operations</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>P-loop</td>
<td>Pore loop</td>
</tr>
<tr>
<td>PTM</td>
<td>Post-translational modification</td>
</tr>
<tr>
<td>RED</td>
<td>Reverse electrodialysis</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td><em>Saccharomyces cerevisiae</em></td>
</tr>
<tr>
<td>SGP</td>
<td>Salinity-gradient power</td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco Etch Virus</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEC</td>
<td>Size-exclusion chromatography</td>
</tr>
<tr>
<td>TM</td>
<td>Transmembrane</td>
</tr>
<tr>
<td>VSD</td>
<td>Voltage sensing domain</td>
</tr>
</tbody>
</table>
Chapter 1 – Introduction

An overlooked approach to renewable energy production

Increased global urbanization, growing populations and rise in economies of less developed countries all lead to a growing pressure on energy resources in the future. In the light of climate change there is a need to replace fossil fuels with renewable energy resources.

Currently the main source of sustainable power in EU and US is biomass, hydropower and wind power [1, 2]. A less known approach is to harvest the energy of mixing waters of high and low salinity, named Salinity-gradient power (SGP). Reverse electrodialysis (RED) is a membrane based SGP technology which can be used to harvest the energy of mixing sea and river water. The energy of mixing 1 m³ seawater with 1 m³ river water is 0.256 kWh and the global potential is 25 PWh(10¹⁵) [3] - Denmarks yearly consumption is 33 Billion kWh [4]. The global energy yield is therefore theoretically 1000 times that of Denmark’s total yearly energy consumption.

A RED unit comprises membrane stacks of alternating cationic exchange membranes (CEM) and anionic exchange membranes (AEM), collectively named ion exchange membranes (Figure 1.1). These membranes separate sea and river water, and allow cations and anions to flow in opposite directions down their concentration gradient. This movement of charge in opposite directions creates a high electrochemical potential difference over each membrane pair and sums up over all membrane stacks to give the total potential difference. This electrochemical potential difference can then be converted into an electric current at the electrodes.
Figure 1.1. Schematic diagram of a membrane stack for RED. Alternating anion exchange membranes (AEM) and cation exchange membranes (CEM) separate fresh water (light blue arrow) and salt water (dark blue arrow). Ions can move across the permselective membranes down the salinity gradient, moving positive and negative ions in opposite directions, creating a salinity gradient over the membrane stack. The energy is harvested as an electric current from the difference in electrochemical potential created; a (reversible) redox reaction is used to transfer the ionic current into an electrical current. For practical application, the system will consist of many membrane stacks. The figure was adapted from [5].

The concept of harvesting energy from mixing fresh and seawater using small membranes, was first described in 1954 [6]. Over the next five decades, apart from the development of theoretical model of the power production of RED, only small advances were made in the development of the method. Not until the early 2000s, did the amount of research in this field increase dramatically, at same time as an great increase in oil prices[7], hereby increasing the interest for alternative energy sources.

The key property of ion exchange membranes is the membrane resistance and the permselectivity; the ability to separate solutes and discriminate between cations and anions,
allowing specific ions to cross. The membrane resistance influences the power output and depends largely on the polymer matrix used to make the membranes. The permselectivity is based on the charged functional groups of the polymer; to allow passage of cations and reject anions, CEMs contain negatively charged groups; to allow passage of anions but exclude cations, AEMs contain positively charged groups [7].

Optimization of components of a RED unit; membranes, spacers and electrodes, among other; has been vital for increasing RED power output performance/power density. Even though there has been huge advances in the design and materials used RED technology, there is still a need for improvements for efficient power generation[7]. Multivalent ions have been shown to be an issue for ion exchange membranes, lowering the efficiency (power density) of the membranes. Unfortunately standard ion exchange membranes have low valance selectivity [8].

**Biological membranes and membrane proteins**

Biological membranes are barriers between all cell compartments and consist mainly of lipids, surface carbohydrate and embedded proteins. Membrane lipids are amphipathic with a polar head-group and a hydrophobic carbon-tail, and it is largely the non-covalent interaction between the hydrophobic tails that holds the membrane together. The lipid bilayer makes up a dynamic but stable barrier where passive diffusion only occurs for small ions and non-charged molecules at a very low rate, meaning that transport of ions, nutrients and solvent need to be facilitated by membrane embedded transport proteins.

These membrane embedded transport proteins can be divided into channels, pumps and transporters (FIGURE 1.2). Channels facilitate passive transport down a concentration gradient, pumps use energy to move molecules against their concentration gradient and transporters can further be divided into uniporters, symporters and antiporters. Uniporters transport one molecule down a concentration gradient, while symporters and antiporters transport one molecule against its concentration gradient, facilitated by the movement of one more molecules down the concentration gradient in the same or opposite direction.
Evolution has optimized ion channels for the fast and specific transport of ions, often $10^6$ ions per second or greater [9]. Ion channels are found in all kingdoms of life and are the driving force behind vital functions like cell volume regulation, establishing the resting membrane potential, shaping action potentials. They are classified by their ion selectivity and gating mechanism, meaning what ion they transport and what type of stimulus that opens and closes the ion pore. Voltage-gated ion channels respond to change in membrane potential, ligand-gated ion channels respond to binding of a ligand molecule, mechanosensitive ion channels respond to change in membrane and there are many other types of gating [9].

**Biomimetic membrane can improve reverse electro-dialysis**

Biomimetics, synthetic biology and bioengineering, all have slight variations of the same scope: To use the brilliance of biological design and processes to develop new technologies. This will often require adjusting or enhancing specific features of a natural system by manipulating biological, chemical or physical properties. Biomimetic membranes use specific biological elements or inspiration to improve efficiency and specificity of artificial membranes.

The key property of biomimetic membranes is the membrane ability to separate or enclose solutes with permeation of a specific molecule. Biomimetic membranes can consist of a wide range of materials with different structure and functions and they can generally be divided into nanostructured polymer membranes, amphiphilic block copolymer membranes and lipid bilayers.
Nanostructured polymer membranes have nanopores with a defined diameter and are coated with functional molecules with a specific charge, hydrophobicity or chemical reactivity, to enhance permeation of certain molecules [10]. Amphiphilic block copolymers are often made from diblock or triblock copolymers, which have either two or three alternating hydrophobic and hydrophilic blocks. Changing the block length and length ratios or functional groups of the polymer blocks makes it possible to customize this type of membrane. Further these membranes can be either free-standing or immobilized onto supports depending on the application and the mechanical stability needed.

Lipid bilayers and amphiphilic block copolymer both form bilayers capable of incorporating membrane proteins but differ in stability, permeability, thickness and flexibility. It is then the membrane proteins intrinsic property, which allows certain molecules to cross these membranes [11].

The state-of-the art approach is to take advantage of the highly selective ion channels embedded in biomimetic membranes to allow permeation of monovalent ions only. Modification of the ion channels will further assist the design of next generation ion-selective biomimetic membranes. Specific modifications are described in chapter X, but the general idea is to remove any kind of gating or unnecessary parts of these membrane proteins while still maintaining the specific and fast ion transport.

Using membranes with ion channels incorporated for RED should help overcome the problem with passage of multivalent ions through the membranes, and improve the efficiency of this technology for sustainable energy production.

**Ion Channels**

This section gives a brief description of the ion channels included in this thesis, their structures can be seen in Figure 1.3

**KcsA**

The prokaryotic potassium channel KcsA from the soil bacteria *Streptomyces lividans* represents the basic unit of a potassium selective ion channel. KcsA is a tetramer with a pore in the middle, formed by the four subunits. Each subunit has two transmembrane domains TM1 and TM2 and
the two helices are linked by an extracellular loop and a short so-called pore helix. The carboxyl end of this helix points towards the center of the channel pore and is followed by the selectivity filter on the pore loop (P-loop). The selectivity filter is a narrow pore of ~3Å, with a signature sequence TVGYG [12]. This sequence is highly conserved within K⁺ channels and provides stabilizing interactions between K⁺ ions and the pore. The stabilizing interaction of the selectivity filter is able to accommodate high ion-flux by multiple ions passing through the pore in a single line[13].

Voltage dependent gating behavior of KcsA has been a mystery, since it does not contain a voltage sensing domain (VSD). Instead the selectivity filter is believed to confer voltage dependence, by reorientation of Glutamine71 in response to voltage[14]. Lastly, C-type inactivation is a mechanism which stops the channel from opening, and is related to the degree of opening of the selectivity filter, but the mechanism is not yet fully understood [15].

The opening and closing of the channel is regulated by two mechanisms: Activation gating and inactivation gating. Activation gating opens and closes the pore in response to voltage and the inactivation gating acts if the activation gate keeps the channel open for a long time, termed C-type inactivation. The activation gate is formed by the so-called bundle crossing of the carboxyl-terminal ends of the four TM2 helices. This bundle crossing sterically limits the access to the pore and is held together by hydrogen bonds. This stabilizing interaction is disrupted upon binding of protons, leading to opening of the bundle activation gate, hence regulated by pH [13].

NavAb

NavAb from *Arcobacter butzleri* belong to the family of bacterial voltage gated sodium channels and their biological functions is not understood yet, but motility has been suggested [4].

NavAb has 6 transmembrane helices (S1-S6). Four subunits come together to form a homotetramers with a central pore surrounded by four VSDs. The VSD of one subunit packs around the pore-domain of the neighboring subunit. S1-S4 make up the VSD and S4 contain arginine residues which act as gating charges [5, 6]. These residues enable the VSD to move up and down in the membrane, in response to change in membrane potential. This opens the pore in the center of the protein where sodium ions flow through.
The pore domain is formed by S5 and S6 and two pore-helices (P1 and P2), with a P-loop containing the selectivity filter. The side chains in this pore convey the selectivity and the pore is thought to bind multiple ions and the same time with a “knock-off” mechanism, elucidated from molecular dynamics (MD) simulations [7].

The first structures of NavAb did not reveal the C-terminal domain (CTD) [5, 6] and was first seen in another Nav protein structure [8]. The structure was later solved in two different conformations for NavAb [9]. The CTD form a four-helix bundle from residue I217 to the C-term and have two regions; a so-called neck region next to the membrane and a C-terminal coiled-coil domain. The bundle crossing moves to open and close access of ions through the pore, with narrowing points located at I217 and M221, acting as a hydrophobic gate [9].

Depolarizing membrane potentials over a long time or repetitive short pulses have been shown to induce a state of inactivation in NavAb, which stops the pore from re-opening over a short or a long time scale [10].

**ClC-ec1, ClC-0 and ClC-1**

The family of chloride transporting membrane proteins consists of both passive channels and secondary active transporters and they are found in organisms across all kingdoms of life [16, 17]. Chloride channels facilitate movement down the Cl⁻ concentration gradient, while transporters transport two Cl⁻ in exchange for one H⁺, resulting in Cl⁻ transport against a concentration gradient and H⁺ down a gradient, or vice versa.

These proteins have a high degree of sequence and structure similarity, which means that the line between channels and pumps is blurry. It has been a puzzle how they can have similar sequence, and accommodate different mechanisms for transporting ions. I turns out that it is minor difference in the ion passage way which separate them from each other, however the exact molecular elements of ion transport is still not fully understood. Paradoxically, while most functional work has been carried out on Cl-channels, all structures solved were Cl⁻/H⁺ transporters up until 2017 where the first bovine channel structure was determined [18], closely followed by the structure of the human ClC-1 channel [19, 20].

General for these chloride transporting proteins are that they are homodimers with an hourglass-shaped ion passage through the protein, and are made up of eighteen helices in each monomer...
The selectivity filter is 15 Å long located in the middle of the protein with three Cl- binding sites seen in crystal structures (S_{int}, S_{cen}, S_{ex}) [21].

Selectivity for Cl\(^-\) over Br\(^-\) and I\(^-\) is very low, presumably because Cl\(^-\) is the only abundant small anion in biological systems[22]. Larger anions like phosphate, sulfate and bicarbonate are too large to pass the selectivity filter[23].

The bacterial CIC Cl\(^-\)/H\(^+\) exchange transporter, CIC-ec1 originates from *E. coli*. The monomer size is 50 kD and it has very short N and C termini, and is therefore mostly embedded in the membrane. [24]. The proton pathway in CIC-ec1 has been extensively studied but remains puzzling, since H\(^+\) is invisible in crystallographic structures, very little is understood about the molecular mechanism of Cl\(^-\)/H\(^+\) exchange. The current model has a glutamate residue (internal glutamate, Glu_{in}), which is only present in Cl\(^-\) transporters, acting like H\(^+\) transfer residues [25, 26]. This creates a transfer route; same place as Cl\(^-\) on the external side, with a different entry than Cl\(^-\) on internal side.

The structure of the CIC channels CIC-1 and CIC-0 was suggested to have same overall structure and fold as CIC-ec1 based on homology models, with the difference of having two CBS domains in the C-terminal end of these proteins, a fold found in many proteins. The CIC-1 monomer is 100 kDa and the CIC-0 is around 90 kDa.

The molecular ion transport properties of CLC proteins were first based on electrophysiology of CIC-0, showing double-barrel channel behavior. This revealed two modes of channel gating; fast gating where each pore opens and closes independently, and common gating where both pores are either open or closed [27]. Gating mechanisms have been shown to be voltage dependent and fast gating is regulated by pH, Cl\(^-\) and voltage.

CIC-ec1, CIC-1 and CIC-0 share the same fast gating mechanism, facilitated by a negatively charged carboxylate group from Glutamate (external Glutamate,Glu_{ex}). This sidechain occludes Cl\(^-\) ions from passing through the pore, when the carboxyl is protonated the side chain moves away and room for chloride ion to enter [21].
Figure 1.3. Structure of ion channels. A, B. Tetrameric structure of KcsA and NavAb shown as side-view and top-view, with the location of mutations indicated in yellow spheres and sequence removed indicated in red. C. The dimeric structure of ClC-ec1 with the monomers colored cyan and purple, shown as side-view and top-view. Mutations are indicated by yellow spheres. D. The dimeric structure of ClC-1 shown as side-view and top-view, representative also for ClC-0, with monomers colored in gray and orange. Mutations are indicated by yellow spheres and the removed CBS domains are colored red. Note that the size of the CBS domains varies between the two chloride channels. Figure was adapted from Manuscript III.
Objectives

This work intends to form the basis for exploring the idea of developing new biomimetic membranes, selective only to monovalent cations and anions to give a high RED membrane ion separation, hereby ensure efficient energy production. To be able to do this, constitutively open ion channels in large quantities would be needed. The specific aims towards this goal are listed below:

1. Identify suitable robust bacterial cation and anion selective channels with high ion conductance
2. Design of constitutive open ion channels from the chosen targets
3. High-yield protein production, Identifying best suited expression platform and establishing a robust and scalable purification procedure
CHAPTER 2 – Manuscript I

Synopsis

Apart from being interesting targets for future bio-based technologies, membrane proteins are targets for many pharmaceutical drugs [28]. Genomic projects have established that 20-30% of genes of a genome encode membrane proteins [29, 30]. This stresses the importance of studying the structure and function of membrane proteins. A fundamental challenge is the expression and purification of large quantities of membrane protein, whether it is needed for bio-technological application or for structural and functional studies.

First step is to have a cost-effective recombinant expression system, which can express high yields of correctly folded membrane proteins. In the manuscript we explore the expression capacity of *E. coli* and *Saccharomyces cerevisiae* for four different membrane proteins; AqpZ, ClC-ec1, KcsA and NavAb. Both expression systems have the advantage of being easy to cultivate, only requiring low-cost media and have well characterized genomes, making them easy to genetically modify to accommodate membrane protein expression [31]. The membrane proteins were expressed from expression plasmids transformed into either *E. coli* or *S. cerevisiae*, and screened for expression via a GFP tag.

All four membrane proteins had previously been purified and crystalized from *E. coli*, and we set out to test if they could also be purified in stable and active form from Yeast.

The present manuscript describes the use of Yeast for expressing bacterial membrane proteins, which has not been exploded in detail before. Our results suggest that *S. cerevisiae* should be considered for production of bacterial membrane proteins on a more regular basis.
Saccharomyces cerevisiae as a superior host for production of prokaryotic membrane proteins

Sarah Spruce Preisler¹, Marc Friis¹, Lasse Kjærgaard¹, Kirstine Callø², Dan Klærke², Per Amstrup Pedersen¹*

¹ Department of Biology, University of Copenhagen, Universitetsparken 13, DK-2100 Copenhagen OE, Denmark
² University of Copenhagen, Department of Veterinary and Clinical Animal Science, Frederiksberg, DK-1870, Denmark

*Correspondence to PAPedersen@bio.ku.dk

Abstract

Background: Escherichia coli is often used for recombinant expression of bacterial membrane proteins. However membrane proteins expression levels are often low or they are expressed in inclusion bodies, which require refolding the membrane proteins.

Results: We compared the expression capacity of Saccharomyces cerevisiae to that of Escherichia coli for four different membrane proteins. We demonstrate that using Saccharomyces cerevisiae as the expression platform enables production and purification of high yields of the selected membrane proteins, and we were able to confirm that two of the membrane proteins retained activity after purification in detergent. Further, we observed an improvement of expression yield and detergent solubilization yield by altering tag localization for one of the target proteins.

Conclusion: Our findings demonstrate the great potential of using Saccharomyces cerevisiae for recombinant production of bacterial membrane proteins.

Introduction

Integral membrane proteins are responsible for efficient and selective transport of organic and inorganic molecules across cellular membranes and many signalling processes rely on membrane-embedded receptors. The impact of integral membrane proteins is emphasized by the fact that they constitute around 30% of the proteome across all kingdoms of life [1].
Access to high resolution structures of the membrane proteome is required to comprehend membrane biology. While more than 120,000 structures are deposited in the Protein Data Bank, presently only 948 are high resolution structures of integral membrane proteins (http://www.blanco.biomol.uci.edu/mpstruc/). This bias exposes how challenging it is to deal with membrane protein structural biology.

Obtaining high expression levels is usually a major bottleneck for producing membrane proteins in large quantities. Next a purification procedure, that preserves protein fold and activity, needs to be established. This is laborious as it generally involves screening the behavior of a large number of homologous proteins from a variety of species.

Prokaryotic membrane proteins for structural biology have almost exclusively been supplied from recombinant production in *Escherichia coli* (*E. coli*) [2], which may partly be attributed to fast and cheap cultivation. Furthermore, *E. coli* has been extensively studied, so it comes with a large toolbox of biochemical methods, including advanced tools for recombinant gene expression. Traditionally it is thought that for recombinant production of prokaryotic membrane proteins, the most suitable lipid environment is provided by *E. coli*; the concept being that the closest related host, is likely to yield the highest amount and quality of the desired proteins [3].

In the present paper we challenged this dogma, as we hypothesized that *Saccharomyces cerevisiae* (*S. cerevisiae*) may have a number of advantages for expression of prokaryotic membrane proteins. First, we avoid having protein expressed in inclusion bodies, hereby circumventing the requirement for membrane protein refolding, which at its best is inefficient, but most often impossible due to the particularly complicated nature of this class of proteins [2, 4]. Secondly, yeast possesses advanced machinery for membrane protein biogenesis that may support correct folding of membrane proteins better than *E. coli*. To investigate the capacity for expression of membrane proteins in *E. coli* and *S. cerevisiae*, we selected four prokaryotic membrane proteins of different bacterial origin and with different transport functions. We compared expression in the commonly used BL21/pET *E. coli* system [5] with expression in our previously described *S. cerevisiae* platform [6]. To our knowledge the capacity for production of prokaryotic membrane proteins in a eukaryotic host has not been thoroughly exploited before. We therefore tested if the proteins could be extracted from the cell membrane and purified in high yields while maintaining stability and activity. At last we tested the effect on expression levels and purification yields of tag localization for one of the target proteins.
Material and methods

S. cerevisiae and E.coli strains

S. cerevisiae recombinant protein production was carried out in PAP1500 (α ura3-52 trp1::GAL10-GAL4 lys2-801 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL) [6] in combination with the expression vector pEMBLyex4[7].

BL21(DE3)[5], BL21(DE3)-pLysS [5, 8], C41 [9] and C43 [9] were used as hosts for recombinant protein production in E. coli along with an E.coli - S. cerevisiae shuttle plasmid, pPAP10286, carrying the pET52(b) expression cassette (Supplementary Figure 1) [5, 10].

Plasmid construction

Two full-length cDNAs for each target were used; one codon optimized for expression in E. coli and one codon optimized for S. cerevisiae, (Genscript, USA). The cDNAs were PCR amplified with primers listed in table x and inserted into the plasmid along with a Tobacco Etch Virus (TEV) cleavage site, a yeast or E. coli enhanced GFP-tag and a poly-histidine tag. All PCR reactions were performed with AccuPol DNA polymerase (Amplicon, Denmark). Each expression plasmid was generated by homologues recombination in S. cerevisiae using the transformation protocol developed by Gietz and Schiestl[11]. PCR primers contained 5’ overhangs designed for homologous recombination, hereby ensuring correct assembly of PCR fragments and the SalI, HindIII and BamHI digested plasmids. (Supplementary Figure 1). Transformed cells were selected on minimal medium containing 2% glucose with a supplement of 30 mg/liter leucine and 20 mg/liter lysine. Nucleotide sequences were verified by DNA sequencing (Eurofins MWG Operon, Germany).

Live cell bio-imaging

Localization of heterologously expressed GFP-tagged membrane proteins was determined by visualizing GFP fluorescence in whole cells at 1,000 × magnification, using a Nikon Eclipse E600 microscope coupled to an Optronics Magnafire model S99802 camera.

Expression screening and temperature optimization of protein production

E.coli pre-cultures were inoculated from frozen stocks, grown O/N at 30 °C and used to inoculate 50 ml 2 x YT medium to OD450=0.05. At OD450 = 1.0 one half of the culture was transferred to
15 °C and the other to 30°C. After thermo-equilibration protein production was induced with 0.1 mM, 0.5 mM or 1 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG).

Yeast pre-cultures were inoculated from frozen stocks in 5 ml minimal medium containing leucine and lysine at 30°C until saturation. 200µl was subsequently transferred to 5 ml glucose minimal medium containing lysine, but lacking leucine and grown for 24 h at room temperature. Then 500µl pre-culture was transferred to 50 ml glucose minimal medium containing lysine and lacking leucine and grown for further 24 h at room temperature. The pre-culture was used to inoculate 100 ml of expression media (YP medium with 0.5% glucose and 3% glycerol) to OD$_{450}$ = 0.05. At OD$_{450}$ = 1.0, half of the culture was transferred to 15°C and the other half to 30°C. After thermo-equilibration, recombinant protein production was induced by adding 11 ml induction medium (20% galactose dissolved in YP medium containing 3% glycerol and no glucose) to each flask.

For both *E. coli* and *S. cerevisiae* expression, OD$_{450}$ was measured at different time points up to 120 hours post induction. For each time point whole-cell fluorescence was measured in 1 OD$_{450}$ unit in a white micro plate in a spectrofluorometer (Fluoroskan Ascent, Thermo Scientific) with excitation at 485 nm and emission at 520 nm.

**Small scale crude membrane preparation and SDS-PAGE separation for comparing expression in *S. cerevisiae* and *E. coli***

Pairs of 1 L cell cultures were grown as described above. At OD$_{450}$ = 0.6 one the cultures were transferred to 15 °C and the other one to 30 °C, induced with IPTG and galactose for *E. coli* and *S. cerevisiae* respectively. 25 ml cell culture samples were collected at 0, 24, 48, 72, 96, 108 and 120 hours after induction.

Yeast cells from 25 ml cells were harvested at 1,000g for 10 min and the cell pellet was resuspended in 400 µl cold lysis buffer (1M NaCl, 10% glycerol, 25 mM imidazole, 1mM EGTA, 1mM EDTA, 10% glycerol; pH 7.5 containing 1 mM phenylmethysulfonyl fluoride (PMSF) and 1µg/ml of leupeptine, pepstatin and chymostatin (L,P,C)) and transferred to small homogenisator tubes together with glassbeads, followed by homogenization using a Bertin Precellys 24 (Bertin Instruments, France), for 15 seconds, 4 times. Tubes were placed on ice for 2 minutes between each homogenization. The lysed cells were transferred to an Eppendorf tube.
and glass beads were washed several times and liquid transferred to the Eppendorf tube. Cell debris was removed by centrifugation at 3,000 g in a benchtop centrifuge for 10 min at 4 °C. The crude membranes were isolated from the supernatant by ultracentrifugation at 210,000 g at 4 °C for 20 min (Beckmann rotor TLA 100). The crude membrane pellets were homogenized in 200 µl lysis buffer and stored at -80 °C.

*E. coli* cells from 25 ml cells were harvest at 6000 rpm for 20 min at 4 °C in a SLA-3000 rotor, Sorvall). Cell pellets were resuspended in 1ml cold lysis buffer (1M NaCl, 10% glycerol, 25 mM imidazole, 1mM EGTA, 1mM EDTA, 10% glycerol; pH 7.5 containing 1 mM PMSF and 1µg/ml L,P,C) and cells were disrupted by sonication on ice for 15 min using pulsation with 5 seconds intervals and amplitude 50% (Bandelin-Sonopuls 3100). Cell debris was removed by centrifugation at 15,000 g for 10 min at 4 °C (Sorvall SS-34). The crude membranes were isolated from the supernatant by ultracentrifugation at 160,000 g at 4°C for 20 min (Beckmann Optima™TLX ultracentrifuge, S.N. 96U 826 rotor). Crude membranes were resuspended and homogenized in 500 µl cold lysisbuffer with protease inhibitors, and stored at -80 °C.

Crude membranes from each time-point were mixed with 5x SDS-PAGE sample buffer and incubated for 10 min at room temperature prior to loading on a 15 % SDS-gels and all gels were analyzed simultaneously by in-gel fluorescence using a LAS 4000 imager (GE Healthcare, USA).

**Production of membrane proteins**

Yeast cells were pre-cultured and grown at room temperature as described above. 2 L expression medium (YP supplemented with 0.5 % glucose and 3% glycerol) was inoculated to an OD450 of approximately 0.05. When OD450 reached 1 the culture was transferred to 15°C and galactose was added to a final concentration of 2%. Cells were harvested at their expression maximum as established by the expression screens.

Yeast cells from 2L cultures were harvested at 1,600 g for 10 minutes at 4°C. Cells were lysed by glass bead homogenization in ice cold Lysis buffer (1M NaCl, 10% glycerol, 25 mM imidazole, 1mM EGTA, 1mM EDTA, 10% glycerol; pH 7.5 containing 1 mM PMSF and 1µg/ml L,P,C) as described previously (30). The cell lysate was centrifuged at 4°C at 3,000 g for 10 min to remove cell debris. Crude membranes were pelleted from the supernatant by ultra-
centrifugation at 160,000 g at 4°C for 90 min (Sorvall T865 rotor). Crude membranes were re-
suspended and homogenized in lysis buffer containing protease inhibitors and kept at -80°C until use.

To normalize expression yields, GFP fluorescence in 25 µg crude membranes was measured in a
spectrofluorometer (Fluoroskan Ascent, Thermo Scientific) using buffer as a blank. Excitation
was at 485 nm and emission at 520 nm. Fluorescence was converted to pmol protein-TEV-GFP-
His10 from a standard curve generated from purified GFP mixed with yeast membranes as
previously established [12] and normalized to the molecular weight of the protein to calculate the
percentage of total membrane protein content.

**Detergent screens**

Crude membranes were incubated in buffer B (25 mM Tris-HCl, 10 mM imidazole, 0.5 M NaCl,
10% glycerol, pH 7.6) supplemented with protease inhibitors (1 mM PMSF and 1µg/ml L,P,C) at
a protein:detergent ratio (w/w) of 1:3. The screen included the following detergents FC-12, n-
dodecylphosphocholine; FC-13, n-Tridecylphosphocholine; LDAO, Lauryldimethylamine N-
oxide; Cymal-5, 5-cyclohexyl-1-pentyl-β-D-maltoside; DDM, n-dodecyl-β-D-
maltopyranoside;DM, n-decyl-β-Dmaltopyranoside; Detergents were either of Anagrade quality
and purchased from Affymetrix, UK or from Glycon, Germany. Solubilization was performed at
slow rotation at 4°C for 1 hour. Un-solubilized material was removed by ultra-centrifugation at
160,000 g at 4°C for 20 min (Beckmann Optima™TLX ultracentrifuge fitted with an S.N. 96U
826 rotor). Fluorescence was detected in white microplates in a spectro fluorometer (Fluoroskan
Ascent, Thermo Scientific) using buffer as a blank. Excitation was at 485 nm and emission at
520 nm. Solubilization efficiency was estimated as GFP fluorescence in the supernatant after
ultracentrifugation normalized to the total GFP fluorescence in the sample.

**Fluorescence-detection size exclusion chromatography**

Solubilized crude membranes were analyzed by fluorescence-detection size exclusion
chromatography (FSEC) on a Superose 6 Increase 200 10/300 GL column (AqpZ) or and
Superose 12 10/300 GL column (KcsA, NavAb, ClCec1) attached to an ÄKTA Purifier (GE
Healthcare, USA), coupled to a fluorescence detector (Shimadzu Prominence RF-20A) to
visualize the elution profile of the GFP tagged membrane proteins. All experiments were carried out using FSEC buffer (20 mM TRIS-HCl, 0.15 M NaCl, 0.03% DDM pH 7.5).

**Small scale purification**

Crude membranes were solubilized for four hours at 4°C by slow rotation in the detergent found to be most optimal from the detergent screen. Un-solubilized material was removed by ultracentrifugation at 160,000 g at 4°C for 1.5 hour (Sorvall T-865 rotor). Solubilized protein was diluted to a detergent concentration corresponding to 1.5-2 times CMC of the detergent used and incubated overnight in a beaker with Ni-resin (Genscript, USA). The beaker content was transferred to a gravity column and the flow through was collected. The column was washed with Buffers containing 10 mM, 30 mM, 100 mM, 250 mM or 500 mM imidazole. All buffers contained detergent at a concentration corresponding to 1.5 times CMC. Fluorescence in each fraction was quantified using a spectrofluorometer (Fluoroskan Ascent, Thermo Scientific) using buffer as a blank. Excitation was at 485 nm and emission at 520 nm.

GFP fluorescence in 25 µg crude membranes was measured in a microplate reader (Fluoroscan Ascent, Thermo Scientific, USA) at excitation 485 nm and emission 520 nm. Quantifying the relative amount recombinant membrane protein in total yeast membrane was calculated by using the linear relationship between molar amount of GFP and GFP fluorescence generated from purified GFP mixed with yeast membranes as previously established [12].

**Ni-affinity purification**

For purification of the membrane protein-TEV-GFP-His_{10} fusions, the choice of detergent was based on the results from detergent screen and FSEC. Crude membranes were solubilized in the detergent at slow rotation at 4°C for 1 hour. Non-solubilized material was pelleted by ultracentrifugation at 160,000 g at 4°C for 1.5 hour (Sorvall T-865 rotor). Solubilized membranes were diluted to a detergent concentration corresponding to 1.5 times CMC (final concentration of 25 mM Tris-HCl, 10 mM imidazole, 500 mM NaCl, 10% glycerol, pH 7.6) Two purification procedures were used:

1. AqZ. Solubilized protein was loaded onto a HisTrapFF column (GE Healthcare), and washed with a linear gradient of increasing imidazole concentration (from 0 to 500 mM). The eluted
protein was collected in fractions of 0.5 ml and the amount of recombinant protein was determined by GFP fluorescence.

2. ClCec1, NavAb, KcsA. Solubilized protein was poured onto a Econo-column (Biorad, USA), flow-through was collected and the column was washed with different concentrations (10 mM, 30 mM, 100 mM, 250 mM and 500 mM) of imidazole to removed unwanted protein and subsequently elute the recombinant protein. All fractions from the wash steps were collected in fractions and the amount of recombinant protein was determined by GFP fluorescence.

All IMAC buffers (25 mM Tris-HCl, 10 mM imidazole, 500 mM NaCl, 10% glycerol, pH 7.6) contained detergent corresponding to 1.5 times CMC. Fluorescence in each fraction was quantified using a spectrofluorometer (Fluoroskan Ascent, Thermo Scientific) using buffer as a blank. Excitation was at 485 nm and emission at 520 nm.

**Protein quantification**

Protein concentrations in crude membranes were determined by the bicinchoninic acid (BCA) assay according to the Manufacturer’s specifications (Sigma Aldrich, USA).

**SDS-PAGE**

Protein samples were analyzed by SDS-PAGE and western blotting. SDS-gels (10% or 12%) were prepared according to Laemmli, U. K. (1970)[13] or purchased from GenScript, USA and stained with Coomassie-Blue R250. Protein samples were denatured with SDS sample buffer with a concentration of 4% SDS at room temperature for 20 minutes before loading. Imaging of in-gel fluorescence and Coomassie stained gels were done using the Image Quant LAS-4000 (GE Healthcare, USA).

**Electrophysiology**

Single channel currents from purified proteins were recorded using the Orbit Mini with EDR3 software (Nanion Technologies, Munchen, Germany). A multi-electrode-cavity-array (MECA) recording chip, 50-100 µm (Ionera Technologies GmbH in Freiburg Germany) was mounted in the Orbit Mini and recording buffer was applied to the chip. Planar lipid bilayers consisting of 10 mM 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) and 1 mM Cholesterol in nonane
(Avanti Polar Lipids, Alabaster, Alabama) were painted on the recording wells. 0.2 µl KcsA (1 mg/ml in DDM, recoding buffer: 300 mM KCl, 10 mM Hepes, pH 4) or NavAb (1 mg/ml in DDM, recording buffer 100 mM KCl, 100 mM NaCl, 10 mM Hepes, pH 7.2) was applied to the cis-side of the bilayer. Different voltages were applied to facilitate channel insertion. Sampling rate was 1.25 kHz and after recording, a low pass filter, Bessel (8-pole) with a 3 dB cutoff of 200 Hz was applied. Mean current amplitudes were plotted as a function of applied voltages. The single channel conductances were determined as the slope factor of a linear regression to the current-voltage (IV) plot. Clampfit 10.7 (Molecular Devices, Axon Instruments, USA) and Prism 7 (GraphPad Software) were used for data analysis. Data is shown as Mean ± SEM.

Results

The E. coli and S. cerevisiae expression platforms

To compare the capacity of E.coli and S.cerevisiae for bacterial membrane protein production we selected the BL21(DE3) and pET host-vector system [5], which is the most popular for recombinant protein production in E.coli, and the S. cerevisiae PAP1500 expression platform [14]. To simplify construction and manipulation of E.coli expression plasmids we used a pET52b(+) derivative, pPAP10286, that can replicate in S. cerevisiae as well as in E. coli (Figure 1S). This plasmid carries the exact same expression cassette as pET52b(+). The advantage being that E. coli expression plasmids (like yeast plasmids) can be generated fast, cheap and accurate by homologous recombination in yeast (Supplementary Figure 1). This cloning strategy allows for simple optimization of the nucleotide sequences of expression constructs without introducing any “contaminating” sequences, like restriction sites and site specific recombination sequences.

Expression host and temperature affect protein accumulation

To achieve a valid evaluation of the capacity for prokaryotic membrane protein production in E. coli and S. cerevisiae yeast we selected four membrane proteins representing families of different origin. All proteins were expressed as fusions to a TEV-GFP-His10 tag. We used two sets of cDNAs one codon optimized for expression in E.coli and one optimized for production in S. cerevisiae. The TEV cleavage site enables removal of the tag after purification, the poly-His-sequence allows straight forward affinity purification and GFP can be used for quantification and in vivo localization of fusion proteins.
Before evaluating the expression capacity of all proteins in *E. coli*, we identified the optimal combination of IPTG for induction, temperature and *E. coli* host strain for accumulation of AqpZ. We followed accumulation of GFP fluorescence in the four most used *E. coli* strains (BL21(DE3), BL21(DE3) pLysS, C41(DE3) and C43(DE3) at 15°C and 30°C. The data in FIGX show that accumulation in BL21(DE3) was higher at 15°C, while in C41(DE3) and C43(DE3) accumulation was higher at 30°C. The peak accumulation observed in C41(DE3) and C43(DE3) at 30°C after 30 hours and 50 hours, respectively, was comparable to that observed in BL21(DE3) at 15°C after 120 hours. However, the final OD was much higher at 30°C for all strains. The effect of the concentration of IPTG used for induction of AqpZ production depended on the strain and temperature used.

The expression of KcsA, NavAb and ClCec1 in BL21(DE3), BL21(DE3) pLysS and C41(DE3) was screened at both 15°C and 30°C using 0.5 mM IPTG for induction, based on the data from AqpZ. The highest accumulation of ClCec1 was obtained in BL21(DE3) at 30°C and peaked 24 hours after induction. NavAb accumulated the most in BL21(DE3) and C41(DE3) at 30°C already 8 hours after induction, while KcsA accumulation peaked in BL21(DE3) at 30°C after 20 hours (Figure 2).

To ensure that the low expression levels observed for two of the three prokaryotic membrane proteins in *E. coli* was not due to the pET52b based shuttle plasmid, we tested production of the simpler tetracycline repressor protein, TetR (uniprot P04483) from Tn10. It is evident from Supplementary Figure 2 that the low accumulation of prokaryotic membrane proteins is not due to the engineered pET52b plasmid but rather to their membrane embedded nature.

**Temperature dependent accumulation of prokaryotic membrane proteins in *S. cerevisiae* internal membranes**

To determine how temperature affects production in *S. cerevisiae*, we investigated accumulation of the model membrane proteins at 15 °C and 30 °C. It can be seen from Figure 3A-D that at 15 °C accumulation increased over the entire production period and was much higher than at 30 °C. In contrast, the accumulation at 30°C already peaked around 24 hours, followed by a rapid decline for all the membrane proteins.

All the tested prokaryotic membrane proteins accumulate in the plasma membrane in their native hosts. We used live cell bio-imaging to determine if the GFP-tagged prokaryotic membrane
proteins maintained their membrane localization in the new host. It is seen from Figure 3E that all proteins accumulated in distinct intracellular compartments and not in the plasma membrane.

**Prokaryotic membrane proteins accumulate to a high density in yeast membranes**

All the protein targets investigated in the present paper have previously been purified and crystallized from an *E. coli* expression system. We therefore set out to test if our *S. cerevisiae* expression platform could produce these membrane proteins in large yields, and subsequently if they could be purified in a functional form.

To achieve optimal protein production yeast cells were induced at 15°C and harvested at the established expression optimum for each target protein as visualized in Figure 3A-D. The data in Table 1 show that all the target proteins accumulated to a high density in yeast membranes.

Because the molecular weight of the produced proteins differs significantly, we also calculated the molar amounts accumulating per mg total membrane protein content. This showed that the amount of protein accumulating did not differ much among the different targets. From the estimated amount of protein in the membranes, we calculated to theoretical amount of protein the yeast cells would produce, seen in Table 1. This would give rise to the potential to purify between 11 mg and 42 mg of these target proteins per liter from shake flask grown cells with a rather low cell density.

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (kDa)</th>
<th>pmol/mg</th>
<th>% of total membrane protein</th>
<th>mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>AqpZ</td>
<td>23,703</td>
<td>1135.07</td>
<td>2.7</td>
<td>41</td>
</tr>
<tr>
<td>ClCec1</td>
<td>50,000</td>
<td>1989.05</td>
<td>2.8</td>
<td>42</td>
</tr>
<tr>
<td>NavAb</td>
<td>31,000</td>
<td>901.25</td>
<td>0.8</td>
<td>12</td>
</tr>
<tr>
<td>KcsA</td>
<td>18,000</td>
<td>1302</td>
<td>0.7</td>
<td>11</td>
</tr>
</tbody>
</table>

**Table 1.** The expression level of the fusion proteins was correlated to the total protein content in the yeast membrane from the GFP fluorescence in the crude membranes as described in material and methods. The pmol/mg was then converted to % from the molecular weight (MW) of the protein. The theoretical amount of protein (mg/L) produced: 10 g cells from 1L yeast culture, gives 1.5 g isolated crude membranes; from the % of the proteins then results in the values listed.
Prokaryotic membrane proteins show distinct solubilization profiles

We used a simple screen to compare the ability of six different detergents to solubilize the prokaryotic membrane proteins produced in yeast. The results in Figure 4 show that each target protein could be solubilized in all the detergents tested. AqpZ showed the lowest solubilization efficiency of around 20% in all detergents. ClCec1 generally solubilized well and the highest efficiencies were observed in FC-12, DDM and CY-5 and amounted to 80-100%. KcsA showed between 40-75% solubilization, with the highest observed in DDM. At last, between 80 and 100% of NavAb was solubilized in DM, LDAO and DDM, 50% in CY-5 and 30% in FC-12 and FC-13.

Homogeneity of solubilized prokaryotic membrane proteins are detergent independent

To maintain activity of the prokaryotic membrane proteins during purification, it is essential to identify a detergent that results in a homogeneous- and non-aggregated preparation. We used fluorescent size exclusion chromatography (FSEC) to assess the quality of each detergent solubilized protein. The results in Figure 4 show the FSEC profiles from each protein in all six detergents, with colors corresponding to those used in the detergent screen. In general the protein elution peaks are very narrow and symmetrical, with little elution in the void volume, indicating a high protein quality. Only NavAb showed a double peak, revealing two populations of the protein in LDAO, DM and DDM and to some degree in CY-5, FC-12 and FC-13.

IMAC yields pure fusion proteins

Since the detergent screen and the FSEC results (Figure 4) revealed that several detergents looked promising for purification, AqpZ was solubilized in LDAO and ClCec1 in DM, while NavAb and KcsA were both purified in DDM.

AqpZ, ClCec1, NavAb and KcsA were purified by immobilized metal affinity chromatography (IMAC). Figure 5 shows the IMAC elution profile for each protein. The top fraction containing the most fluorescence was separated by SDS-PAGE, and the in-gel fluorescence and Coomassie stained gel was used to visualize the purity of the protein sample.

AqpZ eluted as a single broad peak starting at 50 mM imidazole (Figure 5A). SDS-PAGE separation of purified AqpZ resulted in a highly pure protein band, with only minor bands of higher molecular weight visible both by in-gel fluorescence and Coomassie staining, indicating that AqpZ tetramers have not been fully dissociated (Figure 5A).
ClCec1 eluted as one very narrow and symmetrical peak at 500 mM imidazole (Figure 5B). IMAC resulted in a large amount of pure protein, with only minor impurities visible in the Coomassie stained SDS-PAGE gel seen in Figure 5B.

KcsA mainly eluted as a single peak with a shoulder at 100 mM imidazole (Figure 5C). The elution profile also revealed a minor amount of protein eluting at 50 mM imidazole and a second small elution peak at 250 mM imidazole. SDS-PAGE of the top fraction revealed one band corresponding to KcsA, but a substantial amount of the fluorescence did not enter the gel. An impurity of higher molecular weight was only seen in the Coomassie stained SDS-PAGE gel (Figure 5C).

The elution profile of NavAb resembled that of KcsA, however the SDS gel revealed more impurities seen only by Coomassie stain; lane 1, Figure 5D. The SDS gel also showed several bands of higher molecular weight visible by Coomassie and in-gel fluorescence, hence not all of NavAb fully dissociate by the SDS denaturation (Figure 5D).

The molecular weights of the purified fusion proteins, as determined by SDS-PAGE, corresponded well to the molecular size of the monomeric proteins (Aqpz, 23 kDa; ClCec1, 50 kDa; KcsA, 18 kDa; NavAb, 30 kDa) when taking into account that correctly folded GFP only contributes 10-15 kDa to the molecular weight of the fusion proteins (Geertsma 2008 Quality).

**The prokaryotic ion channels produced in yeast are functional**

The above described data confirm that prokaryotic membrane proteins can be produced in yeast and purified to homogeneity from yeast membranes, but do not provide information on the functionality of the proteins. To investigate the activity of the recombinant KscA and NavAb, we purified these channels after solubilization in DDM and subsequently reconstituted the channel proteins into lipid bilayers.

Before doing the electrophysiology we used FSEC analysis of IMAC purified KcsA and NavAb solubilized in DDM to show that the detergent solubilized proteins had maintained stability during the purification (Supplementary figure 3).

Channel activity was recorded at a series of different voltages (Supplementary Figure 4A), and a representative KcsA recording at -40 mV is shown in Figure 6A. The reconstituted KcsA channels were functional and the representative recording display channel openings and closings...
of at least three single channels (O1-O3) The mean single KscA channel current amplitude is plotted as a function of voltage (I-V curve) in Figure 6B, and revealed a slope conductance of 77.3 ± 5.6 pS.

The reconstituted NavAb channels were also functional and representative single-channel current recordings between at ± 100 mV with show fast transitions between the open and the closed state (Figure 6C). The recordings also revealed that the activity was voltage dependent, with a larger open probability at positive voltages (Supplementary Figure 4B), suggesting that channels apparently show a preference for opening at positive membrane potentials. It should be noted, however, that due to random insertion into the bilayer, the orientation of the channels cannot be known and the polarity of channel insertion is speculative. The I-V curve in Figure 6D revealed a slope conductance of 57.6 ± 4.0 pS.

N-terminal tagging improves accumulation of AqpZ in E. coli and S. cerevisiae

To test if the position of the tag affects the expression level and quality of recombinant prokaryotic membrane proteins produced in E.coli and S.cerevisiae we used the previously described expression plasmids (Supplementary Figure 1) to express AqpZ N-terminally tagged with His10-GFP-TEV. Data in Figure 7 show accumulation of His10-GFP-TEV-AqpZ in four BL21 derived E. coli strains. It can be seen that BL21(DE3), C41(DE3) and C43(C43) were able to express the N-terminally tagged AqpZ much better at 15°C than at 30°C. By comparing the data in Figure 7 to Figure 1, it can be seen that moving the GFP-His tag from the C-terminal of AqpZ to the N-terminal, greatly increased accumulation of the protein. The expression curves also revealed that expression levels were generally not affected significantly by the IPTG concentration used for induction, except for the expression in BL21(DE3), where the expression was higher after induction with 0.1 mM IPTG .

In accordance with the increased expression in E. coli, Figure 8 shows that N-terminally tagged AqpZ also accumulated to a higher density in S. cerevisiae, than the C-terminally tagged version. It also showed that the accumulation was considerably higher at 15°C than at 30°C.

Whole cell fluorescence does not necessarily reflect the specific accumulation of the fusion proteins in the membranes but may potentially result from GFP being hydrolytically removed from the membrane protein. We therefore used in-gel fluorescence of membranes isolated from E. coli and S. cerevisiae expressing either the N- or C-terminally tagged AqpZ to visualize the
effect of tag position, temperature, expression organism and induction time on AqpZ accumulation. To compare the accumulation of GFP tagged AqpZ all gels were imaged simultaneously (Figure 8C). In agreement with the whole cell fluorescent measurements, it can be seen that the N-terminally tagged AqpZ had a higher accumulation in the membranes at 15°C, for both organisms. The expression levels visualized in the gels in Figure 8C also revealed that S. cerevisiae was superior for production of AqpZ. It was also evident that only the full length proteins accumulate in the membranes, seen from the single band in the SDS-gel.

**N-terminal position of the tag improves solubilization in E. coli and S. cerevisiae**

We performed a detergent screen and FSEC analysis on the His10-GFP-TEV-AqpZ fusion, to investigate the influence of the GFP-His-tag position on solubilization and quality of the recombinant AqpZ produced. The screens displayed in Figure 9A,B, revealed a significant increase in solubilization efficiency for the N-terminally tagged AqpZ compared to the C-terminally tagged version (Figure 4), as the N-terminally tagged AqpZ was solubilized between 80-100%, while only 20% of the latter was solubilized, hence a four to fivefold improvement. The FSEC profiles revealed a wider elution peak in LDAO, but all profiles were narrow, symmetrical and in fact superimposable.

Lastly we purified the LDAO solubilized N-terminally tagged AqpZ produced in yeast by IMAC, using the same procedure as for the C-terminal construct. The chromatogram in Figure 9C revealed that the N-terminal AqpZ construct eluted in a single narrow peak at around 60 mM imidazole, resulting in a pure protein sample as revealed by in-gel fluorescence and Coomassie staining of SDS-PAGE separation of the purified protein (Figure 9C).

**Discussion**

Microbial expression systems play a vital role in producing proteins for biotechnological- and medical applications in addition to basic protein science. Membrane proteins are mainly of interest to the biomedical industry and to basic science as targets for medical treatments and structure-function studies, respectively. The prokaryotic membrane protein structures deposited in the Protein Data Bank (PDB) almost exclusively originate from proteins produced in *E. coli*, reflecting the general belief that *E. coli* is a superior host for production of prokaryotic membrane proteins. In contrast *S. cerevisiae* has been used as a workhorse for production of eukaryotic membrane proteins for structural biology [3, 15].
In the present paper we selected four prokaryotic membrane proteins to systematically investigate if *E. coli* is in fact a superior host for production of prokaryotic membrane proteins or if *S.cerevisiae* possesses so far unidentified potential. Our results demonstrate that *S. cerevisiae* was a more robust expression host as the four test proteins accumulated to almost the same high level. We confirmed that the proteins expressed in yeast could be purified to homogeneity and we were able to confirm the functionality of the purified NavAb and KcsA by single channel recordings in lipid bilayers.

Using a high yield expression host vector system is crucial. A strong inducible promoter will increase expression levels, by ensuring that the transcription is not the rate-limiting step for protein production [16]. For this reason the galactose inducible promoter Gal1p was used for expression in yeast. For expression in *E. coli*, an IPTG inducible promoter was used, together with the T7 expression system. One of the advantages of this system is that IPTG is not metabolized, leading to efficient transcription throughout the production phase. This advantage is also present in our *S. cerevisiae* platform as the inducer, galactose, cannot be metabolized in this strain.

As revealed by our results a drawback from expression in *E. coli* is that the accumulation of membrane proteins differed significantly among the tested strains. AqpZ showed the highest accumulation in C41 and C43, while BL21 was superior at expressing ClCec1, KcsA and NavAb. None of the proteins expressed well in BL21(DE3) pLysS. This means that it may be necessary to screen a number of strains to achieve a satisfactory production level. As the optimal temperature for accumulation was observed to be membrane protein dependent it will also be necessary to investigate expression at various temperatures. This obviously does not make the *E.coli* platform very straight forward and robust.

In accordance with our previous experience with production of eukaryotic membrane proteins in our PAP1500 yeast platform lowering the temperature prior to protein induction from 30°C to 15°C improved accumulation of all tested bacterial membrane proteins. So in contrast to *E.coli* our results demonstrate that a using a single yeast strain and expression at 15°C would ensure high level accumulation of functional membrane protein. This is very encouraging as one might initially doubt that the prokaryotic translocation signals would be recognized in yeast and the bacterial proteins would be inserted correctly into the altered lipid environment of a eukaryotic cell. Prokaryotic membranes do not contain steroids like cholesterol and have a different
phospholipid composition than yeast. However, our results prove such concerns wrong as all tested proteins accumulated to a higher density in the yeast membranes than in the E. coli membranes, and those assayed for activity proved to show the same electrophysiological activity as those purified from E. coli.

E. coli have mainly been used, for historical reasons; when it was discovered it was found to express recombinant proteins well, to be easy to manipulate and cultivate, hence laboratories continued to use this organism as their standard expression host for bacterial membrane protein [2]. Another reason given for choosing E. coli as the expression host of bacterial membrane proteins is offers low cost culturing. However, the price difference between E. coli and S. cerevisiae media is not big and both expression systems have the advantage of being applicable for large scale cultivation in bioreactors.

The present paper therefore only focused on the production of prokaryotic membrane proteins in the membranes of the host organism. Expressing membrane proteins in yeast has the advantage of avoiding expressing the protein in inclusion bodies, hereby circumventing the requirement for membrane protein refolding. Membrane proteins can be expressed in very high quantities in inclusion bodies, however it is often inefficient or impossible to refold them properly due to the complicated folding and hydrophobic nature of this class of proteins [1, 4]. Secondly, yeast possesses an eukaryotic machinery for membrane protein biogenesis that may support correct folding of membrane proteins better than the simpler prokaryotic secretory pathway of E. coli.

Expression in yeast might potentially result in non-native posttranslational modifications including N- and O-glycosylation. We did not look for such modifications but the very sharp bands we observed after purification do not indicate heavy glycosylation (Figure 5 and 9). We observed that proteins accumulated in the membrane when expressed in yeast (Figure 3 and 8). This shows that these proteins are recognized as proteins destined for the secretory pathway, also in yeast.

We used GFP to quantify the solubilization efficiency of six detergents. We were able to screen in a volume of only 500 µl, meaning very small amounts of the crude membranes containing the precious membrane protein had to be analyzed. This is possible because GFP measurements are so sensitive, that as little as a few microliters are needed to measure a fluorescent signal. GFP was also utilized for GFP for FSEC, where a 500 µl sample is sufficient to obtain a good profile.
This allowed for comparison of protein solubilization yield and behavior in the different detergents in terms of monodispersity and aggregation.

Using a single step of IMAC was shown to purify all four target proteins to a high degree of purity. Binding the proteins to the resin via a His10-tag enabled removal of unwanted proteins before eluting the proteins from the columns. ClCec1 eluted at 500 mM imidazole whereas KcsA and NavAb eluted at 100 mM imidazole and AqpZ at 50 mM imidazole. The strong binding of ClCec1 enable thorough washing of the protein with higher imidazole concentration before eluting this protein, however it also means that the protein will elute in a higher salt concentration, which might impact protein stability. The other proteins elute at much lower imidazole concentrations which on the other hand limits the use of washing with imidazole to remove unspecific bound protein. SDS-PAGE separation of the purified KcsA and NavAb revealed impurities however the FSEC profiles of the samples were monodisperse after purification in DDM.

The single step affinity purification makes it simple and easy to purify the proteins. GFP makes it even easier to follow the purification and removes the need of running SDS-PAGE and western blots to confirm which fraction the protein elutes in. Instead SDS-Page was used after purification to evaluate the purity of the protein samples. By using in-gel fluorescence and Coomassie stain it is possible to determine the specific protein band of interest without the use of western blotting, saving both time and money.

While we did not utilize this feature, the TEV cleavage site gives the option of removing the tags, which is often needed for crystallization trials (REF). The tags can be removed by cleaving with TEV protease. The cleaved tag and protein can then be separated by reverse-IMAC where the tag will bind to the resin and the protein will elute straight in the flow-through, which is why the tags have to be after purification. The TEV protease is easy to produce and purify in the lab, which makes this an ideal and cost-effective solution for removal of tags.

We were able to show that expression in yeast and the purification procedure yielded active KcsA and NavAb channels. Recordings of KcsA currents showed a single channel conductance of $77.3\pm 5.6$ pS in accordance to previously reported native channel chord conductance of 56 pS[17]. For NaNavAb we found a single channel conductance of $57.6 \pm 4.0$ pS, which is in agreement with previous reports of related channel NavAb1p with a conductance of 37 pS[18].
We observed a higher open probability at positive voltages. This is in agreement with a molecular dynamics simulation of the NavAb pore, which saw more cations in the selectivity filter at positive voltages, and less at a negatively applied voltage [19]. This is to our knowledge the first single channel recordings of NavAb.

Even though AqpZ expressed in sufficient amounts and we were able to purify it with the C-terminal tags, we wanted to explore if we could improve expression levels by changing the tag localization. Not only did we see a great increase in the expression level, but also a great improvement of the solubilization degree in all six detergents. This opens the idea that target expression levels should be tested with both C- and N-terminal tags, to find the optimal fusion construct. The structure of KcsA has both been obtained from protein expressed with C or N-terminal His tag [20, 21], showing that both tag positions is able to maintain the native fold and express in high yields. We find that it is worth testing which tag position yields the highest expression level for a given protein of interest.

We showed that in-gel fluorescence of whole cell lysates was an easy method to visualize the difference in expression between the N and C terminally tagged (-TEV-GFP-His_{10}) AqpZ over time. In this way we also ensure that we measure the fluorescence localized specific in the membranes, and not from inclusion bodies of GFP that have been hydrolytically removed from the membrane protein.

**Future perspectives**

Lipid composition of the membrane can affect the membrane protein production, and might be one of the differences between expressing prokaryotic membrane proteins in yeast and *E. coli*. We could further engineer the host organism to suit a given type of membrane protein, to improve expression levels and quality of the protein, however as a starting point yeast showed to be a good production organism for prokaryotic membrane proteins.

**Conclusion**

We show that expression of the selected microbial membrane proteins was possible both in *E. coli* and yeast, with a higher expression yield in yeast. Using a GFP tag enables a fast detection of both expression yields, identification of the most optimal detergent for purification, and to visually follow the protein during purification. We find that yeast is superior to bacteria for high yield production of bacterial proteins.
Acknowledgements

The authors thank David Sorensen for excellent technical assistance. The present work was supported by the Innovation Foundation through the MEMENTO grant.

Author contributions

SSP and PAP cloned the expression constructs. SSP, MF and LK performed protein expression and purification. PAP performed expression screens. DAK and KC performed activity measurements. SSP and PAP designed the project. SSP wrote the initial version of the manuscript and the final version in association with PAP. All authors contributed with comments on the final manuscript.

References


**Figures**
Figure 1. Effect of IPTG concentration and temperature for expression of AqpZ-TEV-GFP-His$_{10}$ in E. coli strains.

E. coli strains A, BL21(DE3); B, BL21(DE3) pLysS; C, C41(DE3); D, C43(DE3), were grown O/N at 30 °C until OD$_{450}$=1, and transferred to either 15 °C (Left) or 30 °C (Right) prior to induction with 0.1 mM, 0.5 mM or 1 mM IPTG. Protein accumulation was measured by GFP fluorescence in 1 OD$_{450}$ unit at each time-point after induction; fluorescence measurements shown as full line and OD measurements as dashed line. Color-codes are indicated in the bottom.
Figure 2. Effect of temperature for expression of ClCec1, KcsA and NabAb as TEV-GFP-His<sub>10</sub> fusions in E. coli strains.

E. coli strains BL21(DE3), BL21(DE3) pLysS and C41(DE3), were grown O/N at 30 °C until OD<sub>450</sub>=1, and transferred to either 15 °C (blue) or 30 °C (Red) prior to induction with 0.5 mM IPTG. Protein accumulation was measured by GFP fluorescence in 1 OD<sub>450</sub> at each time-point after induction; fluorescence measurements shown as full line and OD measurements as dashed line.
Figure 3. Temperature dependent expression of AqpZ, CiCec1, KcsA and NabAb as TEV-GFP-His\textsubscript{10} fusions in \textit{S. Cerevisiae}. Yeast cultures were grown at room temperature until OD\textsubscript{450}=1, then half of the culture was transferred to 15°C and the other half to 30°C. After thermo-equilibration, recombinant protein production was induced with 2% galactose. A, Protein accumulation was measured by GFP fluorescence in 1 OD\textsubscript{450} at each time-point after induction at 15°C (Blue) and 30°C (Red). B, Live cell bio-imaging of yeast cells expressing the protein TEV-GFP-His\textsubscript{10} fusions. Pairwise GFP and differential interference contrast (DIC) images are shown.
Figure 4. High-throughput detergent screen of crude membranes showing solubilization efficiency and monodispersity. Crude membranes with A, AqpZ-TEV-GFP-His\textsubscript{10}; B, ClCec1-TEV-GFP-His\textsubscript{10}; C, KcsA-TEV-GFP-His\textsubscript{10}; D, NavAb-TEV-GFP-His\textsubscript{10}, were solubilized for 1 hour in FC-12, n-dodecylphosphocholine; FC-13, n-Tridecylphosphocholine; LDAO, Lauryldimethylamine N-oxide; Cymal-5, 5-cyclohexyl-1-pentyl-β-D-maltoside; DDM, n-dodecyl-β-D-maltopyranoside; DM, n-decyl-β-D-maltopyranoside, in a protein:detergent ratio of 1:3. A. normalized solubilization efficiency estimated as fluorescence in the supernatant after ultracentrifugation (left, diagrams). B. Detergent solubilized crude membranes separated by size
exclusion chromatography and GFP fluorescence was monitored during the elution (right, chromatogram). The colors of the chromatograms (right) corresponds to the colors of each detergent in the bar diagram (left).

Figure 5. Affinity purification of the AqpZ, ClCec1, KcsA and NabAb TEV-GFP-His₁₀ fusions. Each fusion protein was solubilized in detergent for 4h at 4°C; A, LDAO; B, DM; C, DDM; D, DDM, and un-solubilized material was then removed by ultra-centrifugation. Each protein was purified by immobilized metal ion affinity chromatography (IMAC) and the protein elution was monitored by GFP fluorescence in the collected fractions. The dashed line indicate the imidazole gradient used. Left inserts show the SDS-PAGE gels with the top fraction from each elution, visualized by Coomassie-stain (1) and in-gel fluorescence(2).
**Figure 6. Single-channel recordings of KcsA and NavAb reconstituted in planar lipid bilayers.**

KcsA and NavAb were solubilized and purified by IMAC in DDM and reconstituted into planar lipid bilayers on a 50 or 100 µm MECA 4 recording chip in the Orbit Mini system (Nanion Technologies). A. Representative single KcsA channel current recordings at -40 mV. At least three individual channels have been incorporated into the planar lipid bilayer. (C) Indicates the closed state and (O1-3) indicates the open states. B. The Current-Voltage relationship (I-V curve) for KcsA with a linear regression fitted to the data, revealing a slope conductance of 77.3 ± 5.6 pS. C. Representative single NavAb channel current recordings at + 100 mV. D. I-V curve for NavAb with a linear regression fitted to the data, revealing a slope conductance of 57.6 ± 4.0 pS. Results are shown as mean current amplitudes ± SEM, where each data point was based on at least three single channel openings (n = 3-1300 events, from 3 independent experiments).
Figure 7. Effect of IPTG concentration and temperature for expression of His10-GFP-TEV-AqpZ in *E. coli* strains. *E. coli* strains A, BL21(DE3); B, BL21(DE3) pLysS; C, C41(DE3); D, C43(DE3), were grown O/N at 30 °C until OD$_{450}$=1, and transferred to either 15 °C (Left) or 30 °C (Right) prior to induction with 0.1 mM, 0.5 mM or 1 mM IPTG. Protein accumulation was measured by GFP fluorescence in 1 OD$_{450}$ unit at each time-point after induction; fluorescence measurements shown as full line and OD measurements as dashed line.
Figure 8. Temperature dependent expression of His$_{10}$-GFP-TEV-AqpZ in *S. cerevisiae*.

Yeast cultures were grown at room temperature until OD$_{450}$=1, then half of the culture was transferred to 15°C and the other half to 30°C. After thermo-equilibration, recombinant protein production was induced with 2% galactose. A, Protein accumulation was measured by GFP fluorescence in 1 OD$_{450}$ at each time-point after induction at 15°C (Blue) and 30°C (Red). B, Live cell bio-imaging of yeast cells expressing the His$_{10}$-GFP-TEV-AqpZ, GFP and differential interference contrast (DIC) images. C, A comparison of fluorescence in cell membranes of *S. cerevisiae* and *E. coli*, grown as described in material and methods, at 15 °C and 30 °C. 50 ml culture samples were taken out at each indicated time-point, and crude membranes were isolated and separated by SDS-PAGE. In-gel fluorescence, of all four SDS-gels, was imaged simultaneously using the Image Quant LAS-4000 (GE Healthcare, USA).
Figure 9. Detergent screen and affinity purification of N-term AqpZ. Crude membranes His$_{10}$-GFP-TEV-AqpZ were solubilized for 1 hour in FC12, n-dodecylphosphocholine; FC13, n-Tridecylphosphocholine; LDAO, Lauryldimethylamine N-oxide; Cymal-5, 5-cyclohexyl-1-pentyl-β-D-maltoside; DDM, n-dodecyl-β-D-maltopyranoside;DM, n-decyl-β-D-maltopyranoside, in a protein:detergent ratio of 1:3. A, Solubilization efficiency was estimated as fluorescence in the supernatant after ultracentrifugation normalized to the total fluorescence in the crude membranes for each detergent. B, The detergent solubilized crude membranes were separated by size exclusion chromatography and GFP fluorescence was monitored during the elution. The colors of the chromatograms (B) corresponds to the colors of each detergent in the bar diagram (A). C, His$_{10}$-GFP-TEV-AqpZ was solubilized in LDAO for 4h at 4°C and unsolubilized material was removed by ultra-centrifugation, and purified by immobilized metal ion affinity chromatography (IMAC). The protein elution was monitored by GFP fluorescence in the collected elution fractions. The dashed line indicate the imidazole gradient used. Left inserts show the SDS-PAGE gel with the top elution fraction, visualized by Coomassie-stain (1) and in-gel fluorescence (2).
**Supplementary Figures**

**Supplementary figure 1. Structure and assembly of expression plasmids.** A, Map of the yeast expression plasmid pPAP2259 containing CG-P, hybrid promotor with the GAL10 upstream activation sequence in the 5’ non-translated leader of the cytochrome-1 gene; 2µ, yeast origin of replication; leu2-d, β-isopropyldenhydrate dehydrogenase gene with truncated promotor sequence resulting in poor expression of this gene; bla, β-lactamase gene; pMB1, origin of replication; URA3, orothidine-5-phosphate decarboxylase gene. B, Map of the E. coli expression plasmid pPAP10286 containing T7-P, T7 RNA polymerase promotor; lacIg, lac repressor gene; TetR, tetracyclin repressor protein A gene; pMB1, origin of replication; URA3, orothidine-5-phosphate decarboxylase gene. C, Illustration of gene insertion into expression plasmids via homologous recombination. PCR primers designed with sequence overlap at sites for homologous recombination, indicated by black crosses. The plasmids are constructed by co-transfecting PCR products with SalI, HindIII and BamHI linearized plasmid, either pPAP2259 or pPAP10286. FW primer sequence aligns in the reading direction, RV primers are reverse complementary. Blue primer contains the 35 bp homologous sequence to the plasmid, orange primer encode the Tobacco Etch Virus (TEV) cleavage site. Yellow indicate target gene sequence, Green indicate the green fluorescence protein (GFP)-polyhistidine fusion tag.
Supplementary figure 2. Protein expression capability test of the constructed plasmid in BL21. The *E. coli* strain BL21(DE3) was grown O/N at 30 °C until OD$_{450}$=1, and transferred to 15 °C prior to induction with 0.1 mM, 0.5 mM or 1 mM IPTG. Protein accumulation was measured by GFP fluorescence in 1 OD$_{450}$ unit at each time-point after induction; fluorescence measurements shown as full line and OD measurements as dashed line.

**Figure 3. FSEC of KcsA and NavAb after IMAC purification.** The monodispersity of KcsA and NavAb after IMAC purification in DDM was analyzed by FSEC. A small sample was separated on a Superose 12 10/300 GL column and the GFP fluorescence of the elution was monitored. The chromatogram was plotted as normalized fluorescence over the elution volume.
Figure 4. Single-channel current recordings of KcsA and NavAb reconstituted in planar lipid bilayers. NavAb and KcsA was reconstituted in planar bilayers on a 50 µm MECA 4 recording chip in the Orbit Mini system (Nanion Technologies). Representative recordings at a range of voltages show opening and closing of channels. The blue line indicates the baseline.
CHAPTER 3 – Manuscript II

Synopsis

With the goal of creating constitutively open ion channels for developing biomimetic membranes, removal of gating mechanisms in the ion channels would be needed. This in turn leads to creating many different mutations, to be able to find the optimal engineered ion channel. Therefore we needed an efficient tool to create these mutations.

The manuscript focuses on the many potentials of generating expression plasmids and site specific mutations by homologues recombination in *S. cerevisiae*. This is the basis for how the cloning work was performed in this thesis, in an easy and cost-effective manner and with limited hands-on time.
Homologous recombination in yeast: A tailor made tool for manipulating expression plasmids

Sarah Spruce Preisler and Per Amstrup Pedersen *

Department of Biology, August Krogh Building, University of Copenhagen, Universitetsparken 13, 2100 Copenhagen OE, Denmark

*Corresponding author: PAPedersen@bio.ku.dk

Abstract

Background: Fast and efficient generation and manipulation of recombinant expression plasmids is vital to the Biotechnological and Biomedical industries as well as to academic protein science. Proteins that are difficult to express often require optimization of vector constructs. Efficient, fast and flexible plasmid construction is therefore of great importance for achieving recombinant expression of such proteins. The present paper shows how cloning by homologous recombination in S. cerevisiae can assist with specific features, required to optimize the structure of expression plasmids. The method is extremely simple and offers complete flexibility for constructing any desirable vector, with a minimum of in vitro manipulations, and without having to buy expensive reagents.

Results: We demonstrate that homologous recombination in S. cerevisiae is a powerful tool for precise and fast assembly of expression plasmids without introducing any contaminating nucleotide sequences. Distances and nucleotide sequences between important elements required for gene expression can easily be maintained or manipulated. We also demonstrate that homologous recombination is a convenient tool to assemble several DNA fragments in the intended order, simply by including a 25-35 bp long recombination sequences in the 5’ extensions of PCR primers used for target amplification. Site directed mutagenesis is easily performed by including the desired nucleotide alterations in the PCR primers. Furthermore we show that homologous recombination is a very attractive tool for generating plasmids carrying two different expression cassettes (for co-expression studies). Lastly, we reveal how to generate an E. coli pET based expression plasmid that uses homologous recombination in yeast to generate a clean E. coli expression plasmid.

Conclusions: Cloning by homologous recombination in S. cerevisiae is an all-round technology that is completely sequence independent and does not require expensive enzymes. It can be
applied to manipulate any nucleotide sequence. Cloning by homologous recombination therefore contributes the flexibility that is particularly important when it comes to construction and optimization of clean expression plasmids.

**Keywords**
Recombinant proteins, recombinant plasmids, DNA manipulation, genetic engineering, homologous recombination, yeast.

**Background**
Recombinant protein production is central to modern biotechnology that constitutes a growing market predicted to exceed 775 billion $ in 2024 [1]. Optimizing product yields is therefore of utmost importance for biotech companies. Furthermore, modern protein chemistry is crucially dependent on access to recombinant proteins, as only a very restricted number of proteins can be purified from native sources. Therefore, structure and function analysis in basic research and generation of protein products in the biotech and biomedical industries rely on the ability to produce and purify engineered proteins. Obviously such proteins can only be obtained by recombinant gene expression. However, in many cases it is far from being trivial to obtain large amounts of prime quality recombinant protein and a successful outcome is an iterative process often requiring creation and screening of a number of plasmid constructs. Access to methods for simple, cheap and reliable creation of recombinant plasmids is consequently an essential part of biotechnology and basic protein science.

The ability to generate recombinant plasmids was introduced in the early seventies after the discovery of restriction enzymes and DNA ligase [2]. Even though technology has improved substantially, generation of recombinant plasmids can still be a tedious and time consuming process, in particular when stoichiometric co-expression of two or more proteins is required. The advent of PCR seeded the invention of new technologies that improved the speed and effectiveness of generating recombinant plasmids. Gateway cloning [3] exploits phage lambda integration and excision from the *E.coli* genome, USER cloning [4] relies on incorporation of a single deoxy-uracil in the PCR primers used to amplify the target DNA, while LIC [5] uses the 3’-5’ exonuclease activity of the T4 DNA polymerase to generate single stranded extensions, that can base pair with complementary single stranded regions in the cloning vector. These cloning technologies allow generation of plasmids in a high throughput manner but at the cost of introducing additional nucleotides and limitations in the flexibility for DNA inserts.
Cloning requirements are however higher for creation of expression plasmids, as the structure of the final plasmid construct often has a huge impact on the amount of produced recombinant protein [6]. The ideal high throughput cloning technology for generating expression plasmids should be cheap and flexible enough to include all desired features in vector design and possess the ability to introduce randomization of particular sequences in the final construct. Important features include; the distance and nucleotide sequence between promoter elements and the transcriptional start site, the distance and nucleotide sequence between the mRNA sequence recognized by the ribosome (Shine-Dalgarno sequence in bacteria and Kozak sequence in eukaryotes) and the translation initiation codon. Evolution has optimized all these traits to form the structure of native expression cassettes [7], [8] that assures the optimal expression of each gene. To take advantage of the results of evolution we initially want to preserve the natural expression context to optimize the structure of expression plasmids by removing the natural gene, by surgical precision, and substitute it with a codon optimized cDNA encoding the protein of interest. This strategy maintains the entire native expression cassette; the promoter, the transcriptional start site, the ribosome recognition sequence and its distance to the translation initiation codon and the nucleotide sequences separating the individual elements.

Even small cloning scars like unwanted nucleotide sequences introduced by restriction enzyme-ligase mediated cloning procedures, have been shown to reduce recombinant protein production tremendously [6]. This may reflect secondary structures introduced into the mRNA, which will subsequently reduce the translation initiation rate.

The present paper demonstrates that in vivo homologous recombination in *Saccharomyces cerevisiae* is a simple and cost effective tool particularly for generation of expression plasmids, and minimizes the need for in vitro enzymatic reactions for plasmid assembly. This cloning method does not leave cloning scars and we show that it can be used to; generate expression plasmids for any host, generate single or multiple nucleotide substitutions, combine genes to encode fusion-proteins, and generate a single plasmid with multiple expression cassettes to attempt stoichiometric co-expression.

The method described in the present paper has a strict focus on exposing the many potentials of homologous recombination for generation of expression plasmids and site specific mutations. As such it is an expansion of the yeast gap-repair/plasmid shuffle protocol described previously [9].
Results

Cloning by homologous recombination in *S. cerevisiae*, overall strategy

Our approach to create expression plasmids is illustrated in Figure 1. In short, the protocol simply involves co-transformation of one or more PCR fragments with a linearized plasmid into *S. cerevisiae* and plating on selective media. Growth on selective medium requires that the plasmid can replicate, which reflects that the linearized plasmid has been converted into a covalently closed, circular molecule. Homologous recombination between pairwise identical nucleotide sequences joins PCR products and linearized plasmid in a predefined way to create a closed circular plasmid with the designed structure. Being an essential biological process, homologous recombination has evolved not to introduce or remove any nucleotides during the recombination event [10].

Generation of ion channel GFP fusions

To demonstrate the robustness of the homologous recombination cloning strategy outlined in Figure 2, we selected a number of ion transport proteins with different molecular weights, ion selectivity and origin; Three from bacteria, the KcsA K⁺ channel from *Streptomyces lividans*, the NavAb Na⁺ channel from *Arcobacter butzleri*, the ClC-ec1 Cl⁻ channel from *E. coli*; Two from *Homo sapiens*, the Kir2.3 K⁺ channel and the ClC-1 Cl⁻ channel and ClC-0 from *Torpedo californica*. Each sequence was codon optimized for *S. cerevisiae* and expressed with a C-terminal yeast codon-optimized TEV-GFP-His₁₀ tag.

All expression constructs were generated by co-transformation of the *Bam*HI, *Sal*I and *Hind*III triple digested yeast expression plasmid pEMBLyex4 [11], and one or two ion transporter PCR fragment(s) with a standard TEV-GFP-His₁₀ PCR fragment. None of the PCR products were purified but taken directly from the PCR tube after verification of PCR products by agarose gel electrophoresis. The expression plasmid was digested with three enzymes simply to ensure minimal amounts of non-digested plasmid, which would give rise to a large percentage of non-recombinant yeast transformants.

To reduce the costs of purchasing long synthetic DNAs, channel cDNA exceeding 2.5 kb was ordered in two pieces that were PCR amplified individually and designed to create a 25-30 nucleotides long identical sequence used for homologous recombination. The His₁₀-tag and the
Tobacco Etch Virus (TEV) protease cleavage site can be used for metal affinity purification and subsequent GFP-His$_{10}$ removal, respectively.

**Sites directed mutagenesis**

The two most favored site directed mutagenesis protocols involve either several PCR reactions [12] or restriction enzyme digestion of *in vitro* replicated DNA to remove the non-mutated plasmid DNA strand (QuikChange™ Site-Directed Mutagenesis System developed by Stratagene, USA), which are both time consuming and the latter costly. However, homologous recombination offers a simple, cheap and highly competitive mutagenesis protocol involving fewer *in vitro* manipulations.

We used the wild type protein expression vectors to introduce mutations by PCR as depicted in Figure 3.

N-terminal deletions were generated by PCR with a new forward primer together with the wild type reverse primer. Conversely, C-terminal deletions were introduced with a new reverse primer together with the wild type forward primer.

To introduce a point mutation or several point mutations simultaneously, each codon to be mutagenized was mismatched with a different codon in the PCR primer(s). Primers were designed so each sequence flanking the mis-matched nucleotides was estimated to have a melting point of 30°C. The two primers used for introducing the mutations were reverse complementary. Introduction of each point mutation required two PCR reactions. The reverse complementarity of the two mutation primers resulted in incorporation of the same short stretch of nucleotides in each PCR product required for homologous recombination.

From cloning 34 constructs, we obtained an 86% success-rate in the first cloning attempt and 100% after the 2$^{\text{nd}}$.

**Co-expression of two cDNAs from the same plasmid**

Assembling two or more expression cassettes on the same plasmid is often a tedious process. However, homologous recombination offers a simple and efficient way to construct such plasmids. As an example, Figure 5 illustrates the strategy we successfully pursued to create a
plasmid co-expressing human GFP-KCNQ1 with human KCNE1-RFP from identical promoters located on the same plasmid. KCNQ1 is a 676 amino acids long protein with six transmembrane segments that has to assemble into a tetramer to create a functional K⁺-channel. KCNE1 is a 129 amino acids long single pass membrane protein, which regulates KCNQ1 activity [13]. We initially assembled the two individual constructs, GFP-KCNQ1 and KCNE1-RFP in our expression plasmid, using the approach described in Figure 2. We subsequently amplified the Promotor-KCNE1-RFP-Terminator fusion using PCR primers with 35 nucleotides long 5’ extensions, to direct homologous recombination to the identical nucleotide sequences in the GFP-KCNQ1 expression plasmid. Four randomly picked yeast transformants were inspected by fluorescence microscopy, and they showed membrane associated GFP- and RFP fluorescence, confirming correct assembly of both expression cassettes in the same plasmid. Time dependent accumulation of GFP and RFP fluorescence, together with live cell bioimaging, demonstrated that both channel subunits accumulated synchronously and appeared to co-localize in vivo, Figure 4.

Creation of a pET based expression vector for homologous recombination-cloning in yeast

Due to the obvious advantages associated with cloning by homologous recombination in yeast, we generated a pET [14] based E.coli expression plasmid that can replicate in yeast. Figure 5 shows how easily such a plasmid is assembled in yeast. In just two steps we created an E. coli plasmid that carries the region required for T7 promoter dependent expression, together with the Tetracycline resistance gene (TetR) and the lac repressor gene (lacIq). The nucleotide sequence in Figure 6 shows how cDNA sequences could easily be inserted into this plasmid without compromising the optimal sequence and distances e.g. between the Shine-Dalgarno sequence and the translational start site. Clean expression vectors are generated exactly as described in Figure 1 by homologous recombination in yeast.

Discussion

Optimizing recombinant protein production often involves continuous alterations of vector design, a process that can be very laborious and time consuming. Here we demonstrate an efficient and streamlined procedure particularly suited to generate expression plasmids with minimal hands-on time. We have successfully used this approach to optimize expression of...
difficult human membrane protein targets in yeast [15] [16]. We furthermore show how several expression cassettes easily and effectively can be assembled into a single expression plasmid in an attempt to control stoichiometry of co-expression. In addition we demonstrate that the procedure is well suited for introduction of one or more mutations simultaneously. The constructs described in the present paper all encode GFP fusions which simplifies screening for the correct plasmids. However, due to the high efficiency of the homologous recombination cloning procedure we have used the exact protocol to generate plasmids expressing non-fluorescently tagged proteins; one just leaves out the fluorescent microscopy step depicted in Figure 1.

The protocol is based on the unique preference of baker’s yeast to perform homologous recombination; one of the major reasons for the immense success of this organism as a eukaryotic model system. Homologous recombination has been described before as a tool to generate plasmids but here we focus on the specific features required for optimal recombinant protein production.

Homologous recombination is particular valuable for construction of expression plasmids, as it offers all the flexibility desired to generate expression constructs without introducing any contaminating sequences into the final expression construct. Such cloning scars have previously been shown to reduce the expression level significantly [6].

Homologous recombination as a plasmid construction tool is obviously interesting when *S. cerevisiae* is the expression host. However, it is applicable for generating expression plasmids for any host organism. Hijacking the unique homologous recombination capacity of *S. cerevisiae* simply requires addition of the 2 micron yeast origin of replication and a yeast selective marker, e.g. the URA3 gene, to one’s favorite expression plasmid.

Homologous recombination does not require any commercial molecular biology kit and can be used without optimization, which makes it easy for any lab to implement this method. A standard molecular biology laboratory would only have to learn to transform yeast and isolate DNA from yeast; something that is just as simple as transforming and isolating DNA from *E. coli*. The construction of plasmids by homologous recombination in yeast takes a couple of days more than cloning directly into *E. coli*. Nevertheless, there are many benefits, such as; very low cost, a simple procedure, little hands-on time, high success-rate, in particular for assembling several
DNA fragments and can be used to assemble any express plasmid without having to introduce contaminating nucleotides. We find that all of these benefits, more than compensates for the extra couple of days required to finalize plasmid construction.

Conclusions

In conclusion, we demonstrate that homologous recombination is an easy-to-access and versatile cloning technique, which makes design and construction of any kind of expression vector cheap and straightforward with limited hands-on time. The method can be used to generate expression plasmids for any organism. One simply needs once and for all, to equip the favorite plasmid with a yeast 2 micron origin of replication and a yeast selective maker. The more complicated the cloning becomes the more one benefits from using homologous recombination as illustrated for assembly of two expression constructs on the same plasmid.

Material and methods

S. cerevisiae and E. coli strains

S. cerevisiae PAP1500(a ura3-52 trp1::GAL10-GAL4 lys2-801 leu2A1 his3A200 pep4::HIS3prb1A1.6R can1 GAL) [17] and Omnimax F' [proAB+ lacIq lacZΔM15 Tn10(TetR) Δ(ccdAB)] mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 Δ(lacZYA-argF)U169 endA1 recA1 supE44 thi-1 gyrA96 (NatR) relA1 tonA panD, (Thermo Fisher, USA) were used throughout the present study.

Assortment of ion transporting membrane proteins

For the following proteins KcsA (POA334), Kir2.3(P48050), NavAb (A8EVM5), CLC-ec1(P37019), CLC-1(P35525), CLC-0(P35522), KCNQ1(P51789) and KCNE1(P15382) full-length cDNA were purchased as codon optimized for yeast from Genscript, USA. cDNAs exceeding 2.5 kb were bought in two or more pieces to minimize the cost of synthetic DNA.

Design of primers for full-length recombinant protein and mutations

The template specific part of all PCR primers were designed to have a melting temperature, Tm, of 58°C calculated from the equation, Tm = 2 X (A+T) + 4 X (G+C) [18]. Primers for site directed mutagenesis were designed to carry the mismatched nucleotides in such a way that each of the two flanking nucleotide sequences has a melting temperature of 30°C. Primers used for generating PCR products are shown in Table 1.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Forward primer (5'-3')</th>
<th>Reverse primer (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard TEV-GFP-His&lt;sub&gt;10&lt;/sub&gt;</td>
<td>GAAAATTTGATATTTCGAAGATCAATTGATGTCTAAAGGTGAAGAATTATTTCGTC</td>
<td>CTTCAATGCTATTCAATTTCCCTTGGATATGGATCATTTGTACCAATCCACCATCATCACCATTCC</td>
</tr>
<tr>
<td>Variant</td>
<td>Nucleotide Sequence 1</td>
<td>Nucleotide Sequence 2</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>CLC-01-532</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATAAATTTTTCATGCAGAGGACACAAATTTC</td>
</tr>
<tr>
<td>CLC-1a WT</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>CAAAAATACACTCTGGAAATAAAC</td>
</tr>
<tr>
<td>CLC-1b WT</td>
<td>AAATTATGGCAATATTTCTCCAGATGGATTTTCTTA</td>
<td>AAATTGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>E232A</td>
<td>CAGTTGGTAAAGCCAGTGCTTCCATATTGT</td>
<td>ATCATGAGGTAGGCAAGCCAATGAATCCAGATTC</td>
</tr>
<tr>
<td>I290W</td>
<td>GATTTGTTTTCAACACTGAAAGATCAGATTCTTC</td>
<td>GAAGTAGCTTTCAAGGGAATAAACAAAA</td>
</tr>
<tr>
<td>I556W</td>
<td>AAATGCTCTGAAAATAAATTATTTTCATGATTCTTCT</td>
<td>ATCATGAGGTAGGCAAGCCAATGAATCCAGATTC</td>
</tr>
<tr>
<td>CLC-171-598</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>Remove</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>Nterm1-114</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>CLC-ec1 WT</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>I201W</td>
<td>TATCTTGTGTTTTCAAGGAGAAATGATCGGTCTTTGTT</td>
<td>CAAAACACAAA</td>
</tr>
<tr>
<td>I422W</td>
<td>CTACAAATTTATTTTCAGTTTGCTGCTTACCTTCA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>E148A</td>
<td>TTGTAGGTATGCGAGCCATTTTCATACCTT</td>
<td>CGTTGGGACTCTGCTCTACCTAAA</td>
</tr>
<tr>
<td>Y445A</td>
<td>GTAAACCATTGGCCTCTGCTTT</td>
<td>TAAAATGCAAGGGCCATGATTTC</td>
</tr>
<tr>
<td>E203A</td>
<td>GTTTATTGCAGGAGCAGGCAATGAGG</td>
<td>CGTTGGGACTCTGCTCTACCTAAA</td>
</tr>
<tr>
<td>Kir2.3 WT</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>trunkated</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>both N and C-term</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>CICec1 WT</td>
<td>ACACCAAAATACACACACTAAATTACGGATCAATTCTA</td>
<td>AAAATGACTTTGAAAAATACAAATTTC</td>
</tr>
<tr>
<td>E148A</td>
<td>TTGTAGGTATGCGAGCCATTTTCATACCTT</td>
<td>CGTTGGGACTCTGCTCTACCTAAA</td>
</tr>
<tr>
<td>I201W</td>
<td>TATCTTGTGTTTTCAAGGAGAAATGATCGGTCTTTGTT</td>
<td>CAAAACACAAA</td>
</tr>
</tbody>
</table>

67
<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I422W</td>
<td>ACTACCAATTGTGGTTGCAATGAT</td>
<td>ATCATTGGCAACCACAATTGGTAGT</td>
</tr>
<tr>
<td>E203A</td>
<td>GTTTATTATCGAAGCATAATGGCCA</td>
<td>TGGGCTCATTTGCTTCGATAATAAAAC</td>
</tr>
<tr>
<td>Y445A</td>
<td>GTAAACCATGGGCTCTGCTATTTTA</td>
<td>TAAAATAGCAGAGCCCAATGGTTTAC</td>
</tr>
<tr>
<td>ohKCNE1</td>
<td>ACACAAATACACACTAATAATTACCGGATCAATTC</td>
<td>CTTCCTACCATTGAAACCATAATTGGACTTTGAA</td>
</tr>
<tr>
<td></td>
<td>AGATAATATGATTTTATCCAATACAACCGC</td>
<td>AATACAAATTTTCAGGGGAAGGCTTTGCTGG</td>
</tr>
<tr>
<td>KCNQ1</td>
<td>GAAAATTTGTATTTTCAAAGTCAATTTATGGCTGCTG</td>
<td>TCAAAAATCATCGCTGCCTGGTCAAGACGTGTCGGGGCGCAGCC</td>
</tr>
<tr>
<td></td>
<td>CCTCAAGTC</td>
<td>CTTCATTGCTATCATTTCTGATATTGGA</td>
</tr>
<tr>
<td>yoRFP</td>
<td>ATGGTTTCAAAAGGTGAAGAACCTTCATCTGGACCTC</td>
<td>ATGCGAGCTGCGCACATACAG</td>
</tr>
<tr>
<td>KCNE1</td>
<td>CTTCGGCTTCTTGCTTCGACCTTGA</td>
<td>TCAAAAATCATCGCTGCCTGGTCAAGACGTGTCGGGGCGCAGCC</td>
</tr>
<tr>
<td>cassette</td>
<td>TCCTGGCTTCTTGCTTCGACCTTGA</td>
<td>ATGCGAGCTGCGCACATACAG</td>
</tr>
</tbody>
</table>

Table 1: Nucleotide sequence of primers used to PCR amplify targets to be used for homologous recombination cloning. The two purple sequences in the fw and reverse primers are reverse complementary and encode the TEV protease site used for homologous recombination between ion-channel DNA and the standard GFP fragment; the red sequences in the fw and rv primers are reverse complementary and generate a homologous recombination site between 5’ target DNA and the expression plasmid; the grey sequence is the Kozak sequence from the yeast PMR1 gene; underscored sequences are the codons to be altered by site directed mutagenesis; orange sequence is for recombination between the 3’ end of a PCR fragment and the expression plasmid; the turkish sequence encodes a His10 tag; green, recombination site for insertion of an additional expression cassette; dark grey, recombination sequence for insertion of an additional expression cassette. Template specific sequences are depicted without shading. Color codes are the same as used in Figures 1-5.

**Plasmid construction and transformation**

The PAP1500 yeast strain was transformed according to [19] and selected on SD medium containing 2% glucose and 30 mg/liter leucine and 20 mg/liter lysine.

cDNA was PCR amplified with Accupol DNA polymerase (Amplicon, Denmark) and confirmed to give a single band by agarose gel electrophoresis. Expression plasmid pPAP2259 [11] was treated for 2-3h with Fast digest versions of SalI, BamHI, HindIII (Thermo Scientific, USA) and subsequently heat-inactivated at 80°C for 10 minutes.

The expression plasmids were generated by in-vivo homologues recombination by transforming the protein encoding PCR fragments, the standard GFP PCR fragment and the digested vector into S. cerevisiae strain PAP1500[11]. Four individual yeast transformants were picked and inoculated O/N at 30°C in 5 ml SD media. 1 ml cell culture in 25% glycerol was stored at -80°C until later. 100µl cell culture was transferred to 3 ml SD media supplemented with 2% galactose.
and grown O/N at 30°C. Next morning each culture was screened for GFP fluorescence and RFP fluorescence (if cells should also express an RFP tagged protein) at 1,000 times magnification in a Nikon Eclipse E600 fluorescence microscope equipped with an Optronics Magnafire (model S99802) camera attached. The remaining cells were harvested, treated with Lyticase and plasmid was isolated using the Machery-Nagel E. coli miniprep kit as described before and transformed into Omnimax cells according to [20]. Colonies selected on Tetracycline containing medium were picked and transferred to liquid LB medium. Next day 1 ml of each E.coli cell culture was frozen in 25% glycerol at -80°C. Plasmids rescued into E. coli were sequenced at Eurofins Genomics, Germany to verify the expected nucleotide sequences.

**Declarations**

**Ethics approval and consent to participate:** Not applicable.

**Consent for publication:** Not applicable.

**Availability of data and material:** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Competing interests:** The authors declare that they have no competing interests.

**Funding:** Research in the authors’ lab is supported by The National Danish Advanced Technology Foundation, The Innovation Foundation, The Independent Research Fund Denmark and the Novo-Nordisk Foundation.

**Authors' contributions:** SSP planned and generated the majority of expression plasmids under supervision of PAP. SSP wrote the initial version of the manuscript and the final version in association with PAP.

**Acknowledgements:** The authors thank David Sorensen for excellent technical assistance.
Bibliography

1. Insights, G.M. Biotechnology Market will expand at 9.9% CAGR to hit $775 Billion by 2024: Global Market Insights, Inc. 2019.


**Figure 1: Work flow for plasmid construction by homologous recombination.** At day 1 yeast cells are co-transformed with the linearized plasmid and the PCR products required for assembling the desired expression plasmid. After two to three days, four individual colonies are picked, transferred to selective liquid medium and cultured overnight. At day 5 up to three procedures are carried out in parallel: I, glycerol stocks are kept at -80°C of the four transformed *S. cerevisiae* strains; II (only relevant for fluorescently tagged proteins), overnight induction of a small culture of each of the four transformed yeast strains, followed by fluorescence microscopy at day 6 to identify cells expressing the desired fluorescent fusion protein; III, at day 6 total DNA is isolated from the remaining part of the yeast culture and transformed into *E. coli*. At day 7, *E.coli* transformants are picked, cultured overnight in selective medium and at day 8 a glycerol stock of each *E. coli* transformant is frozen at -80°C, while the plasmid is isolated from the remaining cells and sequenced.
Figure 2: Creation of plasmids for expressing wild type cDNA sequences C-terminally tagged with GFP. A, the expression plasmid is linearized by digestion with three restriction enzymes B, BamHI; S, SalI; and H, HindIII solely to minimize the presence of circular plasmids before transformation. Colors indicate identical nucleotide sequences used for *in vivo* homologous recombination in the yeast host strain. Each color represents a 25-35 nucleotides long sequence present as a 5’ extension in the primer used to generate the PCR fragments to be assembled. The site of integration is determined by the red and the orange sequences in the expression plasmid. The figure shows a case where the cDNA sequence is so long that it needs to be purchased in two pieces (P1 and P2) to obtain the lowest price. R1-4 indicate the homologous recombination sites responsible for the four crossover events required to correctly assemble the final expression plasmid in yeast. B, the final structure of the expression plasmid obtained by *in vivo* homologous recombination in yeast after transformation with the linearize plasmid and the three depicted PCR products.
A

In vivo assembly in yeast

B

In vivo assembly in yeast

pEMBLyex4
**Figure 3: Generation of mutations by homologous recombination.** A, Introduction of a single or multiple nucleotide substitution(s) (*) in wild type cDNA, simply requires co-transformation of yeast with the indicated PCR fragments and the linearized expression plasmid. The combinations of forward (fw) and reverse (rv) primers used to generate the individual PCR products are shown. The GFP PCR fragment (PCR3/PCR5) is the same as PCR3 in Figure 2. PCR1-4 indicate cDNAs encoding small successive parts of the protein; B, creation of cDNA encoding an N-or C-terminally truncated protein C-terminally fused to GFP is simply done by co-transforming yeast cells with PCR1 and PCR2. To generate a 5’ deletion, PCR1 is generated by Nmutfw and wtrv primers while a 3’ deletion is created by using primers wtfw and Cmutrv. PCR 2 (the TEV-GFP-His10 fragment) is identical to PCR3 used in Figure 2. DNA sequences of the same color are identical and between 20 and 35 nucleotides long. P, cDNA encoding the wt protein; P’ indicates cDNA encoding truncated versions of the protein. Crosses indicate the *in vivo* recombination events required to assemble the expression plasmid.
Figure 4: Assembly of two expression cassettes on the same plasmid. Application of in vivo homologous recombination in yeast to assemble a plasmid co-expressing the human KCNQ1 K-channel and its regulatory KCNE1 subunit. Plasmids A and B were generated by homologous recombination as depicted in Figure 2. The KCNE1-TEV-RFP expression cassette was subsequently PCR amplified with primes Rec1 and Rec2. The 5’ extensions of these PCR primers carry sequences identical to the red and gray sequences flanking the TthIII and NaeI restriction sites in plasmid “A”. Homologous recombination between the red sequences and between the gray sequences results in the plasmid shown in “C”. Insert “D” shows time dependent accumulation of GFP and RFP fluorescence in yeast cells after induction of GFP-KCNQ1 and KCNE1-RFP expression at time zero, while insert “E” shows a live cell bio image of yeast indicating co-localization of the two tagged K-channel subunits.
Figure 5: An *E. coli* pET58 derived expression plasmid for *in vivo* homologous recombination cloning in yeast. A, the nucleotide sequence responsible for T7 dependent expression in pET58b(+) was PCR amplified using primers Rec1 and Rec2 and co-transformed in yeast with *Bam*HI (B), *Sal*I (S) and *Hind*III (H) digested pEMBLyex4. The recombination sites in pEMBLyex4 are shown in red and orange, respectively. To introduce the lacI<sup>q</sup> allele and a tetracycline resistance gene into the plasmid we co-transformed yeast with *Bst*EII, *Aat*II and *Nru*I digested pPAP9592, a lacI<sup>q</sup> PCR fragment and a tet<sup>R</sup> PCR fragment. Inverse complementary sequences in the 5′ extensions (shown in red) of the lacI<sup>q</sup> reverse primer and the tet<sup>R</sup> forward primer assures *in vivo* homologous recombination between the two PCR products. Targeting to the intended green and yellow positions in the pPAP9592 plasmid is assured by an identical green sequence in the lacI<sup>q</sup> PCR product and an identical yellow sequence in the tet<sup>R</sup> PCR product, respectively. cDNAs to be expressed can be inserted into the blue and grey sequences simply by co-transforming yeast with pPAP10378 linearized with *Kpn*I, *Not*I and *Sac*I located in the multi cloning site (MCS) and the PCR amplified coding sequence to be expressed as depicted in Figure 6. The forward primer for PCR amplification should carry a 5′ extension identical to the blue region while the reverse primer should have a 5′ extension identical to the grey DNA sequence.
Figure 6: Nucleotide sequence from pET58b(+) contained in the pPAP10378 expression plasmid. The blue and the grey nucleotide sequences are the 5’ extensions to be included in the primers used for PCR amplification of the coding region to be expressed (indicated as blue and grey in pPAP10378). Location of the T7 promoter, transcriptional start site (TSS), lac operator site, Ribosome Binding Site (RBS), translation initiation codon (ATG), T7 transcriptional terminator and three unique restriction sites used for linearizing the plasmid are shown. The blue recombination site present in the forward primer assures that the translational initiation codon for the gene to be expressed is located at the optimal distance from the RBS and no cloning scars are introduced. The His10 coding sequence is used for recombination by the grey reverse PCR primer. This means that a C-terminal His10 tag can be introduced directly into the recombinant protein. Simultaneous introduction of several PCR products into pPAP10378 is done exactly as described in Figure 2.
CHAPTER 4 – Manuscript III

Synopsis

Following cloning and finding a suitable expression system for production of the ion channels described in Chapter 2 and Chapter 3, the next step was to establish a robust purification protocol. Membrane protein purification is expensive and involves many steps, including: Cell lysis, isolation of membranes by ultra-centrifugation; extracting the protein from the cell membranes by detergent – termed solubilization and purification by chromatographic methods.

The aim of the study presented in the manuscript was to establish a robust purification protocol to obtain large quantities of purified protein. When establishing the purification procedure, the later application for reconstitution into polymers for making ion selective membranes, had to be kept in mind. This is especially important in regards to choice of detergent and number of steps included in the purification procedure, since these will have a huge effect in the cost and industrial scale applicability.

The manuscript outlines the principle for creating constitutively open ion channels from ClC-ec1, ClC-1, ClC-0, KcsA and NavAb by introduction of various mutations. It shows that we could express all the mutant proteins in S. cerevisiae. Finally it outlines our systematic approach for establishing a purification protocol, using two ClC-ec1 mutant variants to evaluate the robustness of the procedure. On the basis of these results, a streamlined purification of the many mutant proteins would then be achievable.
Design and purification of constitutively open potassium, sodium and chloride channels - application for creating ion selective membranes

Sarah Spruce Preisler¹, Marc Friis¹ and Per Amstrup Pedersen¹

¹ Department of Biology, August Krogh Building, University of Copenhagen, Universitetsparken 13, 2100 Copenhagen OE, Denmark

Abstract

Background: Using purified ion channels for creating ion selective membranes with high ion flux, requires the production and purification procedure to be robust and yield large amounts of pure and stable protein. In addition, having a scalable purification procedure is essential for industrial scale fabrication of biomimetic membranes.

Results: We show that we able to create and express an array of mutations in selected ion channels, ClC-0, ClC-1, ClC-ec1, KcsA and NavAb, using S. cerevisiae as the expression organism. Comparing two purification procedures of two mutant variants of ClC-ec1, enabled us to evaluate the robustness of the purification procedure. We discuss considerations of applying the purification procedure for industrial scale production of the mutant ion channel variants.

Conclusions: Using an automated single step purification procedure yields large quantities of pure protein with limited hands on time. The use of ion channels for creating ion-selective membranes has many industrial applications, opening many future opportunities when paired with a robust purification procedure.

Introduction

Using ion channels for ion-selective membranes requires engineering to render them constitutively open. The ideal design would create “pore-only” channels that maintain ion selectivity, high ionic flow, and could be expressed in high yields in a stable and active form.

Ion channels are pore-forming membrane proteins that facilitate a flux of ions across biological membranes. Their basic molecular features are to create a membrane spanning pore to facilitate ion movement with a selectivity filter specific to a given ion. In addition they often have regulatory elements, which enable opening and closing only upon specific stimuli, such as voltage or binding of ligands. Different ion concentration across membranes is what creates the
membrane potential in living cells, hence permanently open ion channels would lead to cell death. Gating of ion channels is therefore a conserved trade, one which we want to remove to create constitutively open ion channels, with the aim to create ion selective membranes. To create ion-selective membranes using ion channels, the purified ion channels need to be inserted into a polymer, to create an artificial membrane. Many different polymers exist and the type of membrane created depends on the application it is needed for [1].

Recombinant production and purification of ion channels and membrane proteins in general is difficult due to their hydrophobic nature and require the use of detergents for extraction from the cell membrane and therefore detergent also need to be present in all buffers during purification, above the critical micelle concentration (CMC). This is the concentration were the monomer detergent assembles into micelles, and it varies a lot among different detergents; ranging from 0.001/0.01 %/mM to 0.53/20 %/mM among common detergents (Anatrace). Detergents are expensive compounds; therefore one would preferably be using a lower-cost detergent when establishing a large-scale purification protocol.

For this study we aim to create constitutively active selective ion channels which can be purified in high yields. This is done by removing excess parts of the proteins to enhance stability and removing gating mechanisms by mutations to remove the need for stimulus to open the channels. The aim is to be able to produce these ion channels on an industrial scale for insertion into polymers for making ion-selective membranes, which have many industrial applications.

Bacterial ion channels are in general simpler and more robust than their eukaryotic equivalents, which is why we preferred to use bacterial $\text{K}^+$, $\text{Na}^+$ and $\text{Cl}^-$ channels for this study. The structure and function of the ion channels we selected have all been studied extensively by mutational analysis, providing a large selection of mutations to choose from. We have therefore picked mutations based on these previous findings.

**Results**

**Engineering of constitutive open ion channels**

The concept behind engineering the selected ion channels to be constitutively open is described below for each type of ion channel. A detailed description of structure and function of the ion channels is out of the scope of this manuscript and have been described elsewhere [2-12].
KcsA
KcsA is a prototypical potassium ion channel from the bacteria *Streptomyces lividans*. It has a simple structure with two transmembrane (TM) helices and it associates as a tetramer with a central ion pore containing the selectivity filter. The channel has an intracellular gate formed by the ends of TM2. This TM helix further extends into the cytoplasmic domain, which forms a 4-helix bundle in the tetramer, and acts as a second steric gate on channel opening [5, 13]. Therefore we have chosen to remove this C-terminal part of the protein, expressing only residues 1-135 (Figure 1A,B). Furthermore we chose the E71A mutation, which have been shown to remove voltage dependence and inhibit so-called c-type inactivation, which stops the channel from opening [14].

NavAb
NavAb is a voltage-gated potassium channel from *Arcobacter butzleri* and it was the first structure solved of a voltage gated sodium channel [7]. It has 6 TM segments (S1-S6), where S1-S4 forms the voltage-sensing domain (VSD) and S5-S6 forms the pore domain by associating as a homotetramer, with a pore in the middle. It has a C-terminal domain (CTD) that acts as a steric gate for the ion pathway together with a so-called activation gate (Figure 1C,D).

We chose to remove 40 residues of the CTD, expressing only residue 1-226, which have been shown to favor the open state of the NavAb channel pore [6]. To create a simple pore-only Na-channel we removed the VSD and expressed it with and without the 40 residues of the CTD (Figure 1C,D). Pore-only constructed proteins, has shown to function as Na\(^+\) selective pores and to render channel opening voltage-independent, from other Nav channels [15]. The residues of the activation gate were mutated to the neutral residue alanine (I217A/Met221A), to remove the activation gate. NavAb activates at very negative voltages, so the A49K mutation was chosen, because it has been shown to shift the activation to more positive voltage and also removes a type of inactivation seen in this sodium channel [24].

Chloride channels
Chloride channels are homodimeric channels where each monomer contains eighteen helices. Eukaryotic ClCs, like ClC-0 and ClC-1, contain a large C-terminal domain with two so-called systathionine-β-synthase (CBS) folds [10].
We set out to use bacterial ion channels. However, for chloride transport we discovered a dilemma. The bacterial ClC-ec1 is not a channel but a transporter, and the eukaryotic ClC-0 and ClC-1 are very large proteins, possibly making them more difficult to express in high yields and quality. So we chose to use all three; to try to make the large ClC channels smaller by removing their C-terminal domains, ClC-1(1-598) and ClC-0 (1-532) and to try to convert the transporter into a channel by mutating Y445A and E203A, shown to impair H+ transport of this protein [16, 17]. All three ClC channels share the same mechanism of gating via a glutamate residue, which blocks the ion pathway, but moves away when the carboxyl sidechain is protonated. We chose to mutate this residue to an alanine in all three channels (ClC-1 E232A, ClC-0 E166A and ClCec1 E148A), mutating this glutamate has shown to remove the gating [9]. Lastly, we chose to insert two large tryptophans in the interface between the monomers, which has been shown to make ClCec1 monomeric, while maintaining chloride transport function and stability [18].

Creating expression constructs and expression screening

The mutations of each ion channel were introduced into the WT sequence via PCR primers as described in Manuscript II. The DNA sequences that had been codon optimized for expression in \textit{S. cerevisiae}, to generate plasmids for expression of mutated ion channel sequences. All proteins were expressed as fusions with a C-terminal TEV-GFP-His\textsubscript{10} tag to aid protein purification, quantification of expression and subsequent tag removal if needed. We utilized our go-to expression system, where high expression levels are obtained with an galactose inducible promotor, enhanced by the expression of the GAL4 transcriptional activator [19, 20], with the expression yield further increased by using a high copy plasmid [21].

After cloning and construction of expression vectors, we used fluorescence microscopy to determine the localization of the fusion proteins. As seen in Figure 3 and Supplementary Figure 1, all WT and mutant ion channel variants localized to membranes inside \textit{S. cerevisiae}, as seen from distinct localization, rather than the whole cell cytosol being fluorescent.

Our lab has previously shown that across many different types of membrane proteins, the yield and quality of protein expression is higher at a lower expression temperature [20, 22, 23]. Figure 4 shows that all mutants could be expressed in yeast, grown in minimal selective media (SD) at 15 °C, as described in material and methods. In general, ClC-ec1 gave the highest expression level, followed by ClC-0 and NavAb, while ClC-1 and KcsA had the lowest expression for both WT and mutants.
Figure 4A show that most of the ClC-0 mutants reached an expression optimum around 70 hours post induction. The WT and V490W variants had the lowest level of expression, while 1-532 and E166A had the highest.

ClC-1 mutant expression curves showed a varying degree of expression (Figure 4B). The double mutant obtained the highest level of expression, followed by the wildtype, while I556W and 1-598 had the lowest. All variants reaches maximum expression level 72 hours after induction.

The expression curves of ClCe1 showed that WT and E148A reached maximum expression yield after 70 hours, followed by a dramatic decrease in fluorescence (Figure 4C). While the other ClCe1 mutants accumulated slower, they maintained expression levels over time. The triple mutant reached the highest expression level of all the ClCe1 variants, at 130 hours after induction.

Of the three KcsA variants, the WT showed the highest level of expression together with the 1-135 variant (Figure 4D). E71A had a significantly lower level of expression, and did not exceed 40 fluorescence units per OD$_{450}$.

Expression levels of the NavAb mutants varied a lot among them (Figure 4E). The two NavAb mutants M221A/I217A and N49K had around the same expression level as WT, while 123-267 and 123-226 had a higher expression level. Only the 1-226 mutant had a lower expression yield than the WT variant.

Once optimum for protein expression had been established the cells were grown in 2L batches and harvest at this time-point.

**Establishing robust and scalable purification of the ion channels**

We set out to test if a standard one or two step purification procedure would result in high amounts of pure protein. Establishing a robust purification protocol is essential for scaling up to industrial production, which would be needed for creating ion-selective membranes.

We selected two ClCe1 variants; the double and triple mutants as a case example for developing our purification protocol. Both of these mutants expressed well, and we could compare purification results between two variants of the same protein, hereby enable evaluation of the robustness of the procedure.
Prior to this study we had screened a number of detergents for their ability to solubilize WT CIC-ec1 (Supplementary Figure 2), and we assume that the double and triple mutants behave like the WT protein in the detergents, and therefore did not screen detergents on these variants. CIC-ec1 solubilized well and was monodisperse in all detergent tested (described in Manuscript I). We therefore selected two different detergents based on the later application of inserting the protein into polymersomes. Since the protein behaved well in all detergents, we chose to test purification in LDAO because of its low price and DM because it has a high critical micelle concentration (CMC), which would make it easier to remove during reconstitution.

The proteins were purified by utilizing the C-terminal His-tag, with immobilized metal affinity chromatography (IMAC) on a Ni column. First the double and triple mutant CIC-ec1 proteins, solubilized in LDAO, were bound to the HisTrapFF Ni-column. A linear gradient of 0 up til 700mM imidazole was used, which enabled us to identify the imidazole concentration required to elute the protein from the column. An extra washing step, using 1000 mM imidazole, was added to see if all protein had been eluted from the column during the gradient. Figure 5A, B shows the IMAC elution profiles, where we see that both proteins elute at an imidazole concentration of around 200 mM, and that most of the protein eluted in one main peak for both CIC-ec1 mutant variants. Apart from following A280 signal during the purification, we collected the elution as fractions and measured the GFP fluorescence in these, which verified the elution of our specific fusion proteins (Supplementary Figure 3).

SDS-PAGE analysis of the top elution fractions by in-gel fluorescence and Coomassie-stain of the double and triple mutants purified in LDAO by IMAC revealed that in addition to the high concentration of the proteins, they both were very pure. Fluorescence in-gel show the migration of the target protein via the GFP-tag, while Coomassie stain revealed the total protein content of the sample, hence a comparison of the two easily show the purity of the samples. Both proteins were seen as a single band at ~60 kDa, in Figure 5A, B. The in-gel and Coomassie stain showed that the samples were very pure, only with a minor contaminant band around 70 kDa.

We then proceeded to purify the double and triple mutant variants of CIC-ec1 in DM by IMAC, using a step gradient of; 40 mM, 400 mM and 1000 mM imidazole. As expected from the previous linear gradient elution of the double and triple mutant variants of CICec1, both proteins eluted during the 400 mM imidazole washing step in a single symmetrical and narrow peak (Figure 5C,D). Again we verified the elution of our target proteins by measuring the GFP in the
eluted fractions (Supplementary Figure 3). The SDS-PAGE analysis of the top fraction revealed pure protein samples after the IMAC purification in DM (Figure 5C,D).

All top fractions from the IMAC purifications containing the double or triple mutant ClC-ec1, seen by GFP fluorescence (Supplementary Figure 3), were pooled and concentrated to 1 ml. This was done for both ClC-ec1 mutant proteins, purified in both DM and LDAO. The samples were then further purified by size-exclusion chromatography (SEC). Elution from the SEC column was followed by A$_{280}$ measurements (Figure 6) and GFP fluorescence in the eluted fractions was measured to verify elution of our target proteins (Supplementary Figure 3).

The SEC chromatograms in Figure 5 reveal that both the double and triple mutant ClC-ec1 eluted in a single narrow peak in DM. In LDAO, the double mutant had a shoulder on the left side of the peak, and the triple mutant had three elution peaks. The fraction corresponding to the first peak in LDAO, revealed white precipitation visible by the naked eye (Supplementary Figure 4C). This indicated that the protein aggregated during the purification in LDAO.

The top fraction from each SEC elution where separated by SDS-PAGE, and analyzed by in-gel fluorescence and Coomassie staining (Figure 6). This revealed that the purification resulted in pure protein samples, and that the sample purity looked the same as after the IMAC purification (Figure 5). Therefore there was no big improvement in protein purity after this second purification step. The purification yields are summed up in table:

<table>
<thead>
<tr>
<th>ClC-ec1</th>
<th>Solubulization yield (%)</th>
<th>IMAC yield (%)</th>
<th>Yield after SEC mg/L culture</th>
</tr>
</thead>
<tbody>
<tr>
<td>E148A+Y445A in LDAO</td>
<td>80</td>
<td>50</td>
<td>4</td>
</tr>
<tr>
<td>E148A+E203A+Y445A in LDAO</td>
<td>77</td>
<td>65</td>
<td>3.5</td>
</tr>
<tr>
<td>E148A+Y445A in DM</td>
<td>60</td>
<td>80</td>
<td>4.5</td>
</tr>
<tr>
<td>E148A+E203A+Y445A in DM</td>
<td>70</td>
<td>78</td>
<td>7.3</td>
</tr>
</tbody>
</table>

**Table 1.** Yields are given in %. Solubilization is the relative fluorescence in the sample before and after ultracentrifugation. IMAC yields are estimated from the total fluorescence units loaded onto the columns relative to the total fluorescence units eluted in the flow through, hence did not bind to the column. The yield after SEC was calculated from the protein concentration determined by NanoDrop relative to the volume of cell culture the proteins were isolated from.
Discussion and Future outlook

In this study we present the mutations for design of constitutively open ion channels (K⁺,Na⁺,Cl⁻), and their expression in *S. cerevisiae*. We demonstrate how we set out to establish a purification protocol for these ion channels. We selected two ClC-ec1 mutants to compare purification results between the two variants, and evaluate the robustness of the purification procedure.

Creating constitutively active ion channels requires detailed understanding of the structural and functional principles for channel gating and molecular basis for ion permeability. Therefore we took the opportunity to use ion channels that had already been studied in detail, on a functional level and in molecular details of the structure. Hereby we saved an enormous amount of work, avoiding characterization of the different mutations.

Our data demonstrate that it is possible to create an array mutant protein and screen these in a high throughput manner, by utilizing the GFP tag. If needed, this could also be used to screen for optimal expression conditions such as temperature or type of media, to increase the expression yields. The GFP tag furthermore enabled *in vivo* localization, easy quantification of expression yields and optimization of protein during purification. Therefore we find that this is a neat tool to consider using when optimizing protein production strategies.

The expression level of WT protein and mutant variants varied a lot among all the ion channels and it is difficult to say if this is because of differences in stability, if constitutively open channels are more toxic to the yeast cells of differences in biogenesis and folding. However, it is noticeable that the two mutations rendering NavAb as a “pore-only” channel have the highest expression levels, which could be because this more compact structure makes the protein more stable and easier to express and fold. This protein has a swapped-domain fold, meaning that the VSD of one monomer folds around the adjacent PD [7]. It is still not fully understood how the cell achieves this, but by removing the VSD, we remove the need for this complicated folding of the ion channel.

Even though there are currently not much published data on large scale/industrial scale purification available, there is great potential of utilizing the His-tag for IMAC purification in large scale. We demonstrated that using the His-tag for Ni-affinity purification, yielded a large amount of pure membrane protein. This is a good choice of purification method, since the high affinity of the column for the tag ensures high enrichment of the target protein in a single
purification step. In addition this column material can be used with detergent and tolerate the high salt concentrations used when washing the column to remove unspecific bound proteins. However, one needs to keep in mind that this type of column is not suited for samples that require chelators (e.g., EDTA).

Our choice of detergents was mainly based on the later application for incorporation of the ion channels into polymers. Our results revealed a big difference in the final quality between the proteins purified in LDAO and DM, as seen from the SEC profiles (Figure 5). We chose the detergents based on a screen performed on the WT ClC-ec1 variant, and therefore assumed that the mutants behave like the WT protein in a given detergent. While this might be a valid assumption for point mutations, screening of mutant variants with larger parts of the proteins removed might be required.

The choice of detergent needs a quantitative weighing of cost versus quality and yield of the purification, keeping later application for reconstitution in mind. While LDAO has a lower CMC and cost less, it induced aggregation during purification of the ClC-ec1 mutant variants. The triple mutant even showed a high degree of precipitation, visible by the naked eye. DM SEC profiles showed to be very narrow and symmetrical, an indication that this detergent keeps the proteins stable in solution over time. A downside of DM would be that it has a higher CMC, meaning that it needs to be present in higher concentrations in all purification buffers. In addition, DM costs much more than LDAO, which would significantly increase the price of purifications. Finding a cheaper detergent that is able to maintain protein quality would be of great interest prior to scaling up purification volumes. To this end we would probably need to test the purification protocol in more detergents.

In this study we selected two mutant variants of ClC-ec1 for purification in two different detergents for a comparison of the purification outcome, to validate the robustness of the procedure. While there was a difference in stability in the two detergents, the two chromatographic purification steps were highly reproducible. The SDS-PAGE analysis of the IMAC and SEC purified samples revealed that there was no significant difference in purify between them. This would suggest that it would be possible to skip the second purification step, also benefiting the time and equipment needed for purifications, especially for purification on an industrial scale. Before scaling up protein purification, we would also need to consider the availability and cost of equipment and reagents. We chose to use a chromatography platform that
can accommodate large scale purifications and both cleaning of columns and purification procedures showed to be possible to do in a highly automated manner on the Äkta system, hereby minimizing hands-on handling.

It is often difficult to obtain high yields of membrane proteins, but we showed that using the *S. cerevisiae* expression system setup and our purification protocol, we were able to express enough of the ClC-ec1 mutant proteins to give a purification yield between 3.5 to 7.3 mg of protein/L culture. To evaluate this one has to take into consideration that these yields were obtained in shake flasks at a low optical density. Growth in bioreactors would certainly increase the amount produced. This most encouraging is that the density per cell is high, which is the most important with respect to purification of an intracellular protein. However, yields will vary a lot depending on the efficiency of all the steps towards a purified protein sample. Apart from the solubilization efficiency of the detergents, there was also a noticeable difference in how well the protein bound the HisTrap column. For LDAO we lost about 40%-50% of the protein in this purification step and around 20% for DM (Table 1), which show that there would be a lot to gain by improving this step in the purification.

We used the GFP tag to quantify the yield of protein in the different steps of the purification (Table 1). This could easily be applied to screen different conditions for cell lysis, detergent solubilization, IMAC binding and elution, to obtain as large amounts of purified protein as possible.

The purification protocol outlined here showed to be completely reproducible between the two ClCec1 mutation constructs. The next step would then be to verify if this procedure could be applied for the purification of the rest of the ion channel mutation constructs. Our results show that using a single IMAC purification step in a highly automated procedure is a good starting point for the possibility to purify all selected ion channel mutant variants.

Creating ion selective membranes would require the ion channels to be incorporating into polymers, which might require removal of the GFP-tag. This could be done either by cleavage with the site-specific TEV protease [24] or by making new expression constructs fused only to a His-tag without the GFP. Expressing the ion channels only with a His-tag removes the need of tag removal by TEV protease digestion.
For the purpose of making ion-selective membranes, the application of the production protocol for industrial scale needs to be considered. Mainly the cost compared to the yield, of each step in the purification needs to be considered. We find that using *S. cerevisiae* as a production organism gave a high expression yield, and has the advantage of low cost media and can be cultured in large bio-reactors, which would be needed to obtain larger amounts of protein, and also have the additional benefit of controlled growth conditions for the yeast cells.

In this study we presented a novel idea for engineering ion channels for the application of creating ion selective membranes. These membranes have the potential use for many different applications, where biopharma- and food industry would be possible markets for these new technologies, where ion selective membranes could be used to enrich or remove specific ions e.g. avoiding free calcium ions when isolating whey would increase protein stability [25].

**Conclusion**

We describe our efforts to design and produce constitutively open ion channels. We show that it is possible to express these in high yield and we test the robustness of the purification protocol. We discussed different consideration for the purification procedure to be applied for industrial scale production of the mutant ion channel variants. The use of ion channels for creating ion-selective membranes has many industrial applications, making this of great importance to study.

**Material and methods**

**Cloning and construction of expression plasmids**

*S. cerevisiae* strain PAP1500(α ura3-52trp1::GAL10-GAL4 lys2-801 leu2Δ1 his3Δ200 pep4::HIS3prb1Δ1.6R can1 GAL)[19] was used throughout the present study.

PAP1500 cells were transformed according to Gietz el al [26] and selected in SD medium containing 2% glucose and 30 mg/liter leucine and 20 mg/liter lysine.

cDNA was PCR amplified with Accupol DNA polymerase (Amplicon, Denmark) and PCR primers and confirmed to give a single band by agarose gel electrophoresis. Expression plasmid pPAP2259 [19] was linearized by treating with Fast Digest versions of *Sal*I, *Bam*HI, *Hin*III (Thermo Scientific, USA), for 2-3h with and subsequently heat-inactivated at 80°C for 10 min.
Expression plasmids were generated by \textit{in-vivo} homologues recombination by transforming the PCR fragments, the standard GFP PCR fragment and the digested vector into \textit{S. cerevisiae} strain PAP1500\cite{REF}. Correct assembly of the expression plasmid was confirmed by growing and inducing transformants in 3 ml SD media supplemented with 2\% galactose as described (REF). \textit{In-vivo} GFP fluorescence was visualized by 1000 times magnification in a Nikon Eclipse E600 fluorescence microscope with an Optronics Magnafire camera.

All plasmids were designed to express an ion channel with a C-terminal standard TEV-GFP-His$_{10}$ fusion, using the same standard GFP PCR fragment amplified with primers GFPfw and GFPrv (listed below). The GFP forward (fw) primer contains a 5´-overhang (AAATTGACTTTGAAAATACAAATTTTC) encoding a TEV cleavage-site and a 5´-GFP specific sequence, the GFP reverse (rv) primer contains a 5´-overhang of 35 bp reverse complementary to the pEMBLyex4 sequence for homologous recombination, TCA (reverse complement stop-codon), 10xHis, reverse complementary 3´GFP sequence.

The design was made so, that the first forward PCR primer for each protein, contains a 35 bp long 5´-overhang for homologous recombination with pEMBLyex4, a Kozak sequence from the yeast PMR1 gene, and a template specific part. In addition the last reverse PCR primer for each protein, carries a 5´-overhang that is inverse complementary to the TEV cleavage site, plus a template specific part that is inverse complementary to the 3´end of the protein encoding cDNA sequence.

Forward and reverse oligonucleotides were used to introduce mutations by mismatching the PCR primer sequence at the point of mutation.

List of the sequence of PCR primers used: purple sequences encodes the TEV protease site used for homologous recombination; the red sequences generate a homologous recombination site with the expression plasmid; the grey sequence is the Kozak sequence from the yeast PMR1 gene; underscored sequences points of mutagenesis; the orange sequence is for recombination with the expression plasmid; the turkish sequence encodes a His$_{10}$ tag.; template specific sequences are depicted without shading; dark blue generate a recombination site between two parts of cDNA.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Forward primer (5’-3’)</th>
<th>Reverse primer (5’-3’)</th>
</tr>
</thead>
</table>
| Standard TEV-GFP-His<sub>10</sub> | CAGATGCGAGTTGAACAGACACCAAACTTACCATGCACTTTGCAAGAGAATTATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA...
<table>
<thead>
<tr>
<th>Site</th>
<th>Sequence Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>E166A</td>
<td>CATTGGGTAAACAGGGTCCATTTTG</td>
</tr>
<tr>
<td>I223W</td>
<td>GTTTTGTITTTCTACGGTTACATG</td>
</tr>
<tr>
<td>I490W</td>
<td>CAAATCTCACATACTTTGGCATG</td>
</tr>
<tr>
<td>CLC-0/1-532</td>
<td>ACACAAATACACACAAATTACCCGATCA</td>
</tr>
<tr>
<td>CLC-1a WT</td>
<td>ACACAAATACACACAAATTACCCGATCA</td>
</tr>
<tr>
<td>CLC-1b WT</td>
<td>AAAATGCACTTGGAAAATACAAATTTTC</td>
</tr>
<tr>
<td>E232A</td>
<td>CAGTGGTAAAGCAGGTCCATTG</td>
</tr>
<tr>
<td>I290W</td>
<td>GTTTTTTTTCAATCAGATTTCT</td>
</tr>
<tr>
<td>I556W</td>
<td>AAAATCGCTCATACCTTTGGCAATGAT</td>
</tr>
<tr>
<td>CLC-1/1-598</td>
<td>ACACAAATACACACAAATTACCCGATCA</td>
</tr>
<tr>
<td>CLIce1 WT</td>
<td>ACACAAATACACACAAATTACCCGATCA</td>
</tr>
<tr>
<td>E148A</td>
<td>TTATAGGTAGAGGGTCCACAG</td>
</tr>
<tr>
<td>I201W</td>
<td>TAACCTTGTTTATTGGGAAGAAATGAG</td>
</tr>
<tr>
<td>I422W</td>
<td>ACTACCAATTGTGGTTGCAATGAT</td>
</tr>
<tr>
<td>E205A</td>
<td>GTTATTATAGAAGGAGAGGCA</td>
</tr>
<tr>
<td>Y445A</td>
<td>GTAACCATTGGGCCTCTGCTATTTTA</td>
</tr>
</tbody>
</table>
Expression screening of mutant constructs

Yeast cells were inoculated in SD medium containing lysine and leucine and grown at 30°C for 2-3 days. 200µl was then transferred to 5ml SD containing lysine and grown at 30°C for 1 day. 50ml YPG (YP medium with 0.5% glucose and 3.5% glycerol) medium was inoculated with culture from the previous medium resulting in an OD₄₅₀=0.05. The cells were incubated at room temperature until OD₄₅₀=3 and were then moved to 15°C. After 15 min incubation the ion channel expression was induced by addition of 2% galactose induction media (2,5ml 40% galactose and 2,5ml 2xYP with 3% glycerol). The cultures were measured for GFP expression normalized to 1 OD₄₅₀ unit in 24-hour intervals.

Expression in 2L cultures and membrane preparation

Yeast cells were inoculated in SD medium containing lysine and leucine and grown at 30°C for 2-3 days. 200µl was then transferred to 5ml SD medium containing lysine and grown at 30°C for 1 day. 500 µl of the culture was used to inoculate 50ml of the same medium and grown for 1 day at 30°C. The culture was inoculated to OD₄₅₀=0.05 in 2 L of YPG (YP medium with 0.5% glucose and 3% glycerol) medium. The cell culture was grown at room temperature until OD₄₅₀=3 and subsequently moved to 15°C. After 15 minutes incubation the ion channel expression was induced by addition of 2% galactose induction media (100ml 40% galactose and 100ml 2xYP containing 3% glycerol). Cells were then harvested at the optimal expression time-point, determined by the previous expression screen, by centrifugation at 1000 g (Sorvall SS-34 rotor) for 10 minutes. Harvested yeast cells were stored at -80 °C until further use.

Cell pellets from 2L cultures were thawed and re-suspended in ice-cold lysis buffer (25mM Imidazol, 1 mM EGTA, 1 mM EDTA, 10% Glycerol, 1M NaCl; pH 7.5). Cells were disrupted by vigorous mixing using glass beads in cold lysis buffer (25mM imidazole, 1mM EGTA, 1mM EDTA, 10% glycerol) including 1mM PMSF and 1µg/ml leupeptin, pepstatin and chymostatin (LPC). Cell lysates were collected and the glass beads were washed two times in 15 ml lysis buffer. Cell debris was removed by centrifugation (1000 g, 10 minutes, 4°C - Sorvall SS-34 rotor). Crude membranes were pelleted by ultracentrifugation (160,000 g, 4°C, 90 min – Sorvall T865 rotor). The pelleted membranes were re-suspended in 6ml ice-cold lysis buffer containing 1mM PMSF and 1µg/ml LPC. Crude membranes were kept at -80°C until further purification.
Protein purification

Crude membranes were solubilized in 1% detergent overnight at 4°C in solubilization buffer (15mM Tris-HCl; pH7.5, 0.5M NaCl and 10% glycerol) with protease inhibitors and detergent. Unsolubilized material was removed by ultracentrifugation (160,000 g, 4°C, 20 min – Sorvall T865 rotor). GFP was measured before and after ultracentrifugation.

Protein purification was carried out on an ÄKTA explorer system. The supernatant was diluted 4 times prior to loading onto three sequential 1ml HisTrap FF Nickel sepharose columns (GE Healthcare), which had previously been equilibrated with IMAC running buffer (20 mM Tris-HCl; pH 7.5, 150 mM NaCl and 2xCMC of detergent). Bound protein was eluted in running buffer with either a linear gradient (0mM-700mM) or a step-gradient (40 mM, 500 mM, 1000 mM) in 15 column volumes (CV). The elution was collected in fractions of 1 ml and GFP was measured.

Fractions containing the protein were pooled and concentrated using a viva spin filter with MWCO 50kDa (Sartorius, Germany). The concentrated samples were loaded onto an Enrich SEC650 10/300 column (Biorad, USA) equilibrated with SEC buffer (20mM Tris-HCL; pH 7.5, 150 mM NaCl, 10% glycerol and 2xCMC of detergent) with. Fractions were collected and GFP fluorescence was measured. The main elution fractions were pooled and concentrated to 2 mg/ml, as previously described, an stored at -80 °C until further use.

SDS-PAGE analysis

Protein samples were denatured with SDS sample buffer (xx) at room temperature for 20 minutes before loading. Protein samples were analyzed by SDS-PAGE (12%) according to Laemmli, U. K. (1970)[27] and stained with Coomassie-blue R250. In-gel fluorescence and Coomassie Blue stain imaging was done using an Image Quant LAS-4000 (GE Healthcare, USA).
References

Figure 1. Topology and structure of KcsA and NavAb with positions of mutations. A,B. Topology of KcsA and NavAb monomers showing key features; (P-loop), pore-loop; (VSD), voltage-sensing domain. Transmembrane helices are numbered sequentially. C,D. Tetrameric structure of KcsA and NavAb shown as side-view and top-view, with the location of mutations indicated in yellow spheres and sequence removed indicated in red.
Figure 2. Topology and structure of chloride channels with positions of mutations. A. Topology of the ClC-ec1 monomer. Transmembrane helices are numbered A-R. C. The dimeric structure of ClC-ec1 with the monomers colored cyan and purple, shown as side-view and top-view. Mutations are indicated by yellow spheres. B. Topology representative of ClC-1 and ClC-0 monomer showing key features. Transmembrane helices are numbered A-R and the Systathionine-β-synthase (CBS) domains 1 and 2 are indicated in red. D. The dimeric structure of ClC-1 shown as side-view and top-view, representative also for ClC-0, with monomers colored in gray and orange. Mutations are indicated by yellow spheres and the removed CBS domains are colored red. Note that the size of the CBS domains varies between the two chloride channels.
Figure 3. Mutant ion channels locate to the membranes in *S. cerevisiae*. Live-cell bioimaging of *S. cerevisiae* cells expressing the ion channel variants indicated above the images. Yeast cells were grown at 15 °C and induced with galactose until fluorescence was visible (24-72h). Images are shown in pairs of differential interference contrast (DIC) image and GFP fluorescence.
Figure 4. Expression of ion channel mutants in *S. cerevisiae*. Fluorescence of ion channel mutants A; ClC-0, B; ClC-1, C; ClC-ec1, D; KcsA, E; NavAb, in *S. cerevisiae* cells after galactose induction. Cells were grown at 15 °C and fluorescence was measured in 1 OD₄₅₀ and plotted as a function of time. All proteins were expressed as protein-TEV-GFP-His₁₀ fusions. Colors corresponding to each mutation indicated in each plot (A-E).
Figure 5. Immobilized metal affinity chromatography (IMAC) purification of double and triple mutant ClC-ec1. The crude membranes were solubilized overnight in the detergent indicated above each chromatogram. Un-solubilized material was removed by ultracentrifugation, and diluted four times to lower detergent concentration, followed by loading on an IMAC column using the sample pump on an Äkta explorer system. The chromatograms (A-D) show UV$_{280}$ for the protein elution with the imidazole gradient used (dashed line). The eluted protein was collected in fractions. SDS-PAGE analysis of top fraction visualized by Coomassie-stain (1) and in-gel fluorescence (2).
Figure 6. Size-exclusion chromatography (SEC) of purified double and triple mutant CIC-ec1. The immobilized metal ion affinity chromatography (IMAC) purified protein samples were concentrated to 1 ml and separated by SEC on a Bio-Rad column. The chromatograms (A-D) show UV$_{280}$ of the protein elution, plotted over the elution volume. The eluted protein was collected in fractions. SDS-PAGE analysis of top fraction visualized by Coomassie-stain (1) and in-gel fluorescence (2).
Supplementary figure 1. Mutant ion channels show membrane localization in *S. cerevisiae*. Yeast cells were grown at 15 °C followed by galactose induction. Live cell bio-imaging images of *S. cerevisiae* cells expressing the ion channel variants indicated above the images. Images are shown in pairs of differential interference contrast (DIC) image and GFP fluorescence.
Supplementary figure 2. Detergent screen of wild-type ClC-ec1. Crude membranes with wild-type ClCec1-TEV-GFP-His$_{10}$ were solubilized for 1 hour in FC-12, n-dodecylphosphocholine; FC-13, n-Tridecylphosphocholine; LDAO, Lauryldimethylamine N-oxide; Cymal-5, 5-cyclohexyl-1-pentyl-β-D-maltoside; DDM, n-dodecyl-β-D-maltopyranoside; DM, n-decyl-β-D-maltopyranoside, in a protein:detergent ratio of 1:3. A. Solubilization efficiency was estimated as fluorescence in the supernatant after ultracentrifugation normalized to the total fluorescence in the crude membranes for each detergent (left, diagrams). B. The detergent solubilized crude membranes were separated by size exclusion chromatography and GFP fluorescence was monitored during the elution (right, chromatogram). The colors of the chromatograms (right) corresponds to the colors of each detergent in the bar diagram (left).
Supplementary figure 3. GFP profiles of protein elution during Immobilized metal affinity chromatography (IMAC) and Size-exclusion chromatography (SEC). During the IMAC and SEC purifications, the eluted protein was collected in fractions and GFP fluorescence was measured in each fraction. The profiles from IMAC and SEC are showed as depicted above, for: A,B. ClC-ec1 (E148+Y445A) in LDAO; C,D. ClC-ec1(E148A+E203A+Y445A) in LDAO; E,F.
ClC-ec1 (E148+Y445A) in DM; G,H. ClC-ec1(E148A+E203A+Y445A) in DM.

Supplementary figure 4. Purification of ClC-ec1 mutant variants. Pictures taken during the purification procedure; A. isolated crude membranes, B.IMAC column with bound protein, C. Eluted fractions, right tube with turbid aggregation.
CHAPTER 5 – Journal Article I

Synopsis

With the intend of creating a constitutively open chloride channel we encountered the bias, that while most functional work has been carried out on Cl-channels, all structural information on chloride channels (ClCs) were from homology models based on small bacterial Cl/H+ transporters up until 2017, where the first bovine channel structure was determined [25], closely followed by the structure of the human ClC-1 channel described in this manuscript [26, 27].

Initial screening of both ClC-1 and ClC-0, to find the optimal detergent and purification protocol for structure determination by Cryo-EM, revealed that ClC-1 behaved much better during purification than ClC-0. This was a surprise to us, based on the fact that ClC-1 has a larger cytosolic domain and had a much lower expression yield (see Supplementary data ClC-1 and ClC-0).

The paper presents the structure of the human ClC-1 channel, hereby providing structural knowledge on the channel type of ClCs and the chloride conducting mechanism. In addition the paper provides evidence for regulation via a so-called systathionine-β-synthase (CBS) by pH and adenine nucleotides.
Structure of the human CIC-1 chloride channel

This Journal article can be found in the back of the thesis

Abstract

CIC-1 protein channels facilitate rapid passage of chloride ions across cellular membranes, thereby orchestrating skeletal muscle excitability. Malfunction of CIC-1 is associated with myotonia congenita, a disease impairing muscle relaxation. Here, we present the cryo-electron microscopy (cryo-EM) structure of human CIC-1, uncovering an architecture reminiscent of that of bovine CIC-K and CLC transporters. The chloride conducting pathway exhibits distinct features, including a central glutamate residue (“fast gate”) known to confer voltage-dependence (a mechanistic feature not present in CIC-K), linked to a somewhat rearranged central tyrosine and a narrower aperture of the pore toward the extracellular vestibule. These characteristics agree with the lower chloride flux of CIC-1 compared with CIC-K and enable us to propose a model for chloride passage in voltage-dependent CLC channels. Comparison of structures derived from protein studied in different experimental conditions supports the notion that pH and adenine nucleotides regulate CIC-1 through interactions between the so-called cystathionine-β-synthase (CBS) domains and the intracellular vestibule (“slow gating”). The structure also provides a framework for analysis of mutations causing myotonia congenita and reveals a striking correlation between mutated residues and the phenotypic effect on voltage gating, opening avenues for rational design of therapies against CIC-1-related diseases.

Author summary

Chloride transporting CLC proteins are expressed in a wide range of organisms, and the family encompasses several members with numerous roles in human health and disease
CHAPTER 6 – Discussion and Outlook

The overall aim of this thesis was to engineer ion channels to be constitutively open and to establish a platform for high yield production of these membrane proteins. This would be done with the later intention to explore the idea of developing new biomimetic membranes selective only to monovalent cations and anions resulting for the generating sustainable energy by RED

Production of membrane proteins in large amounts requires a high yield expression systems, easy cloning and manipulation with expression constructs and high throughput screening.

Currently *E. coli* is the preferred expression host for bacterial membrane proteins [32, 33]. However, our comparison of the expression capacity for bacterial membrane proteins in *E. coli* and *S. cerevisiae*, revealed a higher expression yield in *S. cerevisiae*, as described in Chapter 2. These proteins could also be purified to a high degree, and the two ion channels, KcsA and NavAb, were shown to be active.

Yeast as a recombinant expression host has the advantage of a more advanced eukaryotic protein processing machinery; folding, assembly and post-translational modifications (PTMs) and avoiding expression of the membrane proteins in inclusion bodies would be another advantage yeast. In Manuscript I and III we do not see any indications of PTMs on the purified membrane proteins, which would be seen as several protein bands or a smear after separation by SDS-PAGE. However, yeast might introduce glycosylations or other PTMs to recombinant proteins. Unwanted PTMs could be avoided by reengineering of the protein sequence to remove any PTM target sequences. Challenges of using yeast as an expression organism over *E. coli*, might be the longer generation time and a tougher cell wall, which implies longer cultivation times and making it more difficult to lyse the cells. As shown in this thesis, the higher yield of membrane protein would however out-weigh these limitations.

Throughout the thesis it was demonstrated that creating recombinant ion channels as –TEV-GFP-His10 fusions, enabled easy screening of expression, stability in detergents and purification yields, by measuring the GFP fluorescence. Utilizing the GFP-tag by in-gel fluorescence in SDS-PAGE analysis circumvents the need for Weston blotting for direct protein detection. It is
beneficial to avoid this extra protein detection step is, since is time consuming and costly, and requires protein- or tag-specific antibodies for detection.

In addition the poly-histidines enabled the use of a Ni-affinity based purification procedure, resulting in pure protein in a single purification step.

**Potential for industrial scale production of ion channels**

There are many considerations for scaling up production of the ion channels described in this thesis. Basically it will often be a tradeoff of between cost and yield of the purification. When choosing the method for purifying the proteins, the main cost lies in the reagents, time and equipment needed, which all need to be compatible with large scale purifications. In this regard, as discussed in Manuscript III, it is also important to take the critical micellar concentration (CMC) value of the detergent into account prior to up scaling purifications. The detergent solubilized ion channels need to be reconstituted into polymers, when the application is to create ion selective membranes. Detergents with a higher CMC value will be easier to remove from the transmembrane part of the protein during reconstitution, however the higher CMC also means that a higher concentration of the detergent is needed, hence increasing the price of the purification. At the same time it is essential the detergent can keep the protein stable in solution during the purification. In Manuscript III, purification in LDAO led to aggregation of the protein. So the decision on which detergent to be used for large scale purifications should be made from weighing out the pros and cons for these many factors.

We demonstrated that our purification procedure yielded several milligrams of highly pure protein from a single culture flask. This proves the potential for purification of hundreds of milligrams from cell culturing in bioreactors. Using bioreactors enables control of expression condition and oxygen supply to the cells, which results in higher cells densities, hence a higher yield of membrane protein pr. L culture.

In general a minimum of 3 purification steps are needed for production of biopharmaceuticals [34], however for our purpose of creating ion selective membranes we do not need this extreme purity. A single purification step using an automated IMAC procedure resulted in pure protein samples in Manuscript III. Using only a single step of purification will decrease the time and cost
of the purification and it will in general make it much easier to scale up the purification when only using resin-based affinity columns.

We did not remove the GFP-tag in this thesis work. However for creating ion selective membranes, it might be desirable to remove the GFP from the ion channels. This might especially be needed for the small and compact channels that do not have many residues outside of the membrane. The recombinant constructs describe in this thesis (Manuscript I, II, III and journal article I) were created with a Tobacco etch virus (TEV) site between the protein and the GFP-His\textsubscript{10} tag, for tag removal by specific cleavage using the TEV protease [35]. This tag removal would traditionally be done by cleaving the IMAC purified protein with a His-tagged TEV protease. The cleaved tag and the TEV protease can then be removed from the protein sample by a second IMAC step [36]. Yet, for industrial scale purification it might be more beneficial to create constructs only carrying the His-tag, after establishing the purification procedure using the constructs containing a GFP tag.

**Successful structure determination using cryo-EM**

There are no golden rules when it comes to pre-determine which MP targets will yield a pure and stable protein sample. This is exemplified by the screening of ClC-1 and ClC-0 prior to determining the structure by cryo-EM (Journal article I), as described in Chapter 5. Therefore it is advisable to screen targets for expression, detergents and purification yields prior to large scale production. As mentioned earlier, the GFP tag is a great tool for screening, since it is an extremely sensitive and fast way of detecting the recombinant protein, and can be measured on a µl scale, saving precious membrane protein samples.

High yields of membrane proteins are needed both for bio-technological applications and academic research for structural and functional analysis. We were able to determine the structure of ClC-1 using cryo-EM. This was made possible by the high yield expression system, but also by the technical advances within the cryo-EM research field the last years. Incredible numbers of cryo-EM structures are being published [37], and an increase in new microscope facilities, encourage laboratories to attempt solving structures using this method [38]. Even though the technical advanced help determine the structure of proteins by cryo-EM, the most important is still to have a good protein sample, to "know your biochemistry". This is especially true for membrane proteins, where stability is in general lower than their soluble counterparts, and
detergents are needed to keep the protein in solution. Using too much detergent risks inducing denaturing of protein-protein interactions, while too little will aggregate the membrane protein. It is also advisable to screen different purification conditions to increase stability e.g. surfactants, pH or salt concentrations. All screening are still done manually by trial and error, which can be extremely labor some and costly, both due to materials and the need for cryo-EM microscopes for screening. It would therefore be of great use to be able to screen samples on lab-scale microscopes prior to taking the sample to an EM facility. But until then; it is of great importance to have deep understanding of how your protein behaves and to obtain the best sample quality before proceeding to cryo-EM.

**Opportunities for engineering constitutively open ion channels and the basis of ion selectivity**

We demonstrated the use of homologous recombination for constructing recombinant protein expression vectors in Manuscript II. In addition this showed to be a neat method for creating an array of mutation variants. In this way we created all the mutations of the ion channels described in Manuscript III, and we also confirmed that all could be expressed in *S. cerevisiae*. We chose to base our mutations on previously established data on the ion channels, which saved us a lot of time, not having to do extensive mutational studies prior to the work in this thesis. However, measurements on the ion channels varied a lot depending on the method used and the experimental setup, as an example mentioned in Manuscript I, the NaChBac has been reported with conductance from $\sim 12\text{ps}[39]$ to $120\text{ps}[40]$. For our purpose of creating constitutively open ion channels, the exact values of the channel conductance is not important. These channels would be used at a zero membrane potential and at a high ion gradient, which would in any case open the channels. In this sense it is smart to use of pre-established data for the function and effect of specific mutations, and focus on the ability to produce these ion channels in high yields in a stable form.

In this thesis work we chose an array of different mutations to remove gating and different parts residing outside the membrane; towards creating constitutively open ion channels or ion pores. The selected mutations and general structure of the ion channels in this thesis work is describe in Chapter 1 and Manuscript III, and the selected mutations were based on previously established knowledge and mutational analyses, as mentioned previously. Elucidating the exact functional and structural properties for ion selectivity of ion channels is of upmost importance for designing
ion channels with a desired function. Electrophysiological measurements and structure
determination traditionally formed the main foundation for understanding the biological function
of ion channels. Molecular dynamics have now accelerated the understanding of facilitated ion
passage across membranes, with many new studies published in recent years on the ion channels
used in this thesis [41-46]. MD simulates the atomic motions in a system and can complement
the knowledge from 3D structures and functional data [47]. It makes it possible to see events on
time-scales not possible to detect from experimental methods, hereby obtain a deeper
understanding of the atomic mechanistic of how ion channels convey selective ion flux across
biological membranes.

We now begin to understand how selectivity of ion channels arises, based on atomic structures
and functional data. Basically selectivity of ion channels arises from the pore size, pore lining
and the interaction the ion. Ion selectivity in NavAb arises from side chains inside the pore, while
it is mediated by the backbone in KcsA [48]. In addition, it has been known for many years, that
K+ ions pass through potassium channels in a fully dehydrated while Na+ ions pass through
sodium channels partly hydrated, and both types of channels have a knock on mechanism where
the ions push the next one through in a single file [49]. Still MD simulations have contributed
with new knowledge on small changes in the selectivity filter and the degree of hydration in the
NavAb channel [23, 43]. In contrast, chloride channels the selectivity filter is located at the
narrowest point in the middle of the hour-glass shaped ion passage. This selectivity arise from
the width of the filter and side-chain interaction with the Cl\(^-\) ions [23].

Knowledge of molecular basis for selectivity in ion channels, can be used design channels *de
novo* with desired properties for transport of a specific molecule of interest. It was possible to
turn NavAb into a Ca\(^{2+}\) selective channel, showing insights to how ion selectivity is provided by
the protein sequence and structure of the ion passage. Three point mutations inserting extra
negative charge tuned the channel into a highly selective Ca\(^{2+}\) [50]. This opens great
opportunities for biotechnological applications in combination with modern membrane
technologies, to create selective biomimetic membranes for any desired molecule. This could for
example be used either for removal of an unwanted molecule from a solution or for isolating a
wanted molecule from a solution.
Future opportunities for ion channels in biomimetic membranes

As described in Chapter 1, biomimetic membranes can be made in different ways to give various functions of separation and permeation. This novel technology still has a number of challenges associated; poor control over polymer, and limited understanding of interactions between functional molecules and the membrane materials; difficulties in upscaling the production of biomimetic materials and the cost associated with this; long term stability and reusability [11, 51]. By engineering ion channels, as described in Manuscript III or de novo designing ion channels for the purpose, it would be possible to enhance the specific permeation of ions, hereby overcoming one of the current challenges of biomimetic membranes with selectivity. The intention of creating simple ion pores spanning the membrane would also be to increase the long term stability of the protein inside the polymer matrix.

Biomimetic membrane technology have great future potential and many uses e.g. water treatment or biomedical drug delivery, molecular bio-sensing [11], as well as energy conversion like RED described in chapter 1.
CHAPTER 7- Conclusion

The work described in this thesis describes the process of design and production of constitutively open ion channels for creating biomimetic membranes.

After identifying the target channels and mutations, we utilized a brilliant cloning method for easily introducing the mutations and manipulating expression constructs. This neat way of introducing mutations by homologous recombination is of great importance for creating array of proteins with modifications in structure and function, cloning work that would otherwise be costly and time-demanding with traditional cloning methods.

Secondly we tested the expression capacity of *E. coli* and *S. cerevisiae*, which showed that yeast could attain higher expression levels of the selected bacterial ion channels.

Thirdly we were able to use the *S. cerevisiae* expression platform for high-yield protein production and establish an automated, robust and scalable purification procedure.

From the discussion in Chapter 6 it was clear that there are many challenges of producing pure and stable membrane proteins and many considerations to take into account before upscaling the quantities.

The future perspectives of this research are bright and exciting. To embed these membrane proteins into polymer membranes and begin unraveling the many uses for this genius way of separating solution, while allowing passage of a specific molecule of interest. All inspired by the brilliance of nature.
REFERENCES


Almen, M.S., et al., *Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin*. Bmc Biology, 2009. 7.


Supplementary figure 1. A. Expression of CIC-0 (Red) and CIC-1 (Blue) in 15 L Bioreactors. Growth was carried out in expression media with 2% glucose and supplemented with amino acids (except leucine). The cultures were induced with galactose when all glucose was used and expression was followed in 1 OD_{450} over time. B,C. Solubilization efficiency of the detergents listed with or without cholesterol (CHS). Crude membranes were solubilized for 1 hour in detergent, followed by ultracentrifugation, % was determined from the fluorescence in the sample before and after ultracentrifugation.
Supplementary figure 2. Fluorescence-detection size-exclusion chromatography (FSEC) of A. ClC-0 and B. ClC-1, solubilized in the detergent indicated above, with and without cholesterol (CHS).
Supplementary figure 3. Sequence alignment of CIC-1 and CIC-0, showing high degree of similarity and larger CBS domains of CIC-1.
Structure of the human ClC-1 chloride channel

Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Yanxiang Cui, Julie Winkel Missel, Christina Grønberg, Kamil Gotfryd, Erik Lindahl, Magnus Andersson, Kirstine Calloe, Pascal F. Egea, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon

1 Department of Biomedicinal Sciences, University of Copenhagen, Copenhagen, Denmark, 2 Department of Microbiology, Immunology & Molecular Genetics, University of California at Los Angeles, Los Angeles, California, 3 California NanoSystems Institute, University of California at Los Angeles, Los Angeles, California, 4 Department of Biology, University of Copenhagen, Copenhagen, Denmark, 5 Department of Biochemistry & Biophysics, Stockholm University, Stockholm, Sweden, 6 Department of Chemistry, Umeå University, Umeå, Sweden, 7 Department of Veterinary and Animal Sciences, University of Copenhagen, Frederiksborg, Denmark, 8 Department of Biological Chemistry, University of California at Los Angeles, Los Angeles, California, 9 Institute of Biophysics, Consiglio Nazionale delle Ricerche, Genova, Italy, 10 Department of Experimental Medical Science, Lund University, Lund, Sweden

Abstract

ClC-1 protein channels facilitate rapid passage of chloride ions across cellular membranes, thereby orchestrating skeletal muscle excitability. Malfunction of ClC-1 is associated with myotonia congenita, a disease impairing muscle relaxation. Here, we present the cryo-electron microscopy (cryo-EM) structure of human ClC-1, uncovering an architecture reminiscent of that of bovine ClC-K and CLC transporters. The chloride conducting pathway exhibits distinct features, including a central glutamate residue (“fast gate”) known to confer voltage-dependence (a mechanistic feature not present in ClC-K), linked to a somewhat rearranged central tyrosine and a narrower aperture of the pore toward the extracellular vestibule. These characteristics agree with the lower chloride flux of ClC-1 compared with ClC-K and enable us to propose a model for chloride passage in voltage-dependent CLC channels. Comparison of structures derived from protein studied in different experimental conditions supports the notion that pH and adenine nucleotides regulate ClC-1 through interactions between the so-called cystathionine-β-synthase (CBS) domains and the intracellular vestibule (“slow gating”). The structure also provides a framework for analysis of mutations causing myotonia congenita and reveals a striking correlation between mutated residues and the phenotypic effect on voltage gating, opening avenues for rational design of therapies against ClC-1–related diseases.

Author summary

Chloride transporting CLC proteins are expressed in a wide range of organisms, and the family encompasses several members with numerous roles in human health and disease...
by allowing movement of chloride ions across the membranes that encapsulate cells and cellular organelles. Structurally, CLCs form dimers possessing a separate ion translocation pathway in each monomer, and they can operate as either channels or transporters that exchange chloride for protons. The CLC channel CIC-1 is critical to skeletal muscle excitability and has been proposed as a target to alleviate neuromuscular disorders. Here, we have analyzed the structure of human CIC-1 and revealed the high similarity of its ion conducting pathway to those observed in other CLC members, including prokaryotic and algal transporters. Our data suggest how CIC-1 is regulated by environmental cues to allow opening and closure, thereby permitting attenuation of muscle function. Our results help with understanding the principal determinants that govern CLC proteins and may guide downstream translational applications to combat muscle pathologies.

Introduction

CLC proteins comprise a large family of chloride (Cl\(^-\))-transporting integral membrane proteins with diverse physiological functions [1–3]. The first identified human member, CIC-1, is essential for maintaining the permeability of Cl\(^-\) across the plasma membrane of skeletal muscle fibers, \(g_{Cl}\), accounting for approximately 80% of the resting membrane conductance and assuring precise neuronal control of muscle contraction [3]. Mutations of the CIC-1 gene cause myotonia congenita, a disease that allows a single nerve action potential to trigger a series of muscle action potentials (myotonic runs), leading to prolonged muscle contraction [4–7].

Despite distinct roles as passively conducting Cl\(^-\) channels and stoichiometrically coupled secondary active Cl\(^-\)/H\(^+\) antiporters [2, 3], members of the CLC family share a common homodimeric core architecture, with each subunit harboring an independent ion translocation pathway [8, 9]. The molecular mechanisms of ion transport in CLC antiporters have been extensively studied functionally and structurally [8, 10–15]. Yet it is poorly understood how the antiporters and channels establish their separate functions. In addition, the complex gating processes that regulate CLC channel activity remain elusive, with only a single available structure of a channel member, namely, that of bovine ClC-K [9]. Each CLC monomer has a gate that operates independently from the other (also known as “protopore” or “fast gate”), structurally attributed to a specific glutamate, “Glut\(_{GATE}\)” [10]. A slower gate controls both conducting pathways simultaneously (“common” or “slow gate”) [16], but the principles and determinants of this regulation are enigmatic. Furthermore, activity of CIC-1 is modulated by cellular cues such as phosphorylation [17], pH, and nucleotides [18, 19] in an unknown manner. Such regulation is, however, physiologically essential because intense muscle exercise leads to acidosis, resulting in an increased nucleotide sensitivity of CIC-1 and consequent reduction of \(g_{Cl}\), thereby assisting in preventing muscle fatigue [20, 21].

The recent ClC-K structure provided the first insights into the differences between CLC channels and transporters; in particular, it revealed a pore widening on the intracellular side. Yet there are surprisingly few known structural differences between the CLC channels and transporters. However, ClC-K channels exhibit only limited gating as Glut\(_{GATE}\) is missing [2, 3], and their activity has not been reported to depend on nucleotide binding [22]. Therefore key questions concerning CLC channel function and regulation remain unanswered. Furthermore, a deeper understanding on structure–phenotype relationships of myotonia-causing mutations in CIC-1 is required to shed further light on how the muscle disease is manifested at a molecular level.
Results

Here, we have determined structures of full-length human ClC-1 using single-particle cryo-electron microscopy (cryo-EM), exploiting a purified protein sample that displays Cl\(^{-}\)-dependent single-channel-derived ion conductance (S1 Fig and S1 Data). For structural characterization, sample in the presence of 100 mM Cl\(^{-}\) at pH 7.5 and in the absence of nucleotides or antibodies was initially employed (Fig 1). Three-dimensional (3D) classification of particles resulted in several different groups, of which one yielded a 3.6 Å overall resolution density map for the transmembrane domain, allowing confident model building (S2–S4 Figs). The final model represents the membrane-spanning portion (note that the N terminus and intracellular αA helix are lacking) as well as parts of two C terminal’s so-called cystathionine-β-synthase (CBS) domains present per monomer (for which some cryo-EM density is left unmodeled) and includes several features that were not observed in the ClC-K structure (S5 Fig).

The homodimeric architecture of ClC-1 is reminiscent of that of bovine ClC-K and available structures of CLC proteins from lower organisms (Fig 2A). The monomers consist of membrane-spanning helices and half-helices (αB to αR) with connecting loops (e.g., αB–C, between αB and αC) as well as the CBS domains (Fig 1). Each protomer holds a separate...
Fig 2. The ion-conducting pathway. Ion transport in CLC proteins depends on extracellular and intracellular vestibules and a connecting pore. In CLC transporters, the pore is marked by chloride ion binding sites (s<sub>ext</sub>, s<sub>cen</sub>, s<sub>int</sub>; not directly observed in this work) as well as specific glutamate (Glu<sub>GATE</sub>, or E232; ClC-1 numbering throughout), tyrosine (Tyr<sub>C</sub>, Y578), and serine (Ser<sub>C</sub>, S189) residues. Chloride conductance in voltage-dependent CLC channels such as ClC-1 may involve shuttling (i) to protonated E232-Y578 (s<sub>cen</sub>) from the vestibules - directly (or through a weak s<sub>int</sub>) from the intracellular side and (ii) through K231/R421 to overcome the hydrophobic barrier (including M485) from the extracellular side. (a) Comparison of the transmembrane domains of ClC-1 (colored as in Fig 1D) and ClC-K (gray), respectively. Helices are labeled with white letters throughout. (b) Schematic overview of the chloride permeation pathway with key residues pinpointed. Labels in the parentheses refer to the corresponding helices and the αC–D loop, respectively. (c–f) Side views of the pore region of ClC-1 (panels c and e; colored as in Fig 1D) with equivalent views of ClC-K (panels d and f, shown in grayscale) for reference.
chloride conducting pathway across the membrane, established by a vestibule on either side of the membrane, and an interconnecting narrow and short pore.

In CLC transporters, the Cl\textsuperscript{−} conducting pore (Fig 2B) is marked by distinct Cl\textsuperscript{−} binding sites (denoted \(s_{\text{ext}}\), \(s_{\text{cen}}\), and \(s_{\text{int}}\), respectively, but no Cl\textsuperscript{−} ions are resolved in the current structure), and the constricting Glu232 (of \(\alpha F\), also known as Glu\textsubscript{GATE}; CIC-1 numbering throughout) and Tyr578 (of \(\alpha R\), Tyr\textsubscript{C}) [9]. Furthermore, Ser189 (of \(\alpha C-D\), Ser\textsubscript{C}) is located in the vicinity of the pore (Fig 2B, 2C and 2E and S5A Fig). In CIC-1, voltage-dependent gating is established by Glu\textsubscript{GATE}, which is perhaps being displaced by competing Cl\textsuperscript{−} ions and/or protonation. In contrast, in voltage-independent ClC-K channels, Glu\textsubscript{GATE} is replaced by a valine, and, indeed, substitutions of Glu\textsubscript{GATE} with uncharged residues render CIC-1 similarly voltage independent [24]. Unfortunately, the Glu\textsubscript{GATE} side chain is not visible in our cryo-EM density maps (S4 Fig), but carboxylate groups of interacting acidic residues are known to be frequently undetectable using cryo-EM due to radiation damage. A similar orientation of the side chain as observed in ClC-K would be in agreement with Cl\textsuperscript{−} passage through a maintained \(s_{\text{cen}}\) site, as a concomitant adaptation of \(\alpha R\) significantly shifts the position of Tyr\textsubscript{C} and thus maintains the Glu\textsubscript{GATE}-Tyr\textsubscript{C} distance (Fig 2B–2F and S5G Fig). However, we cannot exclude that the side-chain of Glu\textsubscript{GATE} is buried deep into the hydrophobic pocket established by Phe279, Phe288, and Phe484 (S5H Fig).

The pore aperture of the extracellular vestibule is constricted by a hydrophobic barrier with Met485 (Met427 in ClC-K), but in contrast to ClC-K, the gate opening is also controlled by Lys231 (of \(\alpha E-F\)) and Arg421 (of \(\alpha L\)) (Fig 2B–2F) that may orchestrate Cl\textsuperscript{−} permeation to or from the extracellular environment [25–27]. This difference can be attributed to \(\alpha E-F\), with its Glu\textsubscript{GATE} and Lys231 adopting a more CLC-transporter-like configuration because this loop is considerably shorter than in CIC-K, alongside a side-chain reorientation of Arg421 (Fig 2C–2F). We also observe a structural adjustment on the intracellular side of the pore, with \(\alpha C-D\) being displaced as compared to the corresponding loop in CIC-K. This rearrangement opens the vestibule even deeper toward Glu\textsubscript{GATE} (Fig 2C–2F), providing intracellular access beyond the \(s_{\text{int}}\) site present in CLC antiporters and suggesting that no tight Cl\textsuperscript{−} binding occurs on the intracellular side, in agreement with electrophysiological data [28]. The wider intracellular vestibule of the CLC channels, as compared to the transporters, has been proposed to allow for the higher Cl\textsuperscript{−} conductance in channels, lowering the kinetic barrier between \(s_{\text{cen}}\) and the cytosol [9]. We note that the vestibule width of CIC-1 is similar to that of ClC-K at Ser\textsubscript{C}, with the side chain of this residue being positioned away from the Cl\textsuperscript{−} permeation pathway in both channels, establishing the Ser\textsubscript{C} location as another of the distinguishing features between CLC channels and transporters.

It remains obscure whether the channel has been captured in the open configuration, a priori induced by the experimental conditions (0 mV, 100 mM Cl\textsuperscript{−}). Molecular dynamics simulations of the CIC-1 structure suggest that Cl\textsuperscript{−} from the intracellular side spontaneously interacts with Glu\textsubscript{GATE} upon protonation of its side chain but that free energy is required to complete the passage across the membrane (S6 Fig). We anticipate that Glu\textsubscript{GATE} and the Lys231–Arg421 constricting interactions attenuate chloride flux, in agreement with the smaller conductance of CIC-1 versus ClC-K [2, 3], and we cannot exclude that Cl\textsuperscript{−} shuttling occurs directly between
protonated Glu\textsubscript{GATE} and Lys231 across the Met485 barrier (Glu\textsubscript{GATE} overlays s\textsubscript{ext} in some CLC transporters [8, 14, 29]); chloride interaction with the latter may be unfavorable, however.

The molecular mechanisms that govern slow gating in CLC proteins remain elusive. It is known that CBS nucleotide binding and low pH inhibit CIC-1 activity by favoring closure of the common gate [19, 29]. Assessment of the 3 major cryo-EM maps obtained in our structural classification (see also S2 Fig and Methods) reveals different arrangements of the CBS domains, suggesting intrinsic domain flexibility at pH 7.5 (Fig 3A and 3B and S7 and S8 Figs). To test this, we determined the structure of CIC-1 also at lower pH (6.2) in the presence of 0.3 mM of the nucleotide nicotinamide adenine dinucleotide (NAD) to unravel the regulation mechanism (S2, S3 and S8 Figs). In these conditions, the CBS domains appear significantly more rigid (in comparison to pH 7.5; Fig 3A and 3B and S7 and S8 Figs). This observation is also supported by CIC-1 size-exclusion chromatography profiles (S9 Fig), with samples at low pH being shifted toward lower molecular weight (more compact). Therefore, the CBS arrangements seem to correlate with slow gating, being rigid at low pH in the presence of nucleotides and more flexible at higher pH in the absence of nucleotides, bringing to mind a mechanism that has been proposed based on electrophysiological data [29]. The complete effects of such putative rearrangements are, however, not demonstrated experimentally by our structures, because they remain closed also at the higher pH (determined from particles in detergent environment).

How then can the Cl\textsuperscript{−} conductance of 2 separate pores be affected by structural shifts of the CBS domains? Examination of the interface between the CBS and the transmembrane domain suggests that CBS2 interacts with \(\alpha\)D–E, a loop previously shown to affect slow gating (Fig 3C and 3D) [25, 31]. Nucleotides may also interact directly with the transmembrane domain when bound in the cleft between CBS1 and CBS2 (the latter observed in structures of isolated CBS domains [13]; Fig 3E). It is conceivable that these structural arrangements and the direct physical connection between CBS and \(\alpha\)R—all structural elements leading to the Glu\textsubscript{GATE} constrictions site—allow structural adjustment of the transport pathway and thus chloride conductance regulation (Fig 3C). Such structural effects will be propagated between the monomers via the CBS domains, in agreement with concurrent modulation of the 2 conducting pathways in the dimer [16]. We note that the CBS portions that interact with the transmembrane and the CBS domain of the adjacent monomer are structurally (and at interaction sites also sequencewise; S10 Fig) conserved (Fig 3E), and therefore this may represent a unifying mechanism of slow gating for CLC proteins.

CIC-1 defects cause recessive (Becker type) or dominant (Thomsen type) myotonia congenita, typically associated with complete disruption of channel function or with a dominant negative effect in heterodimeric wild-type (WT)-mutant complexes [7], respectively. Our structure now allows mapping of such (or other experimental) CIC-1 substitutions for evaluation of structure–function–disease and -phenotype relationships (Fig 4). Several dominant and recessive mutations induce an alteration of the overall gating from depolarization to hyperpolarization activated, yielding a similar intracellular Cl\textsuperscript{−}-sensitive gating as described for CIC-2 [32]. Therefore, the different gating profiles of CIC-1 and CIC-2 likely do not necessitate major structural differences. These residues are generally surface exposed and localized to the extracellular half, including the vestibule and the pore-constricting residues Lys231 and Arg421 (Fig 4B) [26, 27, 32–35]. In contrast, many dominant mutations exert a “shift” of the common gate to open probability to positive voltages, leading to significant reduction of \(g_{\text{Cl}}\) at the physiological membrane potential [36]. Such mutations cluster primarily at the dimer interface and in the intracellular vestibule and pore region (Fig 4C, and 4D and S5D Fig). One is located in CBS2, close to the membrane domain, in agreement with the above-mentioned mechanism of slow-gating regulation exerted via CBS2. Residues that affect binding of one of the most commonly used CIC-1 inhibitors, the lipophilic 9-anthracene-carboxylic acid
(9-AC), are all buried into a CAVER [37]-computed membrane-embedded cavity on the intracellular side that stretches to GluGATE, in agreement with the intracellular mechanism of action proposed for this compound (Fig 4E and 4F and S11 Fig) [24]. Because this pocket is lined by multiple hydrophobic and a few negatively charged residues, it is unlikely to allow chloride conductance (proton access is possible) but rather 9-AC–induced interference of flux across GluGATE and may thus represent a suitable site for future drug-discovery efforts.

**Discussion**

In summary, we report the molecular structure of Cl−-conducting human ClC-1, sharing an overall fold similar to other CLC proteins, with a narrow connecting pore and positively charged vestibules attracting Cl− ions similar to CFTR [38]. The structure exhibits several unique features, including shifts in the central GluGATE-TyrC pair, a more closed extracellular vestibule, and a wider penetration profile from the intracellular side, the latter representing a distinct feature of CLC channels separating them from transporters. We propose a model for adenine nucleotide and pH regulation of the common gate via CBS2 and the intracellular loops congruent with previous functional data. Overall, these findings significantly increase our understanding of Cl− conductance in physiology and open new opportunities for biomedicine. For example, the positively charged constriction of the extracellular vestibule and the putative 9-AC pocket may serve as favorable target sites for stimulators or inhibitors from outside or inside the cell, respectively.

During the course of the preparation of this manuscript, the structure of human ClC-1 was reported by another group [39]. The ClC-1 structures display only limited differences despite that different overproduction hosts were exploited. The authors detected a similar putative 9-AC binding pocket (the alternative pathway) and conformational flexibility in the CBS region (determined at pH 7.4), in agreement with our findings. We anticipate that the pH-dependent conformational changes reported here—in conjunction with mutational efforts using, e.g., single-channel recordings, as for the first time demonstrated in this work, will allow for more refined studies to further resolve the mechanism of slow-gating in CLC proteins.

**Methods**

**Recombinant expression construct**

Yeast codon-optimized cDNA encoding human ClC-1 (UniProt accession P35523) was purchased from Genscript (Genscript, USA). cDNA was inserted into pEMBLyex4 [40] along...
Fig 4. Myotonia-causing mutations and the putative binding pocket of the 9-AC inhibitor. (a–d) Disease-causing and experimental missense mutations in ClC-1. Substitutions that invert (from depolarization to hyperpolarization activated) or shift the voltage dependence are shown in pink (located to the extracellular side) and blue (intracellular vestibule) or green (subunit interface), respectively. Bright colors represent disease-causing (recessive, with a stronger phenotype, but not dominant mutations are underscored), whereas experimental mutations are shown in pale colors. ClC-1 is shown in white and the Cl\(^{-}\) vestibules in purple (calculated using HOLLOW as for Fig 2). (a) Overall view with all known disease and selected experimental mutations. We note that mutations of 5 residues that cause recessive myotonia and inward rectification are facing the extracellular vestibule; 3 located in a row on the same face of helix B (M128, S132, D136) and 2 being the pore-constricting residues (K231, R421). Therefore, the phenotype may reflect a decreased chloride affinity of an extracellularly accessible site. (b) Close view of mutations that invert voltage...
dependece. Helices are labelled with white letters throughout. (c) Close view of mutations that shift voltage dependence (located at the intracellular vestibule). (d) Close view of mutations that shift voltage dependence (located at the monomer:monomer interface). (e−f) A putative binding pocket of 9-AC. Residues known to affect binding of 9-AC are highlighted as spheres (red for strong effect and pink for minor) and overlay a CAVER calculated pathway (shown in yellow) that stretches from the intracellular membrane interface to GluGATE. The clustering hints at a suitable target point for future rational drug-design efforts (see also alternative view in S11 Fig). CBS, cystathionine-β-synthase; 9-AC, 9-anthracene-carboxylic acid.

https://doi.org/10.1371/journal.pbio.3000218.g004

with yeast-enhanced GFP by homologous recombination to encode ClC-1, followed by a Tobacco Etch Virus (TEV) cleavage site, GFP, and a His\textsubscript{10} tag. The correct nucleotide sequence of the expression construct was verified by DNA sequencing (Eurofins MWG Operon, Germany).

**ClC-1 expression and purification**

Human ClC-1 was produced in the PAP1500 strain [41] grown in computer controlled 15-L bioreactors as previously reported but without addition of any chloride salts (such as NaCl) [42]. Yeast cells were harvested approximately 90 hours after induction of ClC-1 expression.

For crude membrane preparations, approximately 25 g of yeast cells were resuspended in 25 mL lysis buffer (25 mM imidazole [pH 7.5], 1 mM EGTA, 1 mM EDTA, 10% glycerol, 5 mM β-mercaptoethanol) supplemented with protease inhibitors (1 μg/mL leupeptin, pepstatin, and chymostatin, and 1 mM PMFS). Cells were disrupted by addition of glass beads (0.4–0.8 mm) and vortexed in 50-mL Falcon tubes 8 times for 1 minute. The supernatant was collected, and glass beads were washed several times in ice-cold lysis buffer. The cell lysate was centrifuged at 1,000 g for 10 minutes to remove cell debris. Crude membranes were pelleted from the supernatant by ultracentrifugation at 160,000 g for 90 minutes; resuspended in a buffer containing 50 mM Tris (pH 7.5), 300 mM NaCl, 10% glycerol, 1 mM PMSF, and EDTA-free protease inhibitors (Sigma); and homogenized in a Potter-Elvehjem homogenizer. Subsequently, membranes were solubilized by adding dodecyl-β-maltoside (DDM) and cholesteryl semi succinate (CHS; from Anatrace) at final concentrations of 1% and 0.33%, respectively, and incubated at 4°C for 3 hours under gentle stirring. Nonsolubilized material was removed by ultracentrifugation at 30,000 rpm for 30 minutes in a Beckman Ti 60 rotor. Ni-beads from 5 mL of slurry (Thermofisher) were incubated with the supernatant for 2 hours under gentle stirring. To prevent unspecific binding, 30 mM imidazole was added. Resin was transferred to a 5-mL Econo column (Bio-Rad) and washed with 10 column volumes of high-salt buffer (50 mM Tris [pH 7.5], 800 mM NaCl, 5% glycerol, 0.4 mg/mL DDM, and 0.04 mg/mL CHS) followed by 10 column volumes of low-salt buffer (50 mM Tris [pH 7.5], 300 mM NaCl, 5% glycerol, 0.4 mg/mL DDM, and 0.04 mg/mL CHS). ClC-1 protein was liberated from the beads by overnight incubating in 10 mL low-salt buffer containing 0.2 mg of TEV protease. Ni-beads were washed twice with 5 mL of low-salt buffer, and all collections were pooled and concentrated to approximately 1 mL using a 100,000 kDa cutoff concentrator device (Sartorius). Amphipol PMAL-C8 (Anatrace) was added to the purified protein at a mass ratio of 1:5 and incubated overnight. To remove DDM, protein was dialyzed overnight against final buffer (20 mM Tris [pH 7.5], 100 mM NaCl, 0.2 mM TCEP) supplemented with 100 mg of SM-2 BioBeads (Bio-Rad). The protein-amphipol complex was applied to a Superdex-200 column equilibrated with final buffer. Peak fractions were collected and concentrated to approximately 0.5 mg/mL. For the low pH samples, the purification procedure was identical except for using 20 mM BisTris (pH 6.2) (instead of Tris [pH 7.5]) in the final buffer (final protein concentration only reached approximately 0.3 mg/mL due to precipitation).
Single-channel ion conductance

Single-channel ion current was recorded using 2 separate methods, as follows:

1. **The Nanion Orbit Mini bilayer system.** Lipid bilayers were formed using 10 mM 1,2-diphtyanyol-sn-glycero-3-phosphocoline (DPhPc) and 1 mM cholesterol in n-nonane (Avanti Polar Lipids), and single channels were inserted by addition of purified protein (0.2 μL of 0.59 μg/μL ClC-1 in DDM) to recording solution at the cis side of the bilayer (150 μL). Current was recorded at ±150 mV in symmetrical solutions containing 1 M KCl and 10 mM HEPES (pH 6.2 with KOH). Recordings were digitized at 1.25 kHz, low-pass filtered at 160 Hz, and analyzed using Clampfit 10 after 100 Hz digital filtering.

2. **The Nanion Port-a-Patch system.** Giant unilamellar vesicles (GUVs) were made from 5 mM DPhPc and 0.5 mM cholesterol in chloroform by electroformation in 1 M sorbitol using the Vesicle Prep Pro (Nanion Technologies). Purified ClC-1 protein in DDM was mixed with GUVs to a final concentration of approximately 50 ng/mL and incubated overnight at 4˚C with SM-2 Bio-Beads (Bio-Rad). Lipid bilayers were formed from the GUVs, and single-channel current was recorded in symmetrical solutions containing 1 M NaCl and 10 mM HEPES (pH 6.2 with NaOH) at ±150 mV for 1 second. Recordings were digitized at 50 kHz, low-pass filtered at 200 Hz, and analyzed using Clampfit 10.7 software (Molecular Devices, San Jose, CA).

Cryo-EM sample preparation and data collection

Cryo-EM grids were prepared with the Vitrobot Mark IV (FEI) operated at 100% humidity at 4˚C. Immediately prior to sample vitrification, Quantifoil 1.2/1.3-μm holy carbon grids were glow-discharged with Easyglow (TedPella), and fluorinated fos-choline-8 (Anatrace) was added to the protein sample to a final concentration of 3 mM, which was an essential step for producing good quality thin ice. For each grid, an aliquot of 3.5 μL was applied and incubated for 20 seconds inside the Vitrobot. Blotting time was set to 2.5 seconds with 2 seconds of drain time. The low pH sample was treated identically, except for incubation with 0.3 mM NAD before freezing (and that no fluorinated fos-choline-8 was added to obtain one of the pH 6.2 data sets). Cryo-EM data sets were collected on a Titan Krios electron microscope (FEI) operating at 300 keV with a Gatan K2 Summit direct electron detector attached to a Gatan imaging filter (GIF). Movies were recorded under super-resolution counting mode at a pixel size of 0.535 Å and a dose rate of 0.876 e/pixel/frame for a total of 60 frames. The total electron dose was 45 electrons per Å² per movie for 9 seconds.

Image processing and 3D reconstruction

Cryo-EM movies were first gain-corrected and 2× binned to a final pixel size of 1.07 Å. Dose-weighted and nondose-weighted summed micrographs were generated with MotionCorr2 [43] using all frames except the first one. Defocus values were calculated with the nondose-weighted micrographs using Gctf [44]. Next, image processing was conducted using dose-weighted micrographs with the predetermined defocus. Template-free particle picking was done using Kai Zhang’s Gautomatch software (https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch). All following processing steps were done in Relion 2.0 [45] using a box size of 288 pixels.

For the pH 7.5 data set, a total of 594,609 auto-picked particles from 4,475 micrographs with a defocus range of −1.0 to −3.0 μm were subjected to several rounds of reference-free 2D classification to remove defective particles. The selected 477,729 particles were sorted using 3D...
classification. Selected classes were refined using masks, either with the complete protein excluding the amphipol belt or with the membrane domain only. Multiple cryo-EM density maps were calculated demonstrating structural heterogeneity of the protein.

3D classification of particles into 5 classes provided the best class consisting of 176,871 particles (representing more than 37% of all particles). A soft mask covering the entire protein without amphipol belt yielded a map with an overall resolution of 4.00 Å, and a tighter mask only containing the membrane domain resulted in map with resolution of 3.63 Å. To further investigate the structure heterogeneity in the cytoplasmic domain, the 2D selected particles were first refined, and then the refined per-particle parameters were applied for 3D classification, only performing local angular searches within ±10 degrees. This local 3D classification resulted in 9 classes, and the 2 major classes differed primarily in the cytoplasmic domain. Refinement of these 2 classes, each representing approximately 15% of all selected particles, yielded overall map resolutions of 4.34 Å and 4.28 Å, respectively.

For the pH 6.2 data set collected with fluorinated fos-choline-8, 552,914 particles were auto-selected from 4,119 motion-corrected micrographs, and 300,572 particles were selected after 2D classification for further processing; 3D classification into 5 classes generated the best class, which eventually was refined to a final resolution of 4.47 Å. Combination of the data collected at pH 6.2 with and without fluorinated fos-choline-8, and a similar local angular search strategy as for the pH 7.5 data set, generated a final map of 4.2 Å of the best class (based on approximately 30% of the total particles).

C2 symmetry was applied for all classification procedures, and all maps were sharpened with a B-factor of −100 Å². Local resolution was calculated using the postprocessed map, and the map was filtered according to the local resolution and used for model building.

Model building and refinement

The initial model was generated using the SWISS-MODEL online server and the ClC-K structure [9] (PDB-ID 5TQQ) as a template. The model was first fitted into the cryo-EM density map and later manually built in COOT [46]. The 3.6 Å membrane domain density map was sufficient for building the entire membrane domain (residues 115 to 589) with only 1 loop missing (residues 254–261). The built model was refined using phenix.real_space_refine of the Phenix software package [47]. C2 symmetry was imposed during the refinement by using strong non-crystallographic symmetry (NCS) restraints. Secondary structure restraints and Ramachandran restraints were also imposed during refinement.

The resolution and connectivity of the cytoplasmic domain was insufficient for de novo model building. Instead, a homology model based on the available structure of the CBS domains of ClC-0 (PDB-ID 2D4Z [30]) was generated and docked into different maps. The refinement of the cytoplasmic domain was conducted by local grid minimization, model morphing, and simulated annealing implemented in the phenix.real_space_refine software [47]. To prevent overfitting, the map resolution was restricted to 5 Å, the local resolution of the cytoplasmic domains as determined by Relion postprocessing. After model building, the models were trimmed to only include the minimal CBS architecture, consisting of 2 helices and a β-sheet. The quality of the models were validated assessed using Molprobity [48] (see S1 Table for statistics). All figures except for Fig 3A and 3B were generated using the model based on the 4.0 Å (Map 1).

MD simulations

The ClC-1 dimer with Glu232 either protonated or deprotonated was inserted into a palmitoyloleoylphosphocholine (POPC) membrane, and CHARMM36 force field parameters [49, 50]
were generated using CHARMM-GUI [51]. The simulations were performed using the GROMACS 2016.4 simulation software [52]. Each system was energy minimized and equilibrated in a stepwise manner using 25-ps NVT simulations with decreasing restraints on the protein and lipid heavy atoms. In these simulations, a 1-fs time step was used and the temperature was maintained at 310 K with a Berendsen temperature-coupling scheme [53]. The following set of NPT simulations further released heavy-atom restraints for 0.1 ns, 10 ns, and 10 ns, respectively. Here, a 2-fs time step was used and the pressure was kept constant at 1 bar using a Berendsen pressure barostat [53]. In a 100 ns production simulation, all atoms were unrestrained, and the temperature and pressure coupling schemes were Nose-Hoover [54, 55] and Parrinello-Rahman [56, 57], respectively. The GROMACS pull code with a force constant of 1,000 kJ mol\(^{-1}\) nm\(^{-2}\) was applied for 300 ps to the Cl\(^-\) ion in closest vicinity of Glu232 in 1 monomer. The pull rate was 0.1 Å per ps, and the pull force was directed along the vertical axis of the membrane. The potential of mean force (PMF) was calculated using umbrella sampling from 1 Å windows along the ion path. The figures were generated using VMD software [58].

Supporting information

**S1 Fig. Single-channel recordings of purified ClC-1 channels.** (a) DDM-solubilized protein was incorporated into planar lipid bilayers consisting of 10 mM DPhPc and 1 mM cholesterol dissolved in n-nonane. Single-channel activity was measured using symmetrical solutions containing 1 M KCl and 10 mM HEPES (pH 6.2) at holding potentials of ±150 mV in the Orbit Mini system (Nanion Technologies). Openings and closings of the incorporated channels are marked, and zero current is indicated by blue lines. (b) DDM-solubilized protein was incorporated into GUVs consisting of 10 mM DPhPc and 1 mM cholesterol, and planar lipid bilayers were formed on an NPC-1 chip using symmetrical solution containing 1 M NaCl and 10 mM HEPES (pH 6.2). Single channel currents were recorded at ±150 mV using Port-a-Patch system (Nanion Technologies). Openings and closings of the incorporated channels are marked, and zero current is indicated by blue lines. (c) Amplitude histogram of single channel recordings obtained at −150 mV under same conditions as in panel b. The distribution of amplitudes was fitted with the sum of 3 Gaussian distributions. Single-channel conductance was calculated to 4.0 ± 0.2 pS (n = 21) for recordings obtained in 1 M NaCl and to 3.5 ± 0.1 pS (n = 90) for recordings obtained in 1 M KCl. The calculations were based on >3 independent experiments. (d) Single-channel recordings obtained at +200 mV using similar experimental conditions as in panel b, but in the absence and presence of 100 μM the chloride channel inhibitor 9-AC. The channel activity could be recovered after washout of 9-AC. The shown traces are representative of 3 independent experiments. It should be noted that reconstituted ion channels may incorporate with random orientation into the membrane. Therefore, the applied voltage is not necessarily reflecting the direction of the physiological membrane potential, and single-channel rectification properties of ClC-1 may not be correctly reproduced. The large chloride concentration employed (1 M) likely leads to complete opening of the fast gate, explaining why the double-barrelled appearance of ClC-1 is not apparent. Taken together, the measurements in the presence of NaCl or KCl suggest a Cl\(^-\)-dependent single-channel activity resulting from ClC-1. This is further supported by the fact that the current could be totally inhibited by 9-AC. We also note that the ClC-1 overproducing yeast cells were unable to thrive in standard media containing 1.7 mM NaCl and that minimal media without chloride was required for yeast growth and protein production. The underlying data for S1C can be found in S1 Data. CLC, chloride channel; GUV, giant unilamellar vesicle; 9-AC, 9-anthracene-carboxylic acid. (TIF)
S2 Fig. Cryo-EM image processing for the pH 7.5 and 6.2 data sets. (a) The 4 maps (Maps 0–3, respectively) generated using the pH 7.5. Maps 1–3 represent the overall structure refined by applying a mask that only covers the protein part without the amphipol belt but with differences in the cytoplasmic CBS domains. Map 0 represents the membrane domain map derived from focused refinement covering the membrane domain only. Map 1 with overall resolution 4.0 Å was generated by applying a mask covering the entire protein excluding the amphipol belt. After 3D refinement with a membrane domain mask, Map 0 with a resolution of 3.63 Å was obtained. Maps 2 and 3 were produced by 3D refinement of 2 major classes obtained from 3D classification using a local angular search strategy based on the model generated from 477,729 particles by 3D refinement directly (see Methods for further details). (b) pH 6.2 is suboptimal for ClC-1, leading to partial aggregation during purification and freezing. Hence, the collected data set at pH 6.2 is of less quality than that collected at pH 7.5. To obtain the pH 6.2 structure, we combined 2 data sets: (i) a data set collected with fluorinated fos-choline-8 (as the pH 7.5 data set) processed to an overall resolution of 4.47 Å (derived from 34.2% of the particles following 3D classification into 5 classes; we did not identify secondary structure features for the remaining 4 classes, suggesting that there is a large fraction of low-quality particles in the data), and (ii) a second data set without fluorinated fos-choline-8. The second pH 6.2 dataset yielded nonoptimal ice thickness but provided views that were not observed in the first one. The final map derived from combination of these two data sets, following 3D classification with a local angular search strategy as for the pH 7.5 data set, produced 5 classes, of which the best was refined to an overall resolution of 4.2 Å. This class represents 30.6% of the particles, of which 76,881 particles were from the first data set and 49,221 particles were from the second data set. New orientation that was not observed in the first data set is highlighted with a red square in the lowest image. The density of the cytoplasmic CBS domains are better resolved at pH 6.2 compared to the pH 7.5 data set (see also S7 Fig). The maps are all contoured at level $\sigma = 0.013$ in Chimera. CBS, cystathionine-$\beta$-synthase; cryo-EM, cryo-electron microscopy.

(TIF)

S3 Fig. Evaluation of the local resolution of the cryo-EM maps. From top to bottom: (a) Map 0, (b) Map 1, (c) Map 2, (d) Map 3 (all of the pH 7.5 data set), and (e) the final map from the pH 6.2 data sets. See S2 Fig for further information regarding the generated maps. From left to right, Euler angle distribution, FSC, masks exploited for the refinement evaluations, and color-coded local resolution distribution calculated by Relion in two different views (the maps are contoured at level $\sigma = 0.03$ in Chimera, except for Map 0, which is at level $\sigma = 0.044$). The angular distribution plots suggest a high degree of anisotropy. Note that more density features and better connectivity are observed for the CBS domains in the low pH structure. CBS, cystathionine-$\beta$-synthase; cryo-EM, cryo-electron microscopy; FSC, Fourier shell correlation.

(TIF)

S4 Fig. Cryo-EM density of selected parts of the ClC-1 membrane region. The helices are colored as in Fig 1D with the maps contoured at level $\sigma = 0.03$ in Chimera using Map 0. The modelled Glu$_{GATE}$ (E232) is colored blue and an alternative (but not modeled) orientation is shown in gray. Numbers in parentheses indicate shown residues. cryo-EM, cryo-electron microscopy.

(TIF)

S5 Fig. Novel structural features revealed by the ClC-1 structure. Depicted as in Fig 1D and with ClC-K in gray, the new features are highlighted with arrows. (a–f) The $\alpha$C–D, $\alpha$E–F, $\alpha$H–I, $\alpha$L–J, $\alpha$L–M, and $\alpha$N–O loops, respectively. Helices are labeled with white letters throughout.
(d) Details of the extracellular αl–J loop, which is targeted by several dominant disease mutations (see also Fig 4). This loop was not observed in the ClC-K and CmClC (PDB-ID 3ORG) [14] structures. Note the short distance between the αl–J loop and R421 of the vestibule, hinting at a role for αl–J in controlling chloride passage. Residue T335 (which was differently placed in a recent homology model [59]) in the αl–J loop is within reach of Q552 in the αO–P loop, possibly providing a communication bridge of extracellular cues to conformational changes at the dimer interface or in the pore region (mutations of both T335 and Q552 cause inward rectification [35]). In ClC-K channels, 2 symmetrically localized inter-subunit regulatory Ca$^{2+}$ binding sites are formed by αl–J loop residues [60, 61]. The corresponding residues in CIC-1 are not oriented in a manner consistent with Ca$^{2+}$ binding. (g, h) Details of E232 (GluGATE) and a possible (not modeled) alternative orientation of its side chain in panel h (see also S4 Fig).

S6 Fig. Molecular dynamics simulations suggest that protonated GluGATE (E232) primes CIC-1 with Cl$^-$ for ion conductance, but Cl$^-$ transfer across GluGATE is nonspontaneous. Simulations were performed in the presence of a POPC membrane and 100 mM NaCl (no gradient). (a–c) Protonated E232 primes CIC-1 with Cl$^-$ for conductance without applied force. Observed ion positions (gray spheres) during the course of the simulations with E232 being (a) protonated and (b) deprotonated. Notably, Cl$^-$ ions reach R421 from the extracellular side and GluGATE from the intracellular side. For clarity, only GluGATE, TyrC, and R421 are visualized (as sticks). (c) The number of Cl$^-$ ions within 5 Å of GluGATE in both CIC-1 monomers in simulations with protonated GluGATE (red) and deprotonated GluGATE (black). Protonated GluGATE typically coordinates a single Cl$^-$ in the latter half of the simulation. In contrast, Cl$^-$ comes into proximity only transiently and is repelled by the negative charge of deprotonated GluGATE. (d–e) Cl$^-$ movement across CIC-1 with a protonated E232 from the primed Cl$^-$ position at the GluGATE-TyrC pair (observed in panel a) appears nonspontaneous. (d) Free energy associated with moving (by applying force) Cl$^-$ from the primed position to the extracellular side. Positive free energy barriers in the range of 3 kcal/mol indicate that movement along the sampled reaction coordinate is a nonspontaneous process. The distance moved is relative to the TyrC (Y578) hydroxyl oxygen. (e) The observed (not necessarily native) Cl$^-$ (gray spheres) transport pathway exploited for calculating panel d is shown on the overall structure. For clarity, only GluGATE, TyrC, and R421 are pinpointed (sticks). Collectively, these MD simulations suggest that although the determined structure may be closed, Cl$^-$ ions may spontaneously penetrate deep into the vestibules from both sides of the membrane. MD, Molecular Dynamics.

S7 Fig. Maps of the overall structures and CBS domains, respectively. The maps of the overall structures are shown at level σ = 0.03 in Chimera, whereas those of the CBS domains only are shown at level σ = 0.035. Maps 1–3 of the pH 7.5 data and map of the pH 6.2 data are shown. (a) The overall structures. (b) The CBS domains. CBS, cystathionine-β-synthase.

S8 Fig. Details of the CBS shifts. Alternative views of Fig 3 (colored identically). (a) Alternative view of Fig 3B. (b–d) Identical view as in panel a with comparisons of the structures derived from Maps 1–3 (pH 7.5) and the pH 6.2 map (aligned as in Fig 3A), respectively. (e) Identical view as Fig 3C, including the structure determined at pH 6.2. Helices are labeled with white letters throughout. CBS, cystathionine-β-synthase.
S9 Fig. Comparison of ClC-1 samples at pH 7.5 and 6.2. (a) Size-exclusion chromatography profiles of ClC-1 at pH 7.5 and 6.2. The protein peak at pH 6.2 is shifted toward a higher retention volume indicating a more compact ClC-1. The peak appearing at 0.7 CV represents the signal from free PMAL-C8 amphipol. (b) Micrographs for the pH 7.5 (left) and 6.2 (middle and right) data sets (with and without F-FC-8, respectively), indicating worse behavior of the pH 6.2 sample. CV, Column volume; F-FC-8, fluorinated fos-choline-8.

S10 Fig. Sequence alignment of selected CLC proteins. All human CLC members and structurally determined CLC proteins are displayed. Secondary structure elements are pinpointed, and conserved residues are highlighted in red and green, the latter representing residues relevant for the function of ClC-1 discussed in this work. CLC, chloride channel.

S11 Fig. Details of the putative 9-AC pocket. Alternative views of Fig 4E and 4F (colored identically). 9-AC, 9-anthracene-carboxylic acid.

S1 Table. Cryo-EM data validation statistics. cryo-EM, cryo-electron microscopy.

S1 Data. Calculation of single-channel recordings of purified ClC-1. Single-channel conductance was determined based on single-channel current recordings and calculated to 4.0 ± 0.2 pS (n = 21) for recordings obtained in 1 M NaCl and to 3.5 ± 0.1 pS (n = 90) for recordings obtained in 1 M KCl. Recording solution, command voltage, and current amplitude of the individual recordings are listed.

Acknowledgments

We thank the members of ZHZ’s laboratory for suggestions in cryo-EM sample preparation and data processing and David Sørensen and Vibeke Grøsfjeld Christensen for technical assistance.

Author Contributions

Conceptualization: Kaituo Wang, Per Amstrup Pedersen, Pontus Gourdon.

Data curation: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Yanxiang Cui, Magnus Andersson, Kirstine Calloe, Dan Arne Klaerke, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Formal analysis: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Yanxiang Cui, Julie Winkel Missel, Christina Grønberg, Kamil Gotfryd, Erik Lindahl, Magnus Andersson, Kirstine Calloe, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Funding acquisition: Kaituo Wang, Erik Lindahl, Magnus Andersson, Kirstine Calloe, Pascal F. Egea, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Investigation: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Yanxiang Cui, Julie Winkel Missel, Christina Gronberg, Kamil Gotfryd, Erik Lindahl, Magnus Andersson, Kirstine
Methodology: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Yanxiang Cui, Julie Winkel Missel, Christina Gronberg, Kirstine Calloe, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Project administration: Kaituo Wang, Pontus Gourdon.

Resources: Erik Lindahl, Magnus Andersson, Kirstine Calloe, Pascal F. Egea, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Supervision: Kaituo Wang, Erik Lindahl, Pascal F. Egea, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Validation: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Yanxiang Cui, Julie Winkel Missel, Christina Gronberg, Magnus Andersson, Kirstine Calloe, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Visualization: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Julie Winkel Missel, Christina Gronberg, Magnus Andersson, Kirstine Calloe, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Writing – original draft: Kaituo Wang, Kamil Gotfryd, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

Writing – review & editing: Kaituo Wang, Sarah Spruce Preisler, Liying Zhang, Julie Winkel Missel, Christina Gronberg, Magnus Andersson, Kirstine Calloe, Pascal F. Egea, Dan Arne Klaerke, Michael Pusch, Per Amstrup Pedersen, Z. Hong Zhou, Pontus Gourdon.

References


Supplementary figures

S1 Fig – single channel recordings of purified ClC-1 channels

(a) A single channel recording showing the transition from open to closed states at 150 mV and -150 mV.

(b) A comparison of open and closed states at 150 mV and -150 mV, showing the conductance of ions.

(c) A histogram showing the distribution of conductance for NaCl and KCl.

(d) A recording showing the effect of 200 mV, 200 mV + 100 μM 9-AC, and a wash, indicating the transition between open and closed states.
S2 Fig. Cryo-EM image processing for the pH 7.5 and 6.2 data sets
S3 Fig. Evaluation of the local resolution of the cryo-EM maps.

(a) Euler angle distribution | Fourier shell correlation | Mask | Color-coded local resolution distribution
---|---|---|---
Map 0
(b) Euler angle distribution | Fourier shell correlation | Mask | Map 1
(c) Euler angle distribution | Fourier shell correlation | Mask | Map 2
(d) Euler angle distribution | Fourier shell correlation | Mask | Map 3
(e) Euler angle distribution | Fourier shell correlation | Mask | Map pH 6.2
S4 Fig. Cryo-EM density of selected parts of the ClC-1 membrane region.
S5 Fig. Novel structural features revealed by the CIC-1 structure.
S6 Fig. Molecular dynamics simulations suggest that protonated GluGATE (E232) primes CIC-1 with Cl− for ion conductance, but Cl− transfer across GluGATE is nonspontaneous.
S7 Fig. Maps of the overall structures and CBS domains, respectively
S8 Fig. Details of the CBS shifts.
S9 Fig. Comparison of ClC-1 samples at pH 7.5 and 6.2.
S10 Fig. Sequence alignment of selected CLC proteins.
S11 Fig. Details of the putative 9-AC pocket
S1 Table. Cryo-EM data validation statistics.

<table>
<thead>
<tr>
<th>EMDB</th>
<th>EMD-4645</th>
<th>EMD-4647</th>
<th>EMD-4649</th>
<th>EMD-4646</th>
<th>EMD-4657</th>
</tr>
</thead>
<tbody>
<tr>
<td>pdb-id</td>
<td>6QV6</td>
<td>6QVC</td>
<td>6QVD</td>
<td>6QVB</td>
<td>6QVU</td>
</tr>
<tr>
<td>Data statistics</td>
<td>map 0, pH 7.5</td>
<td>map 1, pH 7.5</td>
<td>map 2, pH 7.5</td>
<td>map 3, pH 7.5</td>
<td>map pH 6.2</td>
</tr>
<tr>
<td>Model-map-fit CC</td>
<td>0.81</td>
<td>0.81</td>
<td>0.762</td>
<td>0.756</td>
<td>0.736</td>
</tr>
<tr>
<td>R.m.s.d. bonds (Å)</td>
<td>0.010</td>
<td>0.008</td>
<td>0.016</td>
<td>0.012</td>
<td>0.009</td>
</tr>
<tr>
<td>R.m.s.d. angles (°)</td>
<td>1.290</td>
<td>1.471</td>
<td>2.060</td>
<td>1.580</td>
<td>1.635</td>
</tr>
<tr>
<td>Molprobity statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran Plot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favor (%)</td>
<td>91.6</td>
<td>91.1</td>
<td>90.3</td>
<td>80.0</td>
<td>89.9</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>8.0</td>
<td>8.4</td>
<td>9.6</td>
<td>10.0</td>
<td>9.5</td>
</tr>
<tr>
<td>Non-favor (%)</td>
<td>0.4</td>
<td>0.5</td>
<td>0.0</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>1.0</td>
<td>2.1</td>
<td>4.4</td>
<td>1.8</td>
<td>2.3</td>
</tr>
<tr>
<td>Clash score</td>
<td>4.7</td>
<td>8.9</td>
<td>14.8</td>
<td>12.0</td>
<td>12.7</td>
</tr>
<tr>
<td>Average B-factor (Å²)</td>
<td>118.0</td>
<td>178.0</td>
<td>267.0</td>
<td>303.0</td>
<td>325.3</td>
</tr>
</tbody>
</table>

S1 Data. Calculation of single-channel recordings of purified ClC-1.

<table>
<thead>
<tr>
<th>Orbit 27.11.17 folder 1, ch 1</th>
<th>Orbit10.04.18 data 0, Ch1</th>
<th>Average</th>
<th>Conductance</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 mV 1 M KCl</td>
<td>120 mV 1 M KCl</td>
<td>0.47519</td>
<td>3,167933</td>
</tr>
<tr>
<td>1 M KCl</td>
<td></td>
<td>0.607129</td>
<td>4,047527</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.332431</td>
<td>2,216204</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.355453</td>
<td>2,369689</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.429242</td>
<td>2,861614</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.819952</td>
<td>5,466348</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.548659</td>
<td>3,657728</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.328216</td>
<td>2,188105</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.250584</td>
<td>1,670563</td>
</tr>
<tr>
<td>AV</td>
<td>0.51375</td>
<td>3,468332</td>
<td>0.377476</td>
</tr>
<tr>
<td>SD</td>
<td>0.13919</td>
<td>0.932511</td>
<td>0.685835</td>
</tr>
<tr>
<td>n</td>
<td>90</td>
<td>90</td>
<td>0.854219</td>
</tr>
<tr>
<td>sem</td>
<td>0.014672</td>
<td>0.098295</td>
<td>0.55717</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.700375</td>
<td>4,669165</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.506667</td>
<td>3,37778</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.355386</td>
<td>2,369243</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.366893</td>
<td>2,445956</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.457477</td>
<td>3,049849</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.627752</td>
<td>4,185014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.624627</td>
<td>4,16418</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.543624</td>
<td>3,624158</td>
</tr>
<tr>
<td>Orbit 10.04.18 data 0, Ch1</td>
<td>1 M KCl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120 mV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.489812</td>
<td>3.265411</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.333888</td>
<td>2.225919</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.529776</td>
<td>3.531838</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.501794</td>
<td>3.345291</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.488306</td>
<td>3.255373</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.499946</td>
<td>3.332975</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.458155</td>
<td>3.054369</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.44297</td>
<td>2.953135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.563527</td>
<td>3.75685</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.63043</td>
<td>4.253618</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.496067</td>
<td>3.307114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.735949</td>
<td>4.906326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.498431</td>
<td>3.322873</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.633626</td>
<td>4.224173</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.616377</td>
<td>4.10918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.687585</td>
<td>4.583897</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.821533</td>
<td>5.476886</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.543339</td>
<td>3.622263</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.636614</td>
<td>4.244093</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.543411</td>
<td>4.528421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.447043</td>
<td>3.725358</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.407068</td>
<td>3.392234</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.45104</td>
<td>3.758667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.491178</td>
<td>4.093147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.51375</td>
<td>3.468332</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.13919</td>
<td>0.932511</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.014672</td>
<td>0.098295</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>