LUND UNIVERSITY

Expression and Purification of Channel Proteins

Aiming at structural and functional understanding of TRPAs and Aquaporins

Werin, Balder

2024

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Werin, B. (2024). *Expression and Purification of Channel Proteins: Aiming at structural and functional understanding of TRPAs and Aquaporins*. Lund University.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Expression and Purification of Channel Proteins

Aiming at structural and functional understanding of TRPAs and Aquaporins

Balder Werin

LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the Faculty of Science at Lund University to be publicly defended on Monday 3rd of June at 09.00 in Hall A, Kemicentrum, Department of Chemistry

Faculty opponent: Prof. David Drew, Stockholm University

Organization	Document name
LUND UNIVERSITY Division of Biochemistry and Structural Biology	DOCTORAL DISSERTATION
	Date of issue 2024-05-08
Author(s) Balder Werin	Sponsoring organization Formas

Title and subtitle

Expression and Purification of Channel Proteins
- Aiming at structural and functional understanding of TRPAs and Aquaporins

Abstract

Every living cell is surrounded by a cell membrane that is made up of amphipathic phospholipids, separating the aqueous solutions inside and outside the cell with a hydrophobic barrier. This compartmentalization is a prerequisite for life, but so are the many molecules that are floating in the membrane, altering its properties and connecting the inside with the outside. An important group are the transport proteins, that open ways for molecules and ions, that would normally be too large, too charged, or too polar to pass the membrane. The transport proteins are either active transporters – like pumps – or passive transporters – like channels. In this thesis, I put the spotlight on two types of channel proteins: Transient Receptor Potential ion channels, that let ions pass when activated by temperature or pungent chemicals, and Aquaporins (AQP), that are mainly responsible for letting water cross the membrane.

In my work, I have made an effort to study the structure of one TRP member in particular, known as TRPA1 from pine weevil (*Hylobius abietis*). I had to find the best possible conditions to solubilize the protein in detergents, and I have also investigated other tools such as nanodiscs to keep the protein stable in solution. One major hurdle has always been the low yields, and it was therefore that a GFP-tag (Green fluorescent Protein) was added to the protein construct, to facilitate the tracking of the protein and evaluation of purification methods. Coupled with flow cytometry, a method for measuring fluorescence and scattering of individual cells, this proved very useful in designing an expression and purification protocol.

The purified protein was used for Cryo-EM (Electron Microscopy), but the protein was difficult to freeze on grids with a good homogeneous spread of individual particles. The use of SRCD (Synchrotron Radiation Circular Dichroism) proved more successful, and confirmed the secondary structure of the protein, and gave information on the temperature stability of the protein, with and without agonists and calcium ions. The rapid evolution of machine learning in the field of bioinformatics has been of great aid to me, and I have used AlphaFold to predict several TRPA structures, not just of TRPA1.

I also studied two aquaporins, and their interactions with the FERM-domain of Ezrin. I used Microscale thermophoresis to determine the dissociation constant (K_D), and found some weak interactions, that may regulate aquaporin trafficking.

Channel proteins are complicated membrane proteins that are hard to express and purify, but with the help of GFP and various evaluation methods, a lot has been learned about their structure and function.

Key words:

TRPA1, Aquaporins, Channel Proteins, Pine weevil, Cryo-EM, Flow cytometry, GFP

Classification system and/or index terms (if any):

Supplementary bibliographical information:		Language English
ISSN and key title:		ISBN 978-91-8096-052-6 (print) 978-91-8096-053-3 (pdf)
Recipient's notes	Number of pages 55	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature_____ Date_____

Expression and Purification of Channel Proteins

Aiming at structural and functional understanding of TRPAs and Aquaporins

Balder Werin

LUND
UNIVERSITY

Coverphoto: Cross-stitched pine weevil by Anna-Karin Aronsson

© pp 1-53 Balder Werin 2024

Paper I © by the Authors (Manuscript unpublished)
Paper II © by the Authors (CC BY 4.0)
Paper III © by the Authors (CC BY 4.0)
Paper IV © by the Authors (Manuscript unpublished)

Faculty of Science, Department of Chemistry

ISBN: 978-91-8096-052-6 (print)
ISBN: 978-91-8096-053-3 (pdf)

Printed in Sweden by Tryckeriet E-huset, Lund University, Lund 2024

– Varför blir du inte akademiker då, du som älskar att gå i skola.

– Du menar doktorera? Men det är ju en livsuppgift. Det är just en livsuppgift jag aldrig kan bestämma mig för.

– Det är inte mer än högst en femårsplan nu för tiden, säger han och visar mig en tidningsartikel om den nya doktorsexamen: Epadoktor kallar de den, för att markera hur lättköpt den är.

– ur ”Maken” av Gun-Britt Sundström

Table of Contents

Abstract

Every living cell is surrounded by a cell membrane that is made up of amphipathic phospholipids, separating the aqueous solutions inside and outside the cell with a hydrophobic barrier. This compartmentalization is a prerequisite for life, but so are the many molecules that are floating in the membrane, altering its properties and connecting the inside with the outside. An important group are the transport proteins, that open ways for molecules and ions, that would normally be too large, too charged, or too polar to pass the membrane. The transport proteins are either active transporters – like pumps – or passive transporters – like channels. In this thesis, I put the spotlight on two types of channel proteins: Transient Receptor Potential ion channels, that let ions pass when activated by temperature or pungent chemicals, and Aquaporins (AQP), that are mainly responsible for letting water cross the membrane.

In my work, I have made an effort to study the structure of one TRP member in particular, known as TRPA1 from pine weevil (*Hylobius abietis*). I had to find the best possible conditions to solubilize the protein in detergents, and I have also investigated other tools such as nanodiscs to keep the protein stable in solution. One major hurdle has always been the low yields, and it was therefore that a GFP-tag (Green fluorescent Protein) was added to the protein construct, to facilitate the tracking of the protein and evaluation of purification methods. Coupled with flow cytometry, a method for measuring fluorescence and scattering of individual cells, this proved very useful in designing an expression and purification protocol.

The purified protein was used for Cryo-EM (Electron Microscopy), but the protein was difficult to freeze on grids with a good homogeneous spread of individual particles. The use of SRCD (Synchrotron Radiation Circular Dichroism) proved more successful, and confirmed the secondary structure of the protein, and gave information on the temperature stability of the protein, with and without agonists and calcium ions. The rapid evolution of machine learning in the field of bioinformatics has been of great aid to me, and I have used AlphaFold to predict several TRPA structures, not just of TRPA1.

I also studied two aquaporins, and their interactions with the FERM-domain of Ezrin. I used Microscale thermophoresis to determine the dissociation constant (K_D), and found some weak interactions, that may regulate aquaporin trafficking.

Channel proteins are complicated membrane proteins that are hard to express and purify, but with the help of GFP and various evaluation methods, a lot has been learned about their structure and function.

Popular summary

Let's talk about pain. It is an experience that most people know, but that few can describe precisely. If you wanted to find someone who could define pain, you would have to ask either a poet or a physician, and I am neither of those. While I cannot tell you what pain is, I can tell you a little about how you get to know more about it, and that is in a sense what this thesis is about.

What I have made my focus, is the part of the body furthest from the brain. If you were to touch or taste something warm or irritating, your body would react to this, and the first part to react would be *proteins*. Proteins are molecules in all living organisms that are responsible for everything from breaking down food and junk, to making up the stabilizing structures in cells, and of course, sensing the world around or inside your body.

The cells of our bodies can be seen as bubbles that are both surrounded and filled with water. The shell of such a bubble is made up of an oily layer called the cell membrane. Oil and water will not mix, and the cell membrane can therefore be seen as a wall between the outside and the inside of the cell. Proteins in cell membranes are fittingly known as membrane proteins, and among them is a group of proteins known as *channel proteins*, that open a passage through the cell membrane. They usually tend to be very picky about what they let through, however. During my PhD, I have studied TRPA1 that allow ions (charged particles) to pass, and aquaporins that are open to water.

Proteins are big and complex molecules, and one of the most important parts of studying them is looking at their three-dimensional structure. When you have the structure of a protein, understanding how it works becomes more of a physical or even mechanical question than a biological one. It is possible to predict how new drugs will work or how a protein has evolved by letting a computer twist and turn a structure.

So how do you get a structure? First of all, you need a really pure sample of a protein. This is especially difficult for membrane proteins, as you must imitate the oily cell membrane surrounded by water. The most common way of doing this is with detergents – yes, the same type of chemicals that you would find in cleaning products – as they can capture tiny droplets of oil in a water solution.

I have, throughout my PhD-studies, studied how to best handle membrane proteins, but I have also given the actual structure determination a shot. I managed to describe the stability of the protein TRPA1 from pine weevil, and I discovered hints to the substance BITC being able to open the channel.

Populärvetenskaplig sammanfattning

Jag vill klargöra en grej: Jag har inte drivits av att besvara en stor och konkret fråga. Mitt bidrag till forskarvärlden består inte av att jag grubblat över universums mysterier och gett mig ner i den mörkaste vrån där ljuset från kunskapens fackla aldrig någonsin nått. Jag har arbetat där jag står. Med mina bara händer (och en pipett) har jag stått precis i kanten där ljuset flackar, och lagt pusselbit för pusselbit. De har inte alltid passat, och jag har ofta behövt ta ett steg tillbaka för att jag inte ens känt igen bitarna jag tidigare trodde låg stabilt.

För att dra den här långsökta metaforen ytterligare, så har jag faktiskt inte ens hållit på med att lägga pussel så stor del av min doktorandtid. Jag har behövt karva fram pusselbitarna ur de hårdaste och motsträvigaste av material, och när jag varit klar har de inte alltid varit vad jag hoppades på.

Vad är det för pussel jag har hållit på att lägga? Jag har haft målet inställt på att beskriva proteinet TRPA1 från snytbagge. Metoderna har varierat, men gemensamt har ändå varit att jag har haft den tredimensionella strukturen i sikte hela tiden.

Snytbaggar är skadeinsekter som angriper träd, och hela idén bakom mitt projekt har varit att studera hur en smärtreceptor – TRPA1 – reagerar på irriterande kemikalier så som ämnen från wasabi och pepparrot. Med den kunskapen så vore det möjligt att utveckla avskräckande medel som kunde ersätta bekämpningsmedel, men med mindre negativ påverkan på naturen.

Även om mitt forskningsprojekt inte har drivit det här från idé till färdig produkt, så har jag jobbat med att fylla hål och bygga broar, för att i framtiden kunna nå lösningar på problem. Det är därför svårt att beskriva vad det är jag har gjort under mina år som doktorand, men jag är stolt över mitt bidrag till vårt samlade vetande om TRPA1. Om proteinproduktion i jästceller. Om användningen av ytaktiva ämne i membranproteinforskning. Om användningen av ”AI” för att förutsäga proteinstrukturer. Om hur vattenkanalproteiner påverkas av andra proteiner.

Pusslet är inte klart. Jag har kanske inte lagt en hel pusselbit ens. Men världens samlade vetande är lite lite större, och det är det som den här avhandlingen handlar om.

Acknowledgments

When hiking through a tangled forest of a PhD, having an experienced guide that leads you by the hand is so important. I want to thank you **Urban**, my supervisor through these years. I am very grateful for the support you have offered, and your gentle encouragement. I do also appreciate your deep knowledge of proteins, and your eagerness to share it with me at all times (no matter if I have the time to listen or not). We didn't take the straightest path, but the lessons I learnt were well worth the trip.

I also want to thank all the members of the group. **Angelica** helped me getting started. **Jonna** started on the same square as me, and we struggled together my first year. I hope Umeå treats you well! Thanks also to **Veronika**, **Oliwia**, **Erikas**, **Wille**, **Titus** and **Isabelle** for sharing my journey!

Susanna, your group has been a bonus-group for me, and our weekly meetings were really important for me, both socially and scientifically. **Helin**, thank you for your warm support and deep knowledge of life and LOTR. Thank you, **CJ**, for good laughs. Sorry that I didn't make a good climber! Thank you, **Simon**, for the eternal smile and quick brains, both in our and Susanna's group. **Tamim**, you have saved me so many times at the FPLC, and you are always so nice and helpful. Thank you!

CMPS has been a wonderful place to do my PhD, and I am so happy to have met all the people that makes it so. Thank you, **Maryam**, for your kindness. You are the social glue that keeps that division together, and make sure that everyone is seen. And **Magnus**, you are the duct tape that not just keeps the instruments running, but also always makes sure that coffee time will happen. There is no choice, good times will be had! Thank you, **Camille**, for the straightest back of CMPS. I really appreciate your company, both when teaching long days in biology building, and when talking about anything important or silly. Thank you, **Niels**, for our long language discussions. I guess I just taught you obscure Swedish proverbs, but they might be useful too? I am so sorry that I didn't get to know you sooner **Ipsita**, but I will always remember you as the giggling sun ray of the division. I hope there is origami, wherever you are. Thank you, **Carmen**, for all the nice chats and the excellent co-hosting of Dora, with the relaxing style of **Mads**, and with **Andreas**. I really appreciated teaching with you. Thank you, **Lovisa**, you were a great teaching partner! An thank you to all the people of CMPS that I didn't mention, but that I feel so grateful for getting to know!

Thanks **Magnus**, **Nina** and **Tova** over at applied microbiology for valuable discussions and help with the flow cytometer.

Thank you, **Marjorie**, for the excellent collaboration with those insect TRPAs.

Thank you, **Olle Bergman**, for words of wisdom and for pointing to the world outside of the lab.

Thank you, **Anna-Carin Bergman** (I don't think you are related?), for all the comparison between the worlds of researchers and teachers. You managed to convince me to try my pedagogical wings, and I look forward to it.

Thank you, **Alte Kamereren**, for supplying me with my weekly dose of hmm…, maybe not… and marching music.

Erik, tack för regelbundet intag av skräpmat och skitsnack. Fattar inte att du bestämde dig för att doktorera, trots mitt varnande exempel, men lycka till!

Astrid och **Lars**, kan det vara på grund onsdagsmiddagarna efter schackklubben som jag kom hit? Oavsett så har ni varit oslagbara, och jag hoppas att ni njuter av läsningen

Pappa, jag har inte alltid varit mottaglig för dina tips, men jag är så glad för att du ändå gett dem, och att du korrekturläste avhandlingen! Och tack för alla stöttande ord!

Mamma, du är bäst på att påminna om att det är viktigt att fira, och det ska vi göra!

Atle, ledsen att jag mest snackar tarmar när jag ringer, men det skulle ju inte blivit någon avhandling utan dem!

Finn, jag vet att jag sysslar med pseudo-magi-vetenskap-ish, men alla kan ju inte vara fysiker. Och snytbaggarna torterar inte sig själva. Tack för att du påminner om hur viktig disputationen är, jag skulle nog inte ta det på tillräckligt allvar utan dig.

Alfie, jag visste inte att jag behövde dig, men nu kan jag inte tänka mig ett hem utan dig.

Anna-Karin, kan du förklara för mig hur vi de senaste fem åren, inte bara överlevt, men kommit ut på andra sidan med flaggan i topp? Nu kan det bara bli bättre. Jag älskar dig!

List of Papers

Paper I

Structural characterization of HaTRPA1

Balder Werin, Erikas Blazukas, Wilhelm Hansson Wennersten, Niels Meijer, Oliwia Kołodziejczyk, Urban Johanson
Manuscript

Paper II

Evaluation of heterologous expression in *Pichia pastoris* of Pine weevil TRPA1 by GFP and flow cytometry

Balder Werin, Wilhelm Hansson Wennersten, Robin Olsson, Oliwia Kołodziejczyk, Martin N. Andersson, Magnus Carlquist, Urban Johanson (2024)
Microb Cell Fact, 23(1): 110, https://doi.org/10.1186/s12934-024-02382-5
Reproduced under Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/legalcode

Paper III

***TRPA5* encodes a thermosensitive ankyrin ion channel receptor in a triatomine insect**

Marjorie A. Liénard*, David Báez-Nieto*, Cheng-Chia Tsai, Wendy A. Valencia-Montoya, Balder Werin, Urban Johansson, Jean-Marc Lassance, Jen Q. Pan, Nanfang Yu, Naomi E. Pierce (2024)
iScience, 27(4), https://doi.org/10.1016/j.isci.2024.109541
*These authors contributed equally
Reproduced under Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/legalcode

Paper IV

Structural basis for the interaction between the Ezrin FERM-domain and human aquaporins

Helin Strandberg, Carl Johan Hagströmer, Balder Werin, Markus Wendler, Urban Johanson and Susanna Törnroth-Horsefield
Manuscript

Author's contribution to the papers

Paper I

In this paper, I conceived, drafted and planned the study. I performed the initial detergent screen and the detergent exchange to DDM and GDN, as well as all the SRCD experiments. I also establishing the protocol that was used for the purification of HaTRPA1, and led the experimental work. I analyzed a significant part of the data, and I took a significant part in the drafting and writing of the manuscript.

Paper II

My contribution to the second paper consisted of conceiving, drafting and planning of the study. I made the original truncated construct and performed the small-scale expression. I executed all the flow cytometry and fluorescence microscopy measurements, and analyzed the data. I wrote the majority of the paper.

Paper III

My contribution to this paper consisted of modeling TRPA5 along with a set of homologues in AlphaFold, and analyzing and comparing important features. I also drafted the structure prediction section, along with figures, and partook in the writing of the final version.

Paper IV

In the fourth paper, I expressed the AQP2 peptide and FERM, and performed the MST experiments on the AQP2 peptide. I wrote part of the methods section regarding the expression of the AQP2 peptide.

Introduction

Proteins

If you would have visited the Society for Experimental Biology symposium at University College London in 1957, you would have heard Francis Crick give a talk in which he explained the flow of information that make possible what we know as life (1). In his lecture, Crick drew up something he called the "central dogma", which in its essence explains that genetic information flows from DNA through RNA to protein, but never back from protein. He kept a philosophical and broad tone, much due to the lack of experimental evidence at the time, but his bold ideas are now considered to be fundamental to our understanding of biology. In a sense, that is what life is: transfer of information. DNA is replicated and passed on between generations, RNA is transcribed from DNA, and amino acid sequences are translated from RNA. Proteins form into incredibly complex structures and interact with each other or their environment in unfathomable ways, and yet it all is – at least in theory – predictable if you know how to read the genetic code.

With the advent of "artificial intelligence", some argue that computer code is reaching a completely new level of complexity: it becomes almost alive. That means that you no longer have the ability to treat it like a piece of code, but rather interact with it like an entity of its own. The biological equivalent would be that the introduction of bioinformatics has not yet replaced the need to observe more than just the flow of information. Even with structural prediction by machine learning turning the field of protein science upside-down, information science is not enough to answer all questions. In order to understand proteins, sleeves need to be rolled up, and hands need to get dirty. Enter the biochemist.

The special case: membrane proteins

The cell membrane – not just lipids

Surrounding every living cell is a thin layer separating the aqueous solutions inside and outside the cell. This is the cell membrane, and it forms an oil phase that makes up a barrier between the sensitive machinery of the cell, and the outside world. The

membrane mainly consists of amphipathic molecules known as phospholipids, but there are also a lot of other molecules, among them membrane proteins. The most common way of describing the way cell membranes behave is by the fluid mosaic model (2). The idea is that the membrane is no solid structure with fixed positions for all its components, but rather behaves more like a two-dimensional liquid, where molecules are able to move laterally within the membrane. However, movement *across* a lipid bilayer is strictly limited to small, hydrophobic, non-charged particles. (3)

Membrane proteins are classified into integral and peripheral proteins, depending on if they are embedded in the cell membrane, or just attached to the surface. The structure, especially of integral membrane proteins, is easily recognizable by the hydrophobic part that is required for the protein to be stable within the lipid bilayer (3). The structures of membrane proteins are so tightly dependent on their relationship to the lipid bilayer that many structures have phospholipids integrated in key parts, without which the protein structure would not work (4). These particular features of membrane proteins make them hard to study, but they do make up a rough quarter of the human proteome, and 40% of molecular targets of pharmaceuticals, and the study of them is therefore a mission that should be prioritized (5).

Channel proteins

Since the lipid bilayer of the cell membrane is impermeable to both charged and hydrophilic molecules, there is a need for a helping hand. The help comes in the form of transporters of two kinds: pumps and channels, that allow for active and passive transport respectively. Among the most well-studied channel proteins are voltage-gated potassium channels, that allow for passive diffusion of K^+ to exploit the electrochemical gradient generated by ion pumps. The members of the group usually consist of four transmembrane subunits, that together form the pore that selectively channels ions across the membrane. The gating of the pore is typically controlled by a voltage-sensing domain embedded in the membrane, that opens and closes the channel by conformational changes in response to electrical signals (6). In this thesis however, the main focus will be on the channel protein groups Transient Receptor Potential (TRP) ion channels, and Aquaporins (AQP). (3)

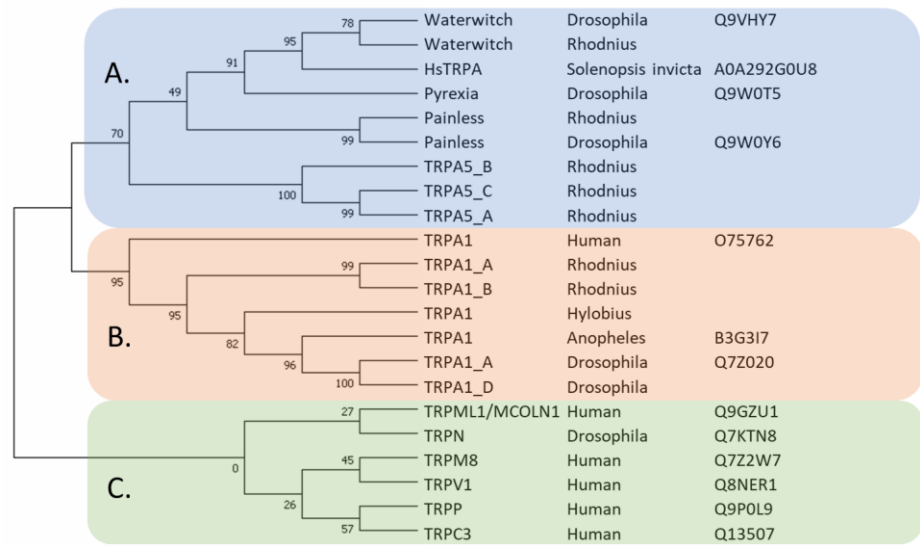


Figure 1: Phylogenetic tree of TRPs. A. TRPAs that are only present in insects. B. TRPA1 from different species. C. Other TRPs. The underscore letters indicate different splice forms or different gene loci. To the right is shown accession numbers for sequences available in data bases. Group 2 constists of TRPP and TRPML, the rest are group 1.

The TRP superfamily

When studying the light sensitivity in *Drosophila*, a gene mutation was discovered that caused a transient voltage response (7). This eventually led to the description of a large and diverse superfamily of proteins known as TRP ion channels (Transient Receptor Potential), characterized by them forming cation channels. They are all made up of six transmembrane helices per subunit that assemble into tetramers, spanning the membranes of almost any animal cell (8). TRPs are divided into two groups, that are further divided into families (figure 1), (9). In figure 2 is shown how the structures of some representatives of different TRP families differ. In the last years, there has been great advances in the description of TRPs, and much has been learned about the diverse nature of their activity and structural features, leading up to the Nobel Prize in 2021 being awarded to David Julius for his work with TRPV1. Some TRPs are well studied because of their involvement in pain responses (10). TRPs are also related to voltage-gated potassium channels (9).

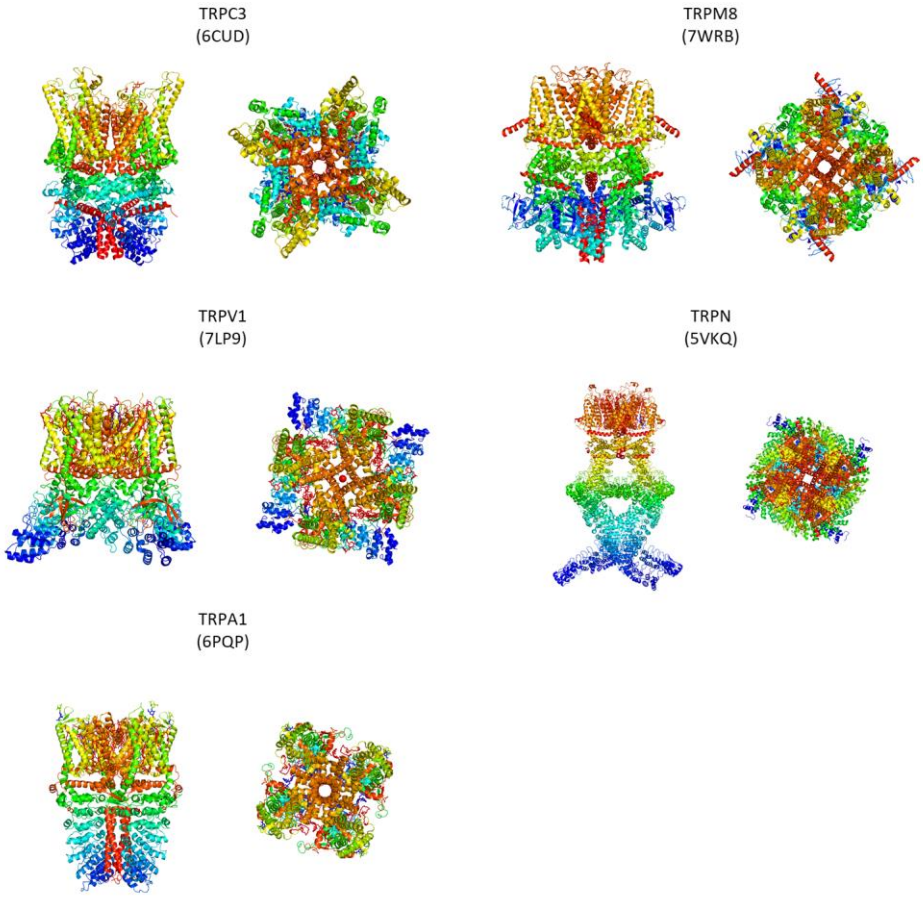


Figure 2: Structural differences between TRPs. Different TRPs are very similar in the transmembrane region (topmost part of each structure), but vary a lot on the cytosolic side (lower part of each structure). Left: side view, extracellutlar side up. Right: top view from the extracellular side.

TRPA1

One of the TRP families that has gotten a lot of attention recently is the TRPA family. In mammalians, only one single member exists – TRPA1. The A in TRPA stands for Ankyrin Repeat Domain (ARD), the most obvious characteristic of the structure, which is made up of a series of motifs called ankyrin repeats, that form into a long and flexible, yet highly ordered, tendril-like shape. The repeats themselves are made up of a simple helix-loop-helix, but together the helices stack up to form a curvature with the inter-connecting loops protruding like grasping fingers (figure 3). The motif was originally described in the ankyrin protein, from whence its name stems, but it is present in a diverse range of proteins (11).

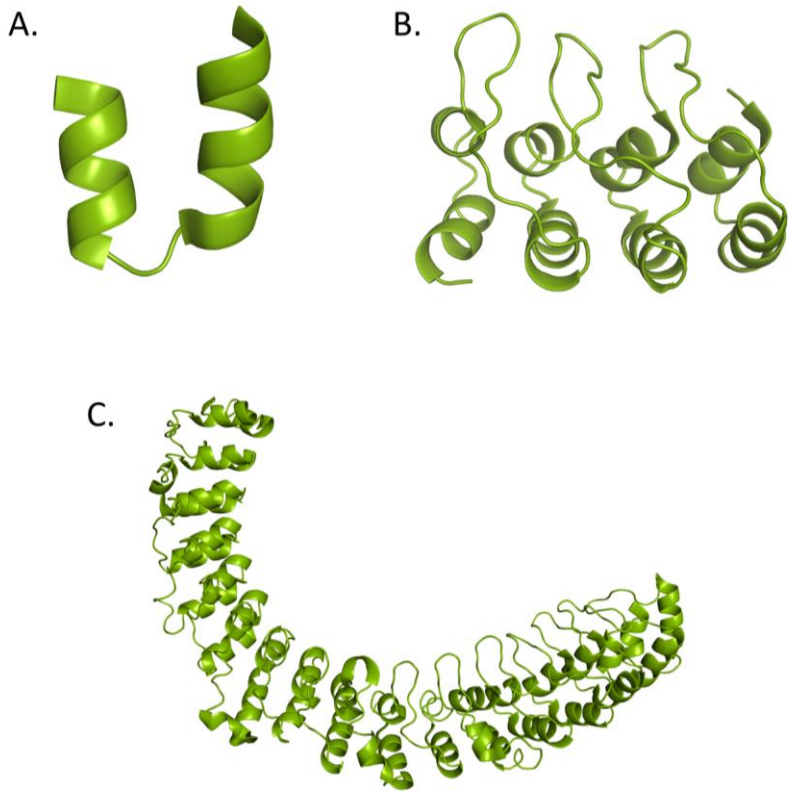


Figure 3: Ankyrin repeats. A. a single ankyrin repeat. B. The way ankyrins arrange with inter-connecting loops. C. The ARD of TRPA1 formes a curve.

TRPA1 was first encountered and described as a TRP by Jaquemar et al. in 1999 (12), but its identity was truly unveiled when Story et al. described the cold temperature activation (13), and Jordt et al. described the sensitivity to pungent chemicals such as mustard oils and cannabinoids (14). Since then, TRPA1 has been well studied, on many levels. The number of compounds discovered to affect the activity of TRPA1 – agonists, antagonists, and other modulators – has exploded, leading it to be known as a promiscuous receptor (15). The temperature activation of TRPA1 is almost as many facetted as the chemical activation, where orthologues of the channels are activated by cold and heat of various levels (16).

With the advent of cryo electron microscopy, the structure of human TRPA1, along with many other previously unsolved proteins, was finally determined in 2015 (17, 18). Many more structures with different ligands soon followed (19-23), and in 2023 the structure of TRPA1 from *Drosophila melanogaster* was determined (24). This structure finally shed some light on an annoying drawback of the previously

published structures, namely that the eponymous ARD could not be resolved. This is important, because the role of the ARD in temperature activation is disputed (15, 25, 26).

Other TRPAs

As mentioned above, TRPA1 is the only member of the TRPA family that is present in mammalians, however, in insects, the story is a completely different one. The most common ones are Painless (27), Pyrexia (28), Water witch (29), and TRPA5 (30), but not all of them are present in all insect species (figure 1). The exact role and structural relationships of these channels remains to be determined, and in **Paper III**, we describe the thermosensing properties of TRPA5 from *Rhodnius prolixius*, as well as investigate the relationship between the structures of the TRPA family.

The usefulness of pain

What is pain?

To survive in a hostile and harsh world, life has evolved a multitude of defensive mechanisms, one of the most important is the ability to detect and react to harmful conditions: pain. Pain is in an anatomical sense a way for certain nerve cells – nociceptors – to register and communicate the presence of intense levels of heat, mechanical stimuli, or chemicals. The purpose is to send a signal, resulting in a response that can protect the body from harm. Many TRPs are present in nociceptors, and are therefore part of a complex system where they play various roles. They react to direct stimuli, such as intense temperatures or pungent chemicals, but are also involved in phenomenon such as persistent pain and inflammation. (31)

Insects

Pain responses is a powerful factor that is often something that medical therapies tries to control and mitigate. However, in certain circumstances, triggering the receptor proteins of nociceptors can be a desirable feature. For example, repelling biting and stinging insects is an objective that could benefit human health, as well as economic interests. TRPA1 has already been shown to be responsible for the repellent effect of catnip towards mosquito (32) and plant-derived essential oils towards red flour beetle (33).

Figure 4: Pine weevil. Photo: Erik Karits.

Hylobius abietis

Another insect that has long been pointed out as a notorious pest is the large pine weevil (*Hylobius abietis*) (figure 4). *H. abietis* targets conifer seedlings, and are therefore extra harmful to the use of clear cutting in modern forestry. Historically, various pesticides have been used with great effectiveness, but with rising concerns for human health and environmental protection, all of the currently used compounds have been or are soon to be banned (34-36). The need for new methods to protect seedlings are therefore preciously needed, and in **Paper I** and **II** we investigate TRPA1 from *H. abietis* (HaTRPA1) to help in the understanding and possible development of future repellants.

Aquaporins

In all niches of life, water is one of the most important substances due to its very specific properties. However, the same properties also make it one of the most enigmatic, and understanding how it works is far from easy. A major riddle was for a long time how the impermeable nature of the lipid bilayer, could coexist with the

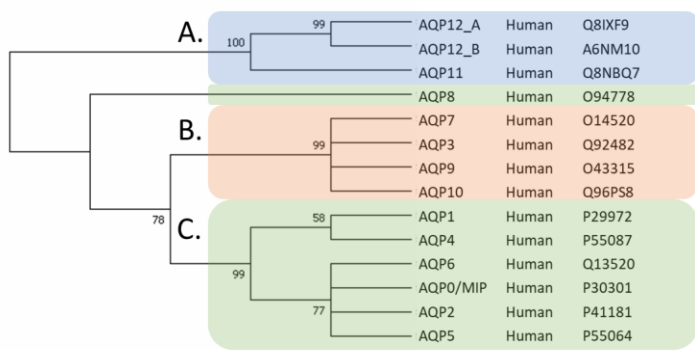


Figure 5: Phylogenetic tree of AQPs. A. Unorthodox aquaporins. B. Aquaglyceroporins. C. Orthodox aquaporins. To the right are shown accession numbers for the sequences.

apparent diffusion of water and solutes in and out of cells. It wasn't until 1992 that the group of Peter Agre discovered the existence of aquaporins (37), a family that in humans is made up of 13 water channel proteins (figure 5), that the puzzle got an explanation. All aquaporins are tetrameric channels with each subunit containing six transmembrane α-helices, forming a separate functional channel. It has two constrictions, the first being selectivity filter, where protons are excluded by the positive charge of an arginine. Further down the pore is the signature NPA (asparagine-proline-alanine) motif, that also limits the movement of protons, by preventing the Grotthuss mechanism (38, 39). The Grotthuss mechanism, lets protons be chained along from water molecule to water molecule, and could therefore effectively let protons pass the selectivity filter. Without the NPA motif, the proton gradient would be disrupted. Water transport is an essential functionality, and it is therefore not surprising that aquaporins are involved in a large range of diseases, from cancer to renal diseases (40). Superficially, aquaporins share several features with TRPs, such as the six transmembrane helices, and a half helix next to the selectivity filter. However, the role of these half helices is rather different. In AQPs they repel cations, whereas in TRPs they stabilize the passing cations. A notable difference is also that TRPs tend to have larger parts outside the membrane compared to AQPs, making them more difficult to handle.

Structure

Secondary, tertiary and quaternary structure

In its most fundamental aspect, a protein is just 20 different amino acids combined in a myriad of ways to form a polypeptide chain. However, the complexity of how these amino acids work together is not limited to a single dimension. The three-dimensional fold of the polypeptide chain, along with other peptide chains, is what makes the true magic of proteins come true (41). The order of the amino acids is typically referred to as the primary structure, whereas the interaction between neighboring amino acids is known as secondary structure. The secondary structures are α-helices, β-sheets and loops, and together they are folded into the tertiary structure. The tertiary structure is essentially the way the entire polypeptide chain folds into a complete structure, with interactions between secondary structures, or with the surrounding solute. Finally, when the need for more complex structures arises, several polypeptide chains can arrange themselves into complexes with a quaternary structure. The subunits in the quaternary structures can all be the same – homomers – or different in any combination – heteromers. For a single protein, there is usually a single correct fold that grants the desired functionality, but even if this fold typically is the most favored, misfolded – denatured – proteins can be hard to return to their native configuration. (42)

Why do you want to determine a structure?

How a protein works and interacts with its environment is largely determined by its structure. It is therefore understandable that in order to properly study proteins, one needs to study their structure. The most immediate thing a protein structure can tell us is perhaps *who* it is. By who, I mean where in a taxonomical context it fits in, and what other proteins it is related to. As protein functionality is dependent more on the structure of the protein than its amino acid sequence, the protein structure tends to be conserved between related proteins, and structural determination can therefore be used as a way of classifying proteins (43-45). Of course, it is also possible to predict protein functionality and ligand interaction from a protein structure, and this could be helpful in a context where experimental methods are too expensive or time-consuming to pursue. It is also useful in the pharmaceutical industry, where alteration or elimination of a certain protein activity is often the goal. The use of protein structures has greatly aided various methods such as structure-based drug design, where large data bases can be screened using computational methods (46).

History from crystals to neural networks

For many years, X-ray crystallography was virtually the only method to determine protein structures. By locking protein molecules in a strictly regular crystal lattice, the nature of diffraction can be exploited to determine the position of every atom in the protein to a resolution down to a few or even below one Å (47). Although it is a powerful technique, it has some limitations. In order to produce light of the right intensity, a synchrotron source is required, and the fact that the protein needs to be crystallized – something that especially membrane proteins do not like – can be a huge hurdle (48).

If you can make do with less detailed information, there are other methods to choose form. Circular dichroism (CD) is a method that utilizes the inherent chirality that all proteins share. By measuring how much left and right hand circular polarized light is absorbed, the secondary structure content of the protein can be determined and, in some cases, it is even possible to see some tertiary structure. If CD is performed at a synchrotron source (and therefore known as SRCD – Synchrotron Radiation CD), access to shorter wavelength information is possible, along with higher signal-to-noise ratio. A big advantage to CD is that protein samples in solution are possible to study, which is much closer to native conditions. (49)

The final experimental method I would like to bring up is Cryo-Electron Microscopy (Cryo-EM). The method of freezing protein samples in a thin layer of vitreous ice, and then capture movies of individual particles using an electron microscope, has gained in popularity over the last few years, in something known as the resolution revolution. The reason for this is a combination of the grid preparation methods becoming more and more sophisticated and the availability of high energy electron microscopes increasing. However, the most important factor here is the development of computational power and clever software that help in analyzing the thousands of micrographs that are generated over a single data collection. To the untrained eye, picking out particles in a noisy micrograph is much akin to finding a needle in a haystack, but with improved algorithms and capability to handle larger datasets, Cryo-EM has come to rival X-ray crystallography as the number one structure determination method. Advantages such as not needing to crystallize the protein is very much welcome to biochemists working with membrane proteins, and the possibility to gather several different conformations or states from a single sample is very tempting. Cryo-EM is especially good for larger proteins. (50)

In 2020, the protein modeling competition CASP14 (Critical Assessment of protein Structure Prediction) was taken by storm when DeepMind's AlphaFold2 outmatched all other competitors by leagues. AlphaFold2 is a neural network-based protein structure prediction model that can, often with high accuracy, return a complete 3D-structure by imputing a protein sequence. It was shortly after made available to the public, with a promise to revolutionize the field of structural biology. Previously, methods such a homology modelling had been useful in certain

situations where closely homologous protein structures are already available. AlphaFold, on the other hand, performs well even when applied to less well studied protein families (51). Still, the need for experimental methods exists, in particular for protein complexes, ligand interactions and multiple conformations.

Producing membrane proteins

Cloning and expression

Heterologous expression

The fundamental explanation that the central dogma gives us to how transcription and translation are connected comes with some interesting implications. A given genetic code (with some exceptions) leads to the production of a given protein, and remarkably, this is true even across different species. The codons that translate into a certain amino acid sequence in a human ribosome, would do the same in a bacterial ribosome and vice versa. There is, in other words, a (nearly) universal language of life!

In nature, viruses – although generally not considered to be alive – use this common ground to hijack the translational machinery of their host, and thereby become able to proliferate. In the era of molecular biology, we humans have through our deepened understanding become able to exploit protein expression in a similar way. Genetic manipulation opened the possibility to tame cells from different organisms and turn them into factories producing alien proteins in abnormal quantities. This is called heterologous expression.

Pichia (*Komagataella phaffii*)

Any protein scientist that aims to overexpress their protein of interest needs to ask themselves what expression host to use. Many times, the easy road of prokaryotes is chosen, as organisms such as *E. coli* are straightforward to genetically manipulate, and can quickly be grown to high densities with little requirements for costly equipment. However, expressing eukaryotic proteins in prokaryotes brings with it a range of problems such as incorrect folding, translocation and posttranslational modifications. (52)

The obvious solution to these problems is to pick an organism much closer to the original organism. Indeed, there are expression systems for insect cells or mammalian cells, even human cells, that of course mimic the conditions of the natural expression much closer, but the problems faced here are instead that the cells are difficult to handle and require expensive setups to thrive (53).

The middle ground exists in the form of yeasts. These single cell organisms are reasonably easy to grow in a liquid culture, while still having the advantage of being eukaryotic organisms. When dealing with membrane proteins, *Pichia pastoris* is the most commonly used yeast due to its excellent properties. *P. pastoris* is a methylotrophic yeast that can use methanol as its sole carbon source, making it useful to combine with an AOX (alcohol oxidase) promoter. This makes for a tightly controlled system with high expression levels, while still maintaining the folding and trafficking necessary for eukaryotic membrane proteins. (52)

Transformation/electroporation and zeocin selection

Transformation is the name used for the method of introducing of alien genes into a cell. It stems from a natural process by which bacteria take up genetic material from their environment, but is used as a term for any type of method that puts DNA into a cell. Typically, this is done by shocking the cell-membrane to make it permeable to the DNA molecules, using methods such as heat-shock or electroporation. For *Pichia*, there is a plasmid called pPICZα, which has the capability of introducing a gene of interest into the chromosomal DNA of *Pichia* (54). The goal here is to get a high copy-number of the gene, which corresponds to a high expression of the protein of interest. A clone with a high copy-number is typically referred to as a jackpot clone, and a common way of identifying such a clone is by screening with a selection marker, such as zeocin resistance. In general, higher levels of zeocin resistance corresponds to higher levels of expression (55).

Fed-batch cultivation

When purifying and analyzing proteins, the quantity needed is often far greater than what would be found in nature. A single cell of the expression host can be turned into an effective machine if transformed with an effective promoter and multiple copies of the gene of interest, but for a large and dense culture to work as a unified protein factory, the growth conditions need to be perfectly balanced at all times. To this end, a fed-batch bioreactor is a valuable tool for any project with a large need of protein. A bioreactor typically has the power to keep pH, temperature, dissolved oxygen, carbon source, and inducer constant, all while the density of the cell culture and the amount of protein increases (figure 6). (54)

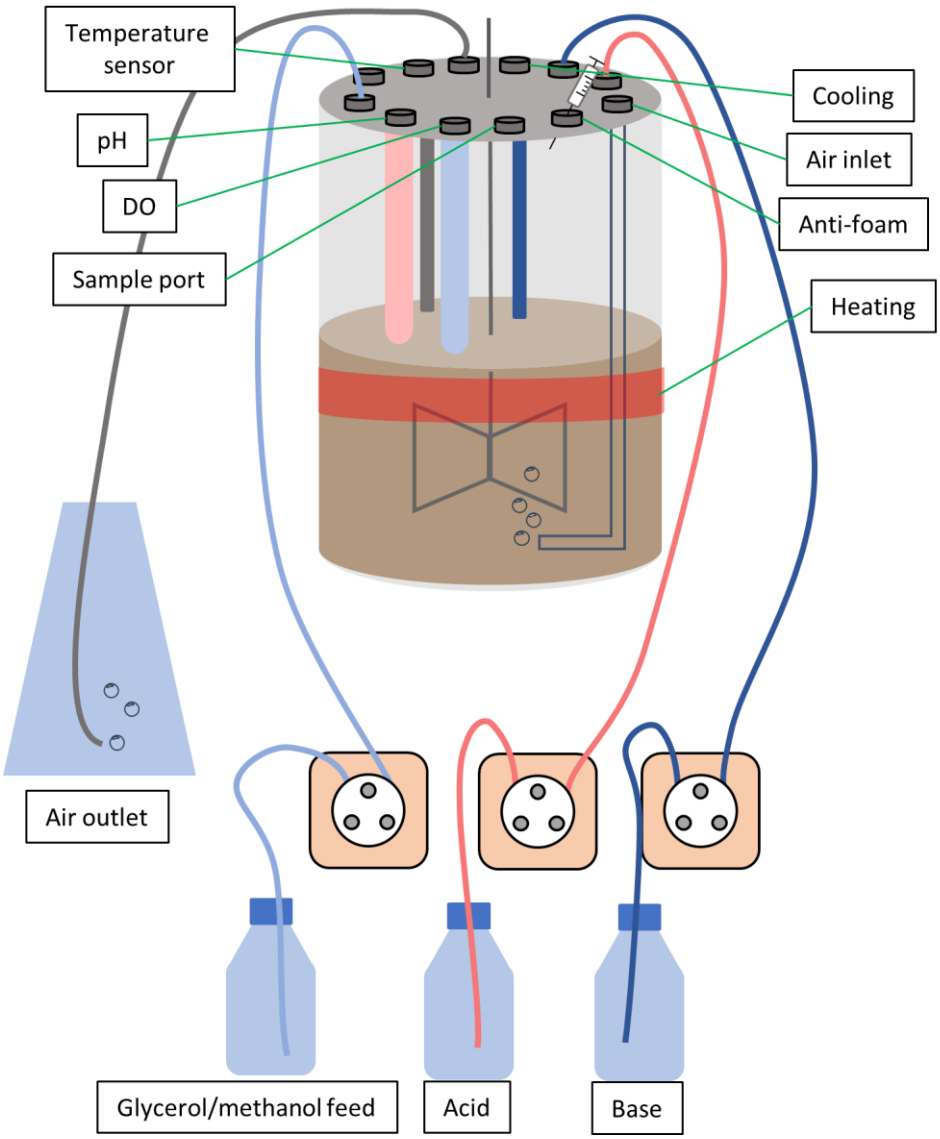


Figure 6: Fed-batch bioreactor.

Solubilization and selection of detergent

Screening for detergents and solubilizing a protein

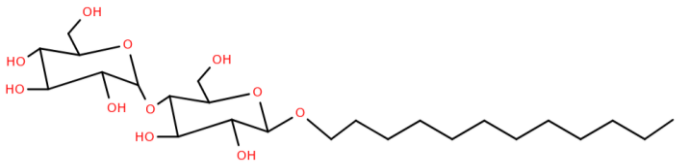
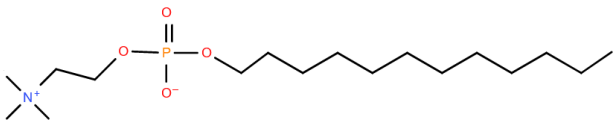
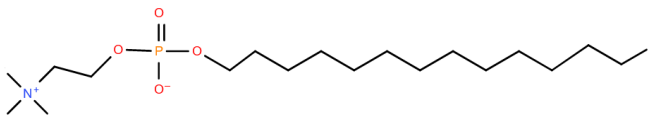
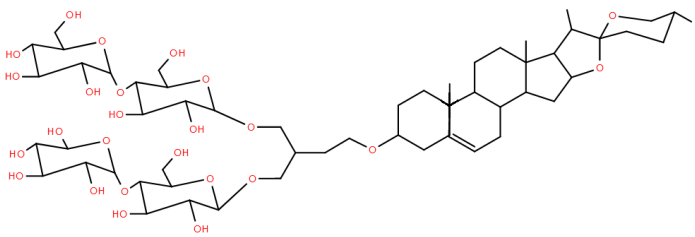
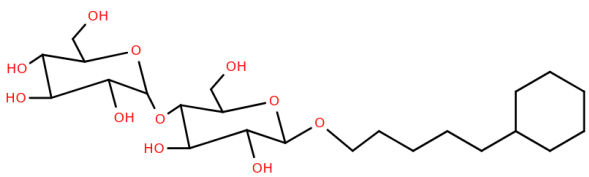
Due to the hydrophobic nature of integral membrane proteins, the fatty nature of the cell membrane must be mimicked to enable the study of the protein in solution – which is important for most in-vitro methods. Detergents make up the most common answer to this problem, as they are simple amphipathic molecules that spontaneously form micelles in an aqueous solution, wherein the transmembrane region of the protein can be stabilized. There are lots of detergents available on the market, with novel ones being released still (56). Detergents are typically classified based on what groups they are made up of, and if they are ionic, zwitterionic or nonionic. In table 1 are listed some detergents that I have used or considered using in my research. DDM and CYMAL-5 are commonly used detergents for membrane protein research due to them being generally effective, but still not so harsh as to draw out unwanted proteins, or even denature the protein of interest (57). The fos-cholines are less commonly used, as they are not as well balanced in that aspect. However, if your protein of interest is impossible to solubilize in any other detergents, fos-cholines might solve your problem (58). Finally, I want to point out one novel detergent known as GDN. It is an engineered version of digitonin that has been successfully been used for different membrane proteins (56), among them TRPA1 (24).

The actual process of solubilizing membrane proteins requires that crude membranes are treated with a high amount of detergent to draw out the protein. With the protein held in the micelles, insoluble parts are pelleted in an ultra-centrifuge. In later stages of the purification, lower concentrations are enough to keep the protein stable. Picking the right detergent is very important, as the wrong choice can either end up not solubilizing the protein at all, or possibly denature the protein. In **Paper I**, I performed a detergent screen by solubilizing small volumes of HaTRPA1 in different detergents and pelleting the insoluble fractions using an airfuge. The protein content was then estimated using western blot.

Picking detergent for Cryo-EM

An important aspect when designing the purification process for a membrane protein is what the sample will be used for. As I was aiming to do a structural determination with Cryo-EM (**Paper I**), any detergent would not fit the bill. A detergent that is good at getting the protein out of the membrane, is possibly not too good at keeping in solution. There are also problems when reaching the microscopy, as detergents can block the signal or make grid preparation harder. (59)

Table 1: List of detergents.

Detergent	Type	Ref
 DDM (n-dodecyl-β-D-maltopyranoside)	Non-ionic	(57)
 Fos-choline 12 (n-Dodecyl-phosphocholine)	Zwitterionic	(58)
 Fos-choline 14 (n-Tetradecyl-phosphocholine)	Zwitterionic	(58)
 GDN (glyco-diosgenin)	Non-ionic	(56)
 CYMAL-5 (5-Cyclohexylpentyl β-D-maltoside)	Non-ionic	(57)

Purification

IMAC

To separate out the protein of interest from the cell lysate, the use of IMAC (Immobilized Metal Affinity Chromatography) is common. Immobilized metal ions, such as nickel, binds covalently to recombinant proteins with a poly-histidine tag and can therefore be used to separate those proteins from other proteins by binding it to a column and then eluting the protein of interest with imidazole.

SEC

SEC (Size Exclusion Chromatography) is a powerful method that can separate proteins based on size. The sample passes through a porous material that prolongs the retention time for smaller particles, whereas larger particles experience a smaller volume, shortening the retention time. This is useful, both to separate out smaller and larger particles from the sample, and to verify the protein's identity and multimeric state.

Saved by GFP

GFP (Green Fluorescence Protein) is a 28 kDa protein with the ability to emit light at a wavelength of 508 nm that was discovered by Shimomura et al. in 1962 (60). It is used as a signaling molecule for expression, or as a tag for other proteins, due to its stability and intense inherent fluorescence. The structure of GFP is a single beta-barrel with an alpha-helix in the middle, and it is a soluble protein – which is important to consider when linking it to a membrane protein. (61)

Constructing a new clone

GFP has been successfully used to tag a membrane protein expressed in *P. pastoris*, and is useful when optimizing expression and purification (62). It has also been used in *Saccharomyces cerevisiae* to produce membrane proteins (63, 64). In my project, a C-terminal GFP-tag was used to investigate and improve the expression of HaTRPA1 in **Paper II**. It was useful to be able to do an initial screening of the expression levels on plates, but also to follow the expression in liquid cultures using flow cytometry and fluorescence microscopy. This in turn gives important information regarding how cell viability and expression levels are related, but also helps to distinguish localization patterns in different constructs.

New doors are opened

Plate screening

After a transformation, one hopefully has a lot of transformants of varying quality that somehow need to be boiled down to one or a few so called "jack-pot" clones. When transforming *P. pastoris* without a replicating plasmid, the goal is to integrate the plasmid into the chromosome, and by getting several copies of the gene into one single cell, the rate of expression can be maximized. A common way to screen for high copy numbers is by plating the transformants on very high levels of antibiotics, thereby causing a high selection pressure that only highly expressing clones can survive. The disadvantage of this method is that it takes a lot of time, and doesn't always give a very accurate prediction of the copy number (55). Antibiotics for eukaryotic cells are also a safety issue as they, per definition, are toxic also to humans. The introduction of a GFP tag opens the way for screening by inducing the clones, directly on plates. By a simple fluorescence scan, the expression levels can then be estimated quickly. In **Paper II**, I employ this method to improve my yields of HaTRPA1. However, this method on its own is not enough to single out the best clone, and therefore the expression has to be evaluated in further steps.

Flow cytometry

An advantage when a GFP tag is used, is that the protein expression levels can be monitored in live cells throughout the expression. Flow cytometry is a method where single cells are passed through a detector, one at a time, and various parameters are measured (figure 7). The result is that a number of events are recorded, that in the ideal case each corresponds to a single cell passing the detector. Among the most common data that is recorded is the forward scattering (FSC) and side scattering (SSC). In rough terms, FSC relates to the size of the cell, whereas the SSC relates to the granularity or shape of the cell. One must, however, keep in mind that there are other factors that can affect these metrics, and they cannot be used to measure the exact size of the cell or its features (65). In addition, a flow cytometer is typically equipped with several channels of fluorescence emission filters that can be used to simultaneously measure the content of different fluorophores. An obvious example of this would be GFP, but there are several fluorescent dyes that can be useful in different situations. Propidium iodide (PI) is a dye that emits red fluorescent light, only if it penetrates damaged cells, and can therefore be useful to estimate cell viability (66).

To make use of all the measured parameters, the strategy of gating is employed. The basic idea is that events that fall outside a given range for some parameter, should be excluded from the analysis. As an example, FSC and SSC is typically used to distinguish cells from noise, or even to distinguish between different cell types in a sample. Further, the PI signal can be used to only include intact or viable cells in the analysis. By layering these gates, very specific populations within a cell sample can be distinguished (figure 8).

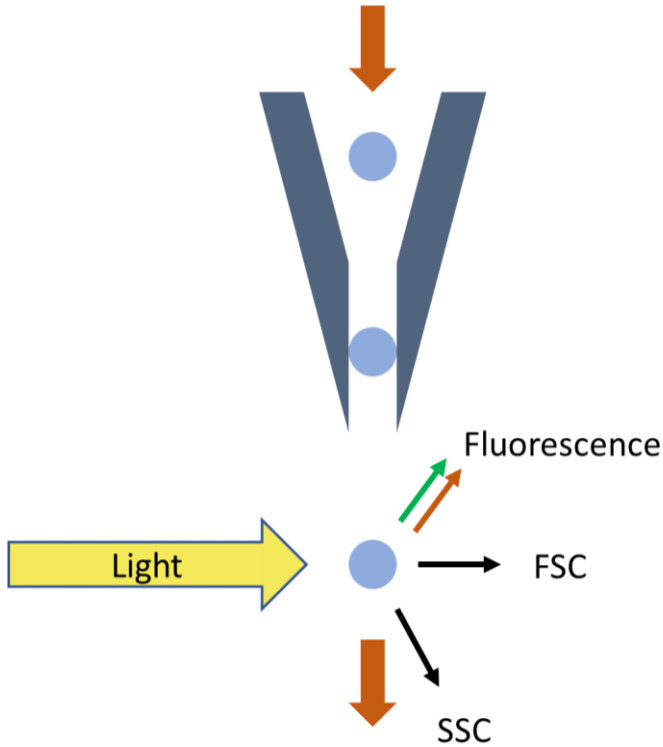


Figure 7: Principle of flow cytometry. Cells pass a detector, one at a time.

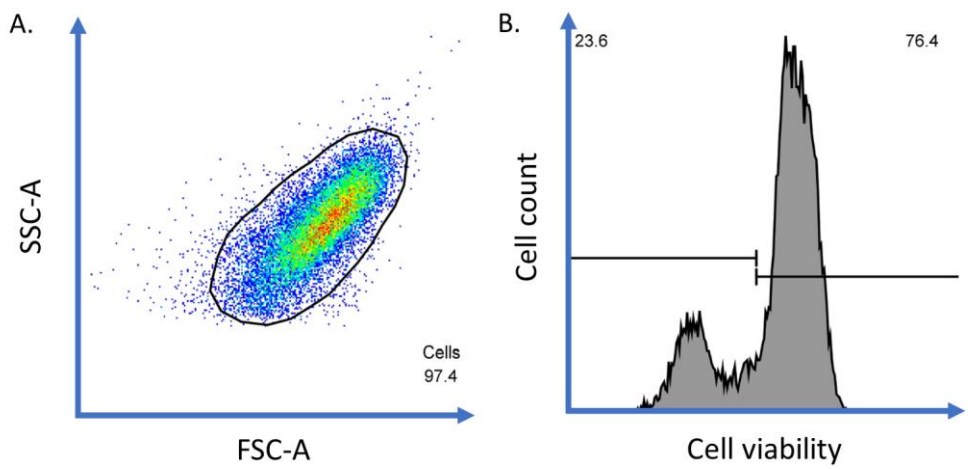


Figure 8: Example of gating of flow cytometry data. This figure is constructed from the data set in **Paper II**, but using a different clone from the figure in the supplemental data. The numbers indicate the percentage of registered events within the gates, marked by black lines. All axis are arbitrary units, except cell count that has number of cells.

Fluorescence microscopy

To review the GFP expression in cells, fluorescence microscopy is a useful method. It is based on the possibility to use a microscope, with a light source and a filter for a particular excitation wavelength, to observe fluorescence in a cell sample. By overlaying the fluorescent signal with the phase contrast signal, the localization and quantity of GFP in cells can be estimated. A common method of preparing the samples for fluorescence microscopy is by casting an agarose pad on the microscope slide, and then adding a small amount of cells. This prevents drying of the sample, and also makes all the cells stay at the same focal distance. This method is most suitable for larger cells, such as yeast cells, and it is therefore compatible with the *Pichia* cells I am using. (67)

Fluorescent gels

Another very straight forward method of evaluating the protein expression is by running the samples from a fermentation in an SDS-PAGE (Sodium Dodecyl Sulfate Poly Acrylamide Gel Electrophoresis), and then exploiting the GFP tag to distinguish the protein of interest. SDS-PAGE is a very common method in protein science that is used to separate proteins based on size. The protein is denatured in the detergent SDS, before being loaded into a poly acrylamide gel where a current is applied. SDS confers charges to the protein that will migrate further the smaller the protein is. Typically, the gel is then stained with a protein binding dye such as Coomassie to show protein content as bands, but with a fluorescence gel, only bands with GFP content will be visible. It is, however, important to note that membrane proteins have been shown to migrate at unexpected speed due to the detergent binding, and therefore bands do not always show up at the expected position on gels (68).

Cleaving the GFP tag – when to do it?

Since a GFP tag is large and likely to disturb further analysis, it is a good idea to remove it during the last purification steps. For this purpose, a TEV protease (Tobacco Etch Virus protease) specific sequence can be included in the protein construct. By incubating the protein sample with TEV protease, the protein is cleaved with high specificity at that site (69), whereafter the GFP with His-tag are fished out using a reverse IMAC, which also improves the purity of the sample. Choosing the right moment in the purification to do this is important. Removing the tags too early, and you lose the advantage of being able to follow your protein easily. Cleaving them too late and they might not be removed effectively from the sample.

To determine a structure

Experimental determination of protein structures is a resource intense challenge. Once a sufficient amount of protein of good enough purity has been acquired, there is still no guarantee that the structural analysis will work out as planned. Still, it is often worth the effort, due to several reasons. Although computer-based prediction models have gotten better in recent years, they are still not good enough in several circumstances, and fundamentally, the predictions are based on training on experimentally determined structures, without which they wouldn't be possible. In my project, determining the structure of HaTRPA1 was an attractive goal because of the many possibilities it would unlock. For one, it would be desirable to perform docking studies with ligands or other proteins, something that is highly dependent on a reliable structure. The goal was also set on determining the structure along with ligands to investigate binding and agonist activation.

The choice of method then quickly fell on Cryo-EM. Only a single X-ray crystallography structure of a TRP exists to date (70), and in 2024 the number of TRP structures in the PDB (Protein Data Bank) is over 350 (71), which is a testament to the difficulty of crystallizing this protein group. It can also be noted that the first ever membrane protein to be determined to atomistic resolution using Cryo-EM is TRPV1 (72), and generally it is the method of choice for structure determination of TRPs.

Solving the TRPA1 structure with Cryo-EM

Cryo-EM is an excellent tool, but not for all projects. One of the first considerations to make when determining if Cryo-EM is a fitting method for a project is the size of the protein, as there is a lower limit around 100 kDa below which it becomes hard to identify the single particles in the micrographs. It is possible to circumvent this problem by increasing the protein size, for example by attaching larger proteins such as legobodies – a scaffold that increases the size of a nanobody, that in turn is attached to the protein of interest (73). The problem with this method is of course the need for a specific nanobody that can bind your protein, and the risk that the protein structure is in some way affected by the binding. Another limitation when it comes to Cryo-EM is the fact that the sample has to be frozen on grids. This can of course not be considered to be native conditions, and some proteins will have

preferred orientations or adhere to the grids, making particle picking and subsequent analysis difficult. There are of course advantages to Cryo-EM as well, like the fact that multiple conformations of the same protein can be resolved from the same dataset. Compared to X-ray crystallography, membrane proteins are far easier to work with in Cryo-EM, as crystallization of membrane proteins is notoriously difficult. (74)

TRPA1 is a fairly large protein, with each monomer at around 130 kDa, and therefore it is no surprise that Cryo-EM has been the structural method of choice when determining its structure. So far, a total of 15 structures of TRPA1 can be found in the Protein Data Bank (PDB), whereof 13 from human and two from *Drosophila melanogaster*. The structures are resolved with various agonists, antagonists, and ligands, mirroring the promiscuous nature of TRPA1's activation. See table 2 for a complete list of the published structures of TRPA1. It is interesting to note that all the structures are solved using the amphipol PMAL-C8, except for the ones from Suo et al. that are solved in nanodiscs made using MSP2N2. The detergents used for solubilization are mild detergents such as DDM. This is a central question in **Paper I**, where I needed to find the optimal conditions for purification of HaTRPA1, considering that different TRPA1 orthologs do not necessarily behave the same way.

Table 2: All published TRPA1 structures. All are from human, except 7YKR and 7YKS that are from *Drosophila*. *No paper has been published along with this structure.

Entry ID	Resolution (Å)	Detergent (solubilization)	Stabilizer for Cryo-EM	Ligand	Ref
3J9P	4.24	MNG-3	PMAL-C8	None	(18)
6PQQ	2.81	Digitonin	Nanodiscs	None	(19)
6PQO	2.88	Digitonin	Nanodiscs	Agonist JT010	(19)
6PQP	3.06	Digitonin	Nanodiscs	Agonist BITC	(19)
6WJ5	3.6	FA-3	PMAL-C8	Antagonist GDC-0334	(20)
6V9V	2.6	CYMAL-5	PMAL-C8	Agonist BODIPY-iodoacetamide and Bound calcium	(21)
6V9W	3.1	CYMAL-5	PMAL-C8	Bound calcium	(21)
6V9X	3.3	CYMAL-5	PMAL-C8	Agonist iodoacetamide	(21)
6V9Y	3.6	CYMAL-5	PMAL-C8	Antagonist A-967079	(21)
6X2J	3.0	FA-3	PMAL-C8	Agonist GNE551	(22)
7JUP	3.05	FA-3	PMAL-C8	Antagonist A-967079	(23)
7OR0	2.64	*	*	Antagonist 2-60	*
7OR1	2.64	*	*	Antagonist 2-60	*
7YKR	3.2	DDM	GDN		(24)
7YKS	3	DDM	GDN		(24)

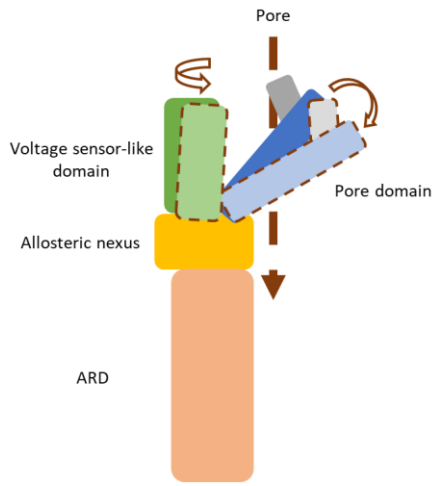


Figure 8: Opening mechanism of TRPA1.

When the structure of TRPA1 was first solved by Julius lab (18), a lot of what was seen was no big surprise as structures of other TRPs where already available and secondary structure predictions could identify major features (75), but a lot of what makes TRPA1 distinct was unknown until the structure became available. One of the most obvious features of TPR channels is the eponymous TRP domain. Given the name, one would assume that this domain is present in all TRP channels, but in several, like TRPA1, analysis of the sequence didn't predict such a domain to be present in TRPA1(76).

However, in the solved structure, a domain with similar structural topology was observed, and therefore dubbed the TRP-like domain. It is made up of a post-S6 helix, that interacts with a pre-S1 helix and a linker domain N-terminally from the pre-S1 helix (18), forming a region later confirmed to be the allosteric nexus, where electrophilic binding of agonists takes place (19). In the allosteric nexus of human TRPA1, the Cys 621 is situated, that is held to be the most important for electrophilic binding, but other cysteines have yet to have their role discerned. The fact that TRPA1 is lacking the canonical "TRP box" that is part of the TRP domain, but still has a domain with a similar structure and function, is talking for my earlier comment that structure generally tends to be more conserved than sequence (18) There has also been a comparison made by David E. Clapham to voltage-gated potassium channels, that share a lot of features with TRPA1 (17). In an extensive publication from 2020, Zhao et al. solve four different structures of TRPA1. One with the irreversible electrophilic agonist iodoacetamide (IA), one with the antagonist A-967079 (A-96), one with bound calcium, and one with bound calcium and a bulkier version of IA called BODIPY-IA (BIA). They focus a lot on the important pore of the channel, that typically is divided into an upper gate, the selectivity filter, and a lower gate, the canonical gate. By comparison of the structure with agonist, and the structure with antagonist, the dilation of both gates to allow ions to pass through the pore could be observed. They could also show that this opening of the gates was accompanied by structural changes throughout the protein, with the pore helices and the helices S5 and S6 shifting to accommodate the wider opening, but even the helices S1-S4 of the voltage sensor-like domain (VSLD) rotating 15° (figure 8). The VSLD is a domain that is structurally related to the voltage sensing domain of potassium channels, but few TRPs have been shown to

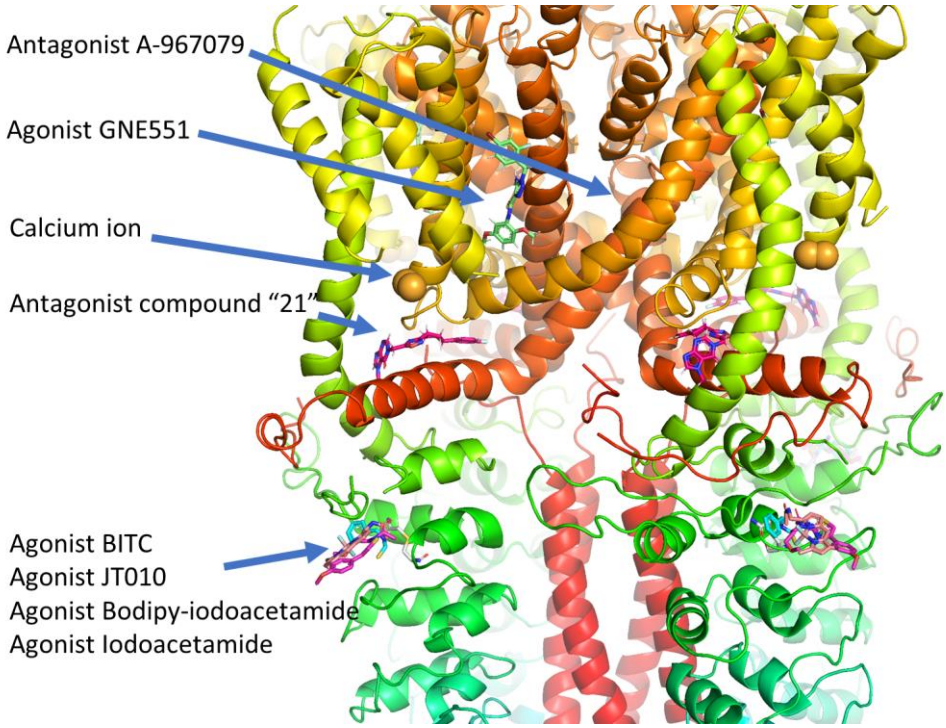


Figure 9: Ligand binding sites of human TRPA1.

have the ability to react to electrical stimuli (77). This is a large movement that involves most of the transmembrane part of the protein, and could be an explanation to how the distant binding of the ligands could affect the pore. IA binds mainly to C621 in the TRP-like domain, whereas A-96 binds in a nook of the S5 helix, which is quite far away (figure 9). It becomes even more complicated when considering the effect of calcium. TRPA1 is a cation channel, and Ca^{2+} can pass through when it is activated, but calcium also has both a potentiating and a desensitizing effect on pore conductivity when bound to a site between the S2 and S3 helices (figure 9). (21) In **Paper III**, the two gates of the pore are compared between other members of the TRPA family.

The resolutions reached are overall high, but most of the structures are lacking in one central aspect. The ARD is only partly solved, with a number of ankyrin repeats missing from the N-terminal (figure 10). When looking at the electron densities of some of these structures, a vague shape corresponding to these ankyrin repeats can be discerned, but it is not enough to model the entire protein accurately (18). The only exception to this is in the structure 7YKR from *D. melanogaster*, where the N-terminal can be seen to turn upwards in a sort of propeller shape (figure 11) (24). It is interesting to note that there is an interaction between a loop at the C-terminal, and the ARD. This, along with a large movement of 12-13 Å up and down of the

coiled-coil of the C-terminal that was observed between two different states suggests that the electrophilic activation of the channel may be tuned by the cytosolic part. Wang et al. also performed a truncation of the C-terminal that proved to be unresponsive to AITC, which further emphasizes the role of this region (24).

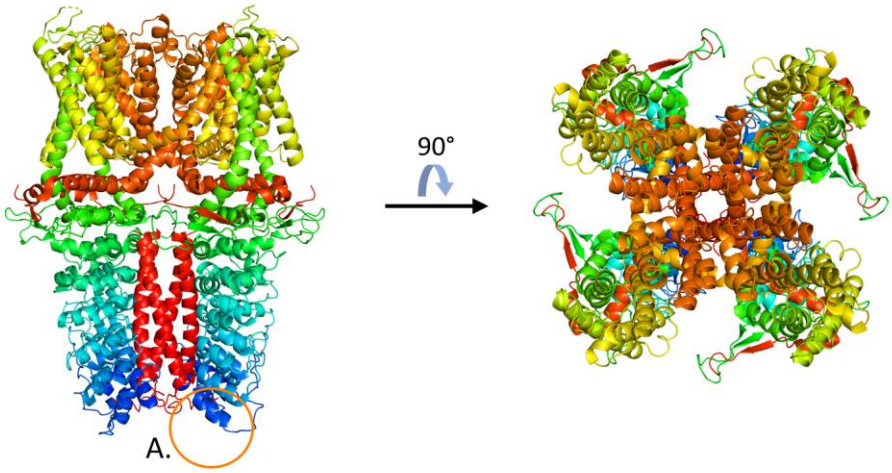


Figure 10: TRPA1 3J9P. A indicates the missing part of the ARD.

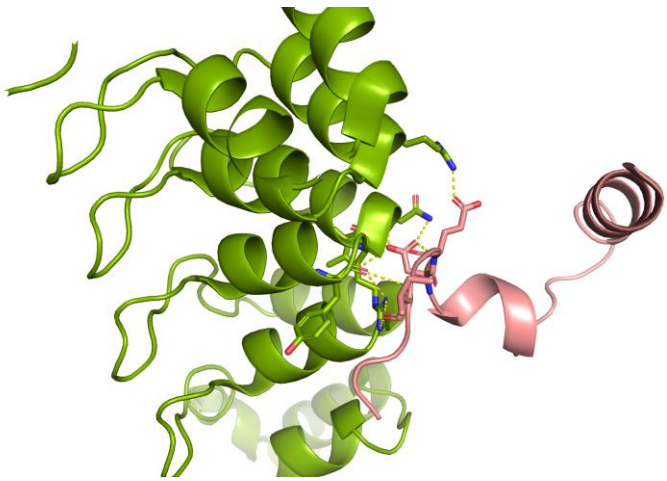


Figure 11: Interaction between the ARD (green) and the C-terminus of *Drosophila* TRPA1 (pink).

Membrane mimetics

As previously mentioned, an integral membrane protein is not stable in a water solution on its own, and therefore one generally has to make sure that there are detergents present at each step of the process. However, being solubilized in a detergent micelle is not really that close to the native environment that a membrane protein would be used to, as interaction with lipids often stabilize or affect the function of the protein (78). There are several membrane mimetics that are commonly used to study membrane proteins, among them are nanodiscs, amphipols, SMALPs (Styrene Maleic Acid co-polymer Lipid Particles), and saposins (79).

Nanodiscs

The introduction of nanodiscs brings a solution to this problem by allowing the protein to be moved from the detergent micelle, back into a lipid bilayer. The lipid bilayer is kept stable by an amphipathic protein, called a membrane scaffold protein (MSP), that forms a girdle around the hydrophobic part of a small circular lipid bilayer, creating a disc-shaped structure that can encapsulate the membrane protein (figure 12) (80). By adjusting the chain length of the MSP and number of subunits, the nanodisc can be made just big enough to fit a single membrane protein molecule, something that is essential for single particle Cryo-EM. Another important parameter to adjust is the lipid composition of the bilayer. Cell membranes contain a complex mix of phospholipids, and mimicking the relevant conditions, or fine-tuning them for a specific application, is important. The process of moving the protein from its detergent micelle to the nanodisc is straight forward in theory. The components for the nanodisc, MSP and lipids, are added to the protein sample, and then the detergent is removed, for example with Bio-Beads – a porous absorbent that binds the detergent, making it easy to remove in a centrifuge. (81)

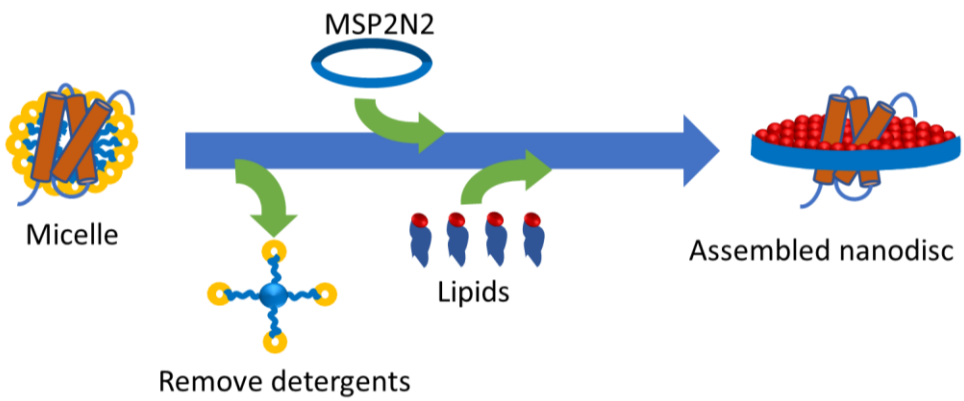


Figure 12: Assembly of a nanodisc.

A good idea to be able to separate out nanodiscs with the integrated protein is to have a tag, such as a His-tag, on the membrane protein to be able to fish out nanodiscs that contain the protein. It is also possible to separate the nanodiscs with SEC, as the nanodiscs with a protein molecule integrated will elute earlier due to their higher size.

As an example of how the process of integrating a protein into nanodiscs work, I will use our attempt at assembling nanodiscs of MSP2N2 with TRPA1 from the malaria mosquito *Anopheles gambiae* (unpublished data, Balder Werin and Veronika Tolevska). The protocol was based on one published by Suo et al. for human TRPA1, that was used to solve the 6PQQ, 6PQO and 6PQP structures (82). When assembling the nanodisc, it is useful to include sodium cholate to stabilize the lipids (83).

In figure 13, a SEC chromatogram is shown with two peaks, one most likely corresponding to nanodiscs with AgTRPA1 inside, and one corresponding to empty nanodiscs. The particles looked well separated and to be of the correct size on negative stain (figure 14), but when analyzing them using Cryo-EM, it proved impossible to find nanodisc particles with protein in them. Instead of optimizing the process to get the nanodisc reconstitution to work, we decided that we would use a detergent based method to minimize the number of purification steps, thereby simplifying the protocol.

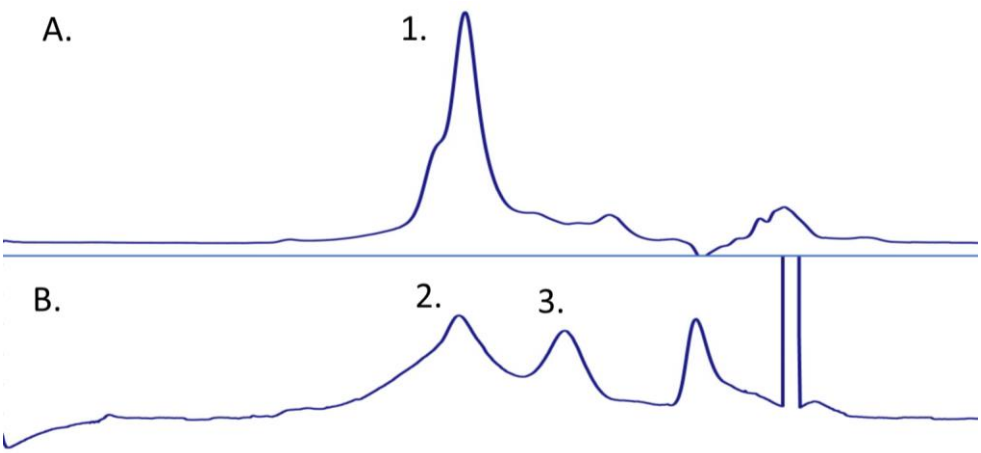


Figure 13: SEC chromatograms from nanodisc assembly. A. SEC chromatogram of truncated HaTRPA1 with a single peak (1). B. SEC chromatogram of truncated HaTRPA1 in nanodiscs (2), and empty nanodiscs (3). The first step of the process was to express MSP2N2 in E. coli, and purify it on IMAC with the His-tag. A TEV site was included that allowed the His-tag to be removed using reverse IMAC. A lipid solution in Tris buffer was prepared with the lipids POPC, POPE and POPG at a ratio of 3:1:1 with a total concentration of 10 mg/ml, and an addition of 26.5 mM sodium cholate and 3 times CMC Foscholine-12. A reconstitution mix of MSP2N2, AgTRPA1 and the lipid solution was prepared at a ratio of 1:3:200 and incubated at 4°C for 30 minutes before transferring the reconstitution mix to 50 mg Bio-Beads, to remove the detergent. After two incubations with Bio-Beads, the empty nanodiscs can be removed by IMAC, or straight away with SEC on a Superose 6 increase 10/300 GL (Cytiva).

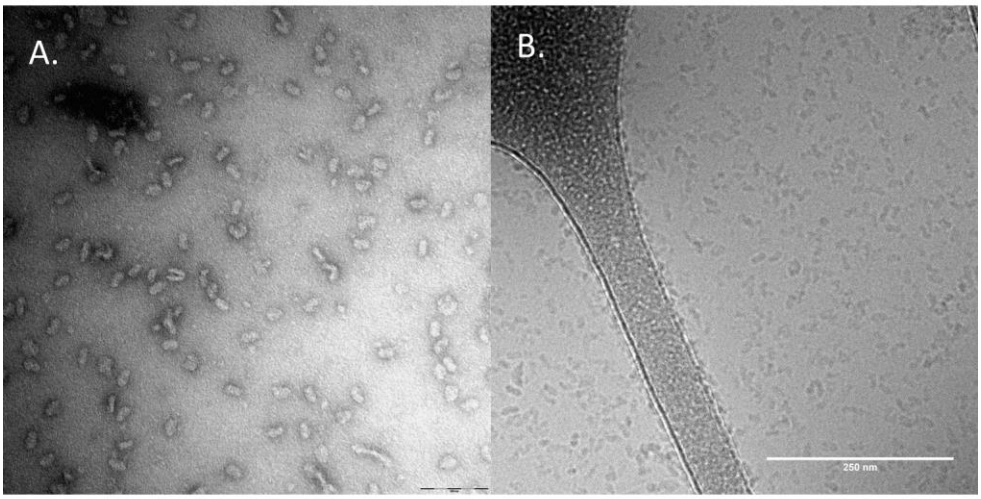


Figure 14: Nanodisc analysis results. A. Negative stain image with particles nicely distributed, although with some tendency to lump together. B. Cryo-EM image of the same sample, where a lot of the particles are seen to adhere to the carbon grid (the darker part).

Amphipols

An alternative to nanodiscs is the use of amphipols (84). Amphipols stands for amphipathic polymers, and is just that: polymers that cover up the hydrophobic parts of the membrane protein, but without additional lipids added. The switch from detergents to amphipols is very similar to the process for nanodiscs, but without the need to add lipids. While this makes them simpler to use than nanodiscs, the emulation of the lipid bilayer is still an advantage for the nanodiscs. However, most of the TRPA1 structures published to date have used the amphipol PMAL-C8 during the Cryo-EM step. An attempt was made to use PMAL-C8 with HaTRPA1, but without success (unpublished data, Oliwia Kołodziejczyk).

SMALPs

Somewhere in the middle we find the SMALPs, molecules that like the nanodiscs incorporate a lipid bilayer that stabilizes a membrane protein. However, unlike nanodiscs, the SMALPs can be assembled directly from crude cell membrane by, simply by adding the polymer and letting it cut out a small portion of the lipid bilayer, along with the protein of interest, thereby avoiding the need for detergent solubilization. The advantage of this method is of course that no need for detergents exists, and that the natural lipid environment is conserved, at least to some degree.

However, there are some problems later in the process, namely that more grid optimization is required, and doublets tend to form where two particles interact. (85)

Saposins

I will also briefly mention the existence of lipoprotein nanoparticles made using the protein family of saposins. The principle is very similar to SMALPs, but instead of a synthetic polymer, a protein, that is more flexible than MSPs, is used to stabilize the lipid bilayer around the membrane protein. (86)

Negative stain

Before taking a sample to Cryo-EM, it is common to do something called negative stain. It is also an electron microscopy method, but instead of freezing the sample, the protein is kept in solution. In order to be able to see the individual particles, a stain such as uranyl acetate is used to create contrast around each particle. These compounds typically contain heavy metals, like uranium, in order to create the contrast, but are because of this often subject to restrictions in their use. In short, the process of analyzing a sample goes like this: a carbon grid is glow discharged to make it more hydrophilic, causing the sample to wet the grid evenly. The sample is then applied to the grid, and lastly the stain is added, before analyzing the sample by transmission electron microscopy. (87)

Preparing grids for Cryo-EM

The process of grid preparation for Cryo-EM is slightly more complicated. To be able to distinguish individual particles, the sample must be frozen in a thin layer of amorphous or vitreous ice – ice lacking crystalline structure that is formed when a solution is frozen incredibly quickly. The sample is applied to a grid and excess liquid is blotted away before plunging the sample at extremely high speed into liquid ethane. All of this can be done either manually, or using a machine such as a Vitrobot, that performs these steps in a quick and reproducible way in a controlled environment. However, there are several challenges to this process, mainly in getting the protein sample to spread evenly within the ice film with the goal of having single particles without preferred orientation. To achieve this, parameters such as grid material or the usage of a support film can be varied. (88)

Processing the data

Once a dataset has been collected, the thousands of micrographs need to be processed. For this purpose, a software such as CryoSPARC (Cryo-EM **S**ingle Particle Ab-Initio Reconstruction and Classification) is often used (89). Usually, to extract the most information from a sample, short movies are captured for each micrograph that can be averaged together into a single picture with the use of motion correction. Next, a CTF (Contrast Transfer Function) estimation is done to take into consideration defocus and astigmatism. When it comes to picking out particles from the micrographs, it can be done either by manual picking, or by the aid of machine learning (90). Either way, the particles are then used to do 2D classification, meaning that classes are gathered with averages of each class shown as representations, and from these classes are then selected templates to extract new particles from the micrographs. Then another set of 2D classes are generated, and the process can be iterated if needed. When the final 2D classes have been generated, an ab-initio reconstruction followed by a homogenous refinement is performed to output the final 3D density, into which the actual structure can be modeled (figures 15 and 16).

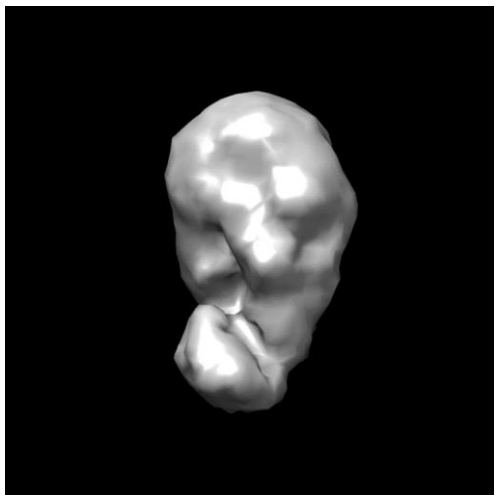


Figure 15: 3D model from negative stain data.

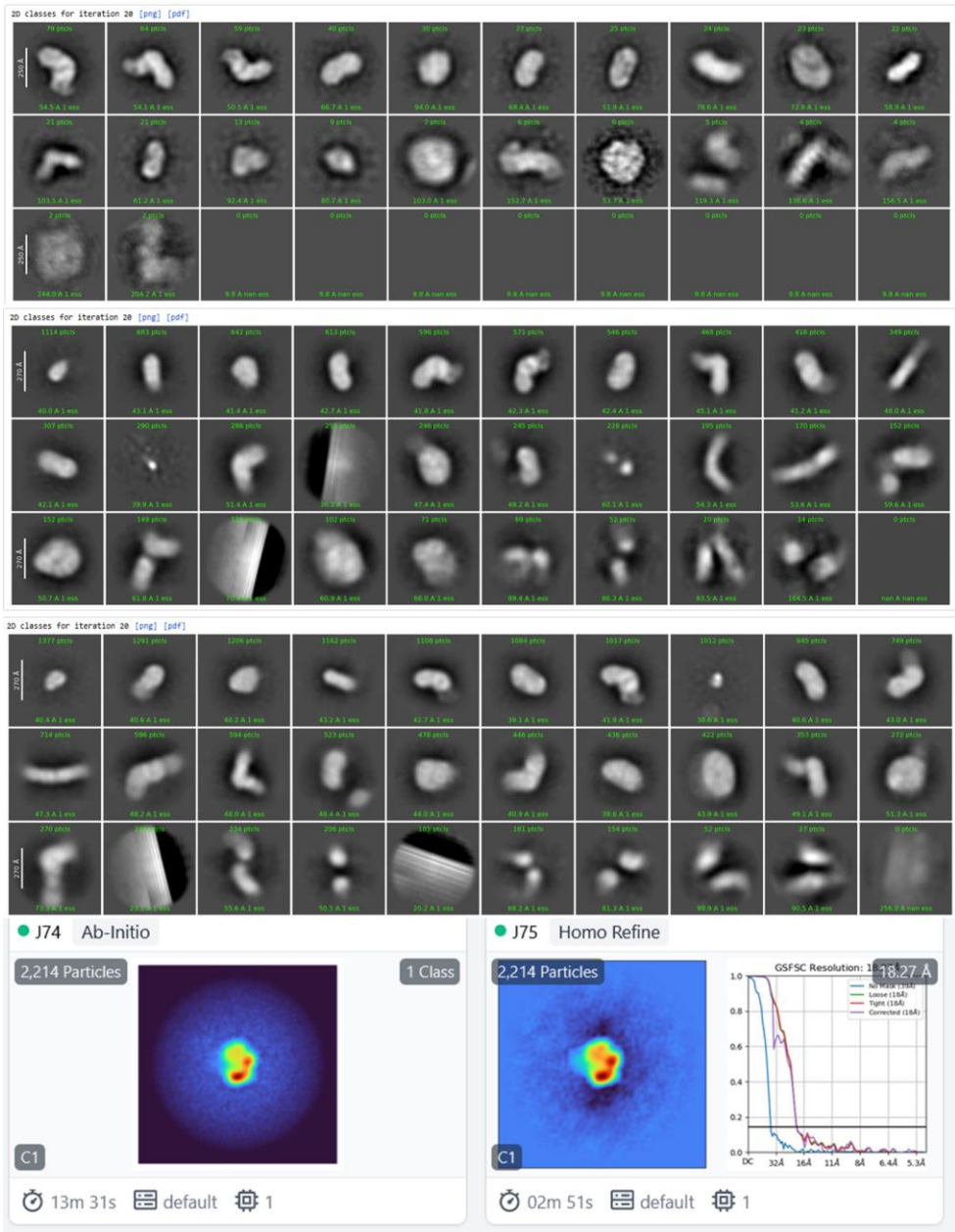


Figure 16: CryoSPARC analysis of negative stain data from HaTRPA1 in GDN. When analyzing the micrographs, one can either do manual picking (top), or use an automated blob picking (second from top). Either way, templates from the first particle picking are used to generate another set of 2D classes (third from top), that are used for Ab-initio reconstruction (bottom left) and homogeneous refinement (bottom right).

AlphaFold

AlphaFold is a deep learning system, that has been trained on a vast database of published protein structures, to be able to accurately predict the folding of any protein with a known amino acid sequence. The way AlphaFold works is by first making a Multiple Sequence Alignment (MSA), querying a large number of protein sequences. This is both to identify conserved regions, but also to find proteins with similar structure to identify pairwise contact or proximity between amino acids. This information is then passed through an iterative process called the Evoformer, that essentially tries to identify the most important pieces of information. The final step is to feed the information into a Structure module, which creates the three-dimensional image from the information extracted in the previous step by predicting the φ and ψ angles. And then the whole process is iterated to refine the structure further. (51)

As the code for AlphaFold2, which is the latest version, is open-source, it is now possible for anyone with access to the required computational power to use the program. If you do not have access to a computer with a high-end graphics card, or a computer cluster, it is possible to use the convenient service of ColabFold (Google) (91), which combines the sequence alignment of MMseqs2 with AlphaFold2 on a platform with access to a computer cluster that is ready and easy to use. In addition, it is also possible to input custom templates, and there is a possibility to perform a side-chain relaxation of the output structures using the AMBER software package.

AlphaFold, for all its power and ease of use, have some drawbacks. The first, and most evident one, is the immense need for computational power. Even with access to the best virtual processing units provided by Colab, larger proteins are not possible to fold in their entirety. An additional problem arises for protein complexes, or multimeric proteins, where AlphaFold struggles to predict interactions. With the launch of AlphaFold-multimer in 2022, this problem was mitigated to some degree (92), but AlphaFold still prefers simpler proteins. When it comes to more complex situations, like ligand interactions and membrane proteins with lipid particles integrated in the structures, AlphaFold still has no dedicated way of dealing with this. It is also evident that for proteins with multiple conformational states available, such as a channel protein that opens and closes mechanically, AlphaFold will provide random conformations, seemingly biased towards the related structures available in structure data bases. When it comes to intrinsically disordered regions, AlphaFold is good at predicting their existence, but worse at modelling them connecting to other domains correctly.

Validating the model

Once AlphaFold has been run for a given sequence, you will be provided with a set of models, ranked from best to worst based on an internal scoring system. You will also get a number of files, providing detailed information on the prediction process, including an MSA, and predictions of alignment error, contacts, distogram, and LDDT (Local Distance Difference Test). The predicted LDDT (pLDDT) is an important metric, that is a confidence measure for each residue in the structure, with a higher pLDDT score corresponding to a better prediction. The pLDDT score is stored in the column for B-values in the .pdb file, and one must therefore be careful when analyzing the data. Typically, the model is colored from low pLDDT score in red to high pLDDT score in blue. It is also a good idea to compare a predicted structure manually to experimentally determined structures to evaluate the reliability. In **Paper III**, we predicted a structure of *Drosophila melanogaster* TRPA1 using AlphaFold. After modelling the structure using AlphaFold, but before we had published, an experimentally determined structure determined using Cryo-EM was released (24). In that case, we used the published structure to validate our predicted model, and as a proof of concept for AlphaFold's ability to accurately predict insect TRPA's, a group of proteins it had not been trained on due to no published structures being available at the time of AlphaFold's training.

Circular Dichroism

Circular dichroism, as mentioned above, might seem like an inferior method to the ones giving higher resolution information, but its advantages should not be overlooked. The most immediate use of CD is the prediction of secondary structure of the protein. This can both be used as a first characterization of the protein of interest, but is also useful to verify the quality and content of the sample. As it can be performed at ambient temperature and pressure, and in solution, the conditions are much more similar to native conditions than other structural methods. Sample preparation is also much more straightforward. Even though there are limitation on the buffer composition such as chloride ions and some detergents not being compatible due to their absorbance at low wavelengths, it is a flexible method (49). As CD can be used at various temperatures, thermal stability measurements – also known as melt experiments – can be done to measure unfolding of the protein at increasing temperatures. This gives a melting temperature (T_m) that is defined as the temperature midpoint of a sigmoidal melting curve where the protein is equally distributed between its folded and unfolded state (93), and that can be used as a comparison between different proteins or conditions. Specifically, this can be used to study ligand interactions, as even small changes in stability can be detected using CD by monitoring shifts, either in the T_m value directly, or in the melting curve shape. It is, of course, impossible to confidently correlate these shifts to changes in

protein activity, but using CD to look for binding and conformational changes is still valuable, especially in an integrative structural biology perspective. (94)

Interaction study (AQP)

Aquaporin localization

AQP2

The aquaporin that is the most well studied based on its role in diseases is human aquaporin 2 (AQP2). AQP2 is involved in a number of clinical conditions, and one of them is nephrogenic diabetes insipidus (NDI), which is characterized by the patient not being able to concentrate urine properly (95). AQP2 is located in the collecting duct of the kidney, and is responsible for reabsorbing water from the urine, but when it no longer functions correctly, the amount of water lost in the urine can cause severe dehydration. The mechanism of this dysfunction includes mutations that cause either misfolding of AQP2 in the endoplasmic reticulum (ER), or failure in the translocation from storage vesicles to the plasma membrane. AQP2 is normally transported back and forth between storage vesicles and the apical plasma membrane to regulate the water permeability of the membrane, and understanding this mechanism is important to understand how NDI functions. A C-terminal helix has been identified as a possible site of protein-protein interaction, and an interesting subject for study. (96)

AQP5

Human aquaporin 5 (AQP5), is another aquaporin whose activity is regulated by trafficking to the plasma membrane, and defects therein are connected to Sjögren syndrome, that manifests as dryness in eyes and mouth. The mechanism of AQP5's translocation is however less well studied than that of AQP2. Still, a C-terminal helix, analogous to the one in AQP2, has been identified as a site of interest. (97)

FERM

There is a protein family known as ERM (Ezrin/Radixin/Moesin), that connect the actin cytoskeleton to the plasma membrane by crosslinking, and importantly, they are also involved in membrane protein trafficking. One member of the ERM family, Ezrin, has been shown to interact with the C-terminal helix of AQP2, and knockout

of ezrin causes reduced AQP2 endocytosis, by limiting its trafficking. The part of ezrin that is involved in this interaction is the highly conserved N-terminal FERM (band Four-point one ERM) domain (98). Likewise, ezrin has also been shown to interact with AQP5, and it is suggested that the FERM domain is involved here as well (99). In **Paper IV**, we study the interaction between the FERM domain of ezrin, and the C-terminal helix of AQP2 and AQP5.

Protein interaction

To determine protein-protein interactions is to study an equilibrium. In principle, a protein-protein interaction between protein A and protein B can be described by the following reaction, where the free proteins form the complex AB when bound:

$$A + B \rightleftarrows AB \tag{1}$$

For this association-dissociation reaction we can define an equilibrium constant, which is usually referred to as the dissociation constant (K_D):

$$\mathrm{K_D} = \frac{[\mathrm{A}][\mathrm{B}]}{[\mathrm{AB}]} \tag{2}$$

It can also be described as the concentration of A, when half of the B molecules are bound to A in the form of AB. To exemplify this, a binding curve is shown in figure 17 with K_D indicated. It is, in other words, possible to describe the affinity between two proteins (or other molecules), by titrating one protein and measuring the concentration of either the other free protein, or the complex. (100)

Microscale Thermophoresis (MST)

MST is a method used to measure interactions between molecules in solution, by measuring how much the movement of one of the molecules generated by a heating IR-laser is affected by different concentrations of the other molecule. For this to work, one of the molecules needs to be fluorescent, either from internal fluorescence, or from an attached fluorophore. The initial fluorescence has to be constant, and it is the non-fluorescing molecule that is titrated. At the time when the IR-laser is activated, there is a change in fluorescence intensity known as a T-jump, caused by movement of the fluorescent molecules. This movement is sensitive to interactions with other molecules, and it is therefore possible to detect protein binding in microliter volumes, without need for immobilization. (100)

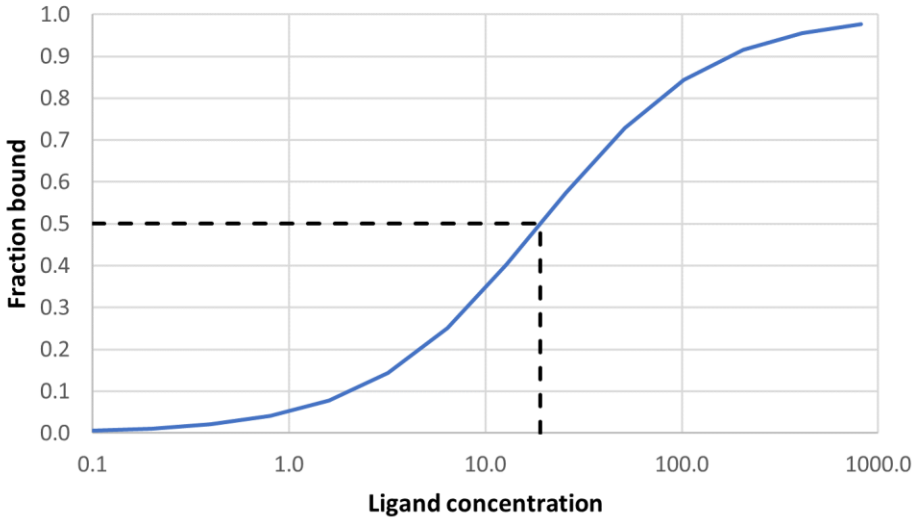


Figure 17: Example binding curve and K_D. The blue curve is an example with arbitrary units of what a binding curve could look like, and the dashed lines indicate that K_D is the point where half of the protein is bound. (logarithmic x-axis)

Results and discussion

Paper I – Structural characterization of HaTRPA1

In this paper, the aim was to study the structure of HaTRPA1. The first step towards this goal was to find a suitable detergent for solubilization. After doing a detergent screen, it was concluded that of the tested detergents, only Fos-choline 12 and 14 were able to solubilize HaTRPA1 in a quantitative manner. However, since neither of these detergents is ideal for Cryo-EM, a test of detergent exchange was done to try and change the detergent to DDM or GDN. It turned out that the exchange went well, and both DDM and GDN were able to hold the protein stable. It should be noted that we faced problems with the purification, and that smaller fragments were seen after SEC, probably originating from protein degradation.

To characterize the secondary structure and stability of the protein, SRCD was performed, with melting studies done from 24.2 °C to 84.5 °C along with ligands benzyl isothiocyanate (BITC), cinnamaldehyde and calcium ions. The midpoint melting temperature for all ligands, as well as the protein without ligand (Apo), was around 50.4 °C. This is comparable to the results presented in a study on TRPA1 from *Anopheles gambiae* of around 55.3 °C (26). It can be noted though, that the protein with BITC had a lower point on the melting curve at 42.8 °C, which probably stems from some type of interaction. It is impossible to say if the effect corresponds to an activation of the ion channel, but we will regard BITC as a putative agonist, and future studies with this compound are recommended.

As a complement to the experimental study, a model of HaTRPA1 was predicted using AlphaFold. The published structures of human and *Drosophila* TRPA1 are definitely recognizable compared to the model, and the site of the cysteine analogous to the Cys621 in human TRPA1 within the allosteric nexus, is preserved. Electrophilic activation of HaTRPA1 is therefore likely possible, and the mechanism might be similar to the one described for human TRPA1 (101). An interesting finding is that we observe a novel C-terminal helix, which interacts with the ARD and the allosteric nexus. This helix is situated in such a way that it would make transition from state 1 to state 2 impossible, as it would block the nexus from moving. However, due to the low pLDDT score of this helix, it should be regarded with some caution. An important consequence of this find is that the C-terminal GFP-tag that was added to HaTRPA1, might affect the formation of these interactions. Cleaving off the GFP-tag is thus more important, but luckily, the TEV

site seems to be readily accessible, as the removal of the GFP-tag worked satisfactorily.

The results in this paper indicate that HaTRPA1 folds much as expected, with a melting profile that is similar to its orthologues. We point out BITC as a putative agonist, but further studies are required to investigate this finding. The quality of the AlphaFold model is overall good, and we value it as tool to complete our structural studies.

Paper II – Evaluation of heterologous expression in Pichia pastoris of Pine weevil TRPA1 by GFP and flow cytometry

In this paper, I dealt with the problem of making the expression and purification yield good enough. To do this, a GFP-tag was added to the C-terminus of HaTRPA1, of both the full-length sequence and the Δ1-708 HaTRPA1 construct lacking the ARD. This allowed for the protein to be monitored, all the way from early screening of highly expressing clones, to optimization of the fermentation process.

The transformation of both constructs into *P. pastoris* gave 200 colony forming units (CFU) for the full-length construct, but a full 600 CFU for Δ1-708 HaTRPA1. By streaking a number of these clones on plates with methanol induction, and then estimating the fluorescence from GFP, a set of promising clones could be selected, along with one weakly expressing representative of each contrast as a negative reference.

The next step was to perform a small-scale induction in 5 mL liquid cultures, where samples were collected at the end of the fermentation to analyze by flow cytometry (FCM) and fluorescence microscopy. For the Δ1-708 HaTRPA1 clones, cell viability was significantly lower, but not for the Δ1-708 HaTRPA1 clone with lower expression levels. The full-length clones did not show the same pattern. The initial conclusion here was that the protein burden of the Δ1-708 HaTRPA1 seemed to be high compared to that for the full-length protein. The highly expressing clones from the plate screen of both constructs had, as expected, higher fluorescence levels in FCM. Also, the full-length clones had slightly higher fluorescence levels when gating for viable cells.

An unexpectedly large difference was seen when cells from the small-scale fermentations were brought under the fluorescence microscope. Where Δ1-708 HaTRPA1 cells featured a green circle around the edge of the cell, looking much like the protein would be inserted into the plasma membrane, the full-length cells had speckles of fluorescence inside the cell, looking more like the protein would be stuck inside some internal structures. The reasons for this could be due to a problem

such as misfolding of the protein, but it is also reasonable to believe that the 2-fold difference in protein size could cause the behavior.

When switching over to a fed-batch bioreactor, the conditions obviously become rather different. It was also possible to collect samples throughout the fermentation, and therefore get time-resolved information about the expression and cell viability. The FCM experiments showed that the fluorescence levels increased rapidly after induction, but then remained stable over time. Compared to the small-scale fermentation, Δ1-708 HaTRPA1 showed higher levels of GFP, which was unexpected, but as mentioned before, the conditions are very different in a fed-batch culture compared to a shake-culture. The decrease in cell viability for Δ1-708 HaTRPA1 is seen here as well, and the change seems to be gradual, with 64% viable cells at the endpoint. The fluorescence microscopy however, showed very much the same difference between Δ1-708 HaTRPA1 and the full-length construct, as seen in the small-scale induction.

Finally, the GFP-tagged protein expressed in the fed-batch reactor was purified using a standard protocol, similar to the one used in Paper I. It was easy to follow the protein throughout the purification steps, but unfortunately, we realized that the protein did not remain intact. Although the protein eluted at expected size on a SEC column, when run on an SDS-PAGE gel, several smaller bands could be seen that we interpret as degradation. This degradation is not limited to the GFP-tag, but seems to affect the actual protein of interest. Still, one should bear in mind that membrane proteins tend to migrate unpredictably on gels, and it is therefore difficult to estimate the full extent of this degradation.

All-in all, we successfully employed several methods to improve and monitor the expression of two HaTRPA1 constructs. Further optimization is however needed in the purification steps, to avoid protein degradation.

Paper III – TRPA5 encodes a thermosensitive ankyrin ion channel receptor in a triatomine insect

The goal of my third paper was to describe the thermosensitivity of TRPA5 from *Rhodnius prolixius* (kissing bug, RpTRPA5). TRPA5 is a TRP-channel that is not present in human, but neither in some insects like *D. melanogaster*.

The thermosensitivity of RpTRPA5 was determined using whole-cell patch clamp with heating delivered by a heat-pulse from a laser. The high temperature coefficient was determined to $Q_{10} = 25$, and currents were induced by temperatures from 53°C to 68°C.

To map its relationship within the TRPA family, a phylogenetic placement was performed with 46 insect families from 9 major orders. In essence, it showed that some of the channels, TRPA1, painless water-witch, pyrexia and TRPA5, are not present in all species. For example, TRPA5 is missing in dipterans – mosquitos and

flies like *Drosophila*, and pyrexia are missing in hemipterans – such as *Rhodnius*. It seems as if different TRPAs fill the same role in different species through convergent evolution.

3D models of RpTRPA5, and a set of homologous proteins, namely the other TRPA members present in *Rhodnius*, and all the TRPA members present in *Drosophila*, were predicted using AlphaFold. Remember that TRPA5 is not present in *Drosophila*, and pyrexia is not present in *Rhodnius*. Due to limited computational power, the full tetramer could not be modelled for any of the homologues, and therefore we needed to find a way around this problem. The first step was that we modelled monomers of each homologue. This was useful to identify major features of the structure, which could be compared to published structures of human TRPA1. AlphaFold was able to predict the transmembrane helices, the TRP-like domain and most of the ARD, but was more at loss when it came to the C-terminal coiled-coil region. This is probably because it relies on interactions with the other monomers to fold the proper super helical secondary structure. Apart from monomers, it was also possible to model a truncated tetramer consisting only of the transmembrane region. By then using that as a template for modelling of a new monomer, we could assemble a tetramer of four of these monomers, by aligning them with the predicted transmembrane region. This resulted in a tetramer model with the important pore region, but also with the full ARD.

Shortly after modelling the tetramers, but before we had time to publish, a *Drosophila* TRPA1 structure with the ARD solved was released (24). We decided that this could actually be a good thing for us, as this provided an opportunity to validate our *Drosophila* TRPA1 structure against a structure that definitely was not part of AlphaFold's training set. We found that the structures agreed well, and not just in the transmembrane region, but also in the ARD. This gave us more confidence in our other TRPA homologue predictions.

An interesting find from analyzing the tetrameric structure, is that when comparing the pore of the modelled RpTRPA5 with a published human TRPA1 structure in closed conformation (21), the selectivity filter is actually narrower in RpTRPA5 due to the positioning of the side chain of Glu914. However, the lower gate is wider, with helix S6 giving more room. Another interesting feature of the pore of RpTRPA1 is that Gly914, which is suggested to be important for gating in human TRPA1 (21), is absent in RpTRPA1. It does, however, seem as if the structure has evolved to accommodate this deletion, and inserting a glycine at the same position will cause significant disruption, especially to the position of helix S6.

We also looked at the ARD, which has been linked to thermosensitivity of TRPA1 (102), and saw a few interesting features. RpTRPA5 has a higher number of ankyrin repeats, but also longer loops between them, compared to the other insect TRPAs. We cannot draw any conclusions regarding the effect these differences would have on the thermosensitivity of RpTRPA5, but they are still noteworthy. It is possible

that future studies could reveal the importance of these variations, and by doing so clarify the role of the ARD for thermosensation in TRPs.

Paper IV – Structural basis for the interaction between the Ezrin FERM-domain and human aquaporins

For the fourth paper, the goal was to investigate interactions between the C-termini of aquaporins 2 and 5, and the FERM-domain of Ezrin. The full-length proteins AQP2 and AQP5 were expressed in *P. pastoris*, solubilized in the detergent n-Nonyl-Beta-D-Glucopyranoside, and purified using IMAC and SEC. To verify that the location of the interaction, a soluble C-terminal peptide of each was expressed in *E. coli* BL21 Star (DE3), and purified using a GST-tag. The FERM domain of ezrin was similarly expressed without the full protein, and then tagged with Alexa488.

The interactions were then measured using MST, and the dissociation constants were estimated to K_D= 5.9 ± 1.6 µM for AQP2, K_D= 17.4 ± 6.8 µM for AQP5, K_D= 7.86 ± 3.22 µM for the AQP2 peptide, and K_D= 2.19 ± 0.68 µM for the AQP5 peptide. This means that the affinity for FERM is slightly higher for full-length AQP2 than for AQP5, but opposite for the peptides.

To further investigate the nature of this interaction, in-silico modelling with AlphaFold was used. The results showed that two parts of the C-terminal of both AQP2 and AQP5 bind to FERM-domain at two distinct sites, in a similar way to other FERM-complexes. The question of whether phosphorylation plays a role in the interaction remains to be answered.

Summary and Outlook

As with any project, this one has to end at some point. However, some of the questions asked remain to be answered, and new ones have arisen over time. Starting with the goal of finding a possible repellent against *Hylobius abietis*, I have come a few steps closer. Firstly, the hint that BITC might have some interactivity with HaTRPA1 is an important piece in describing the effects of pungent chemicals on the channel activity. The structural information from SRCD, and structure prediction from AlphaFold, are useful in comparing TRPA1 orthologs and planning of fusion points, and could be useful in docking studies.

Although there were some challenges, such as optimizing the purification protocol, the process is stable enough to open for future studies. A Cryo-EM structure is not far off, and with it, new doors would open. With purified protein, methods such as MST could be employed for testing interaction between TRPA1 and agonists.

I am also optimistic that the development of AlphaFold, and other machine learning methods, will continue to push the field. Features such as ligand interactions seems to be next (103), and as computer power becomes cheaper, and the code becomes more efficient, reliability will increase.

On the end of aquaporins, AlphaFold multimer has proven to be a useful tool. The exact nature of interactions between AQP2 and 5, and Ezrin, remain to be investigated further, like the role of phosphorylation.

The study of channel proteins is a thrilling subject, and the results and conclusions presented in this thesis can hopefully pave the way for future discoveries. Linking the worlds of different protein families, as well as linking the worlds of in-vitro and in-silico research is what defines my contribution to our common understanding of nature.

References

1. Cobb M. 60 years ago, Francis Crick changed the logic of biology. PLoS Biol. 2017;15(9):e2003243.
2. Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. Science. 1972;175(4023):720-31.
3. Luckey M. Membrane Structural Biology - with biochemical and biophysical foundations. 2nd ed. New York, USA: Cambridge University Press; 2014.
4. Renard K, Byrne B. Insights into the Role of Membrane Lipids in the Structure, Function and Regulation of Integral Membrane Proteins. Int J Mol Sci. 2021;22(16).
5. Choy BC, Cater RJ, Mancia F, Pryor EE, Jr. A 10-year meta-analysis of membrane protein structural biology: Detergents, membrane mimetics, and structure determination techniques. Biochim Biophys Acta Biomembr. 2021;1863(3):183533.
6. Sigworth FJ. The sensor for animal electricity. Proc Natl Acad Sci U S A. 2023;120(1):e2218703120.
7. Cosens DJ, Manning A. Abnormal Electroretinogram from a Drosophila Mutant. Nature. 1969;224(5216):285-7.
8. Clapham DE. TRP channels as cellular sensors. Nature. 2003;426(6966):517-24.
9. Himmel NJ, Cox DN. Transient receptor potential channels: current perspectives on evolution, structure, function and nomenclature. Proc Biol Sci. 2020;287(1933):20201309.
10. Julius D. TRP channels and pain. Annu Rev Cell Dev Biol. 2013;29:355-84.
11. Kumar A, Balbach J. Folding and Stability of Ankyrin Repeats Control Biological Protein Function. Biomolecules. 2021;11(6).
12. Jaquemar DS, Thomas; Trueb, Beat An Ankyrin-like Protein with Transmembrane Domains Is Specifically Lost after Oncogenic Transformation of Human Fibroblasts. Journal of Biological Chemistry. 1999;274:7325-33.

13. Story GM PA, Reeve AJ, Eid SR, Mosbacher J, Hricik TR, Earley TJ, Hergarden AC, Andersson DA, Hwang SW, McIntyre P, Jegla T, Bevan S, Patapoutian A. . ANKTM1, a TRP-like channel expressed in nociceptive neurons, is activated by cold temperatures. Cell. 2003;112(6):819-29.
14. Jordt S-E, Bautista DM, Chuang H-h, McKemy DD, Zygmunt PM, Högestätt ED, et al. Mustard oils and cannabinoids excite sensory nerve fibres through the TRP channel ANKTM1. Nature. 2004;427(6971):260-5.
15. Talavera K, Startek JB, Alvarez-Collazo J, Boonen B, Alpizar YA, Sanchez A, et al. Mammalian Transient Receptor Potential TRPA1 Channels: From Structure to Disease. Physiol Rev. 2020;100(2):725-803.
16. Zhang H, Wang C, Zhang K, Kamau PM, Luo A, Tian L, et al. The role of TRPA1 channels in thermosensation. Cell Insight. 2022.
17. Clapham DE. Pain-sensing TRPA1 channel resolved. Nature. 2015;520:439-40.
18. Paulsen CE, Armache JP, Gao Y, Cheng Y, Julius D. Structure of the TRPA1 ion channel suggests regulatory mechanisms. Nature. 2015;520(7548):511-7.
19. Suo Y, Wang Z, Zubcevic L, Hsu AL, He Q, Borgnia MJ, et al. Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. Neuron. 2020;105(5):882-94 e5.
20. Balestrini A, Joseph V, Dourado M, Reese RM, Shields SD, Rouge L, et al. A TRPA1 inhibitor suppresses neurogenic inflammation and airway contraction for asthma treatment. J Exp Med. 2021;218(4).
21. Zhao J, Lin King JV, Paulsen CE, Cheng Y, Julius D. Irritant-evoked activation and calcium modulation of the TRPA1 receptor. Nature. 2020.
22. Liu C, Reese R, Vu S, Rouge L, Shields SD, Kakiuchi-Kiyota S, et al. A Non-covalent Ligand Reveals Biased Agonism of the TRPA1 Ion Channel. Neuron. 2020.
23. Terrett JA, Chen H, Shore DG, Villemure E, Larouche-Gauthier R, Dery M, et al. Tetrahydrofuran-Based Transient Receptor Potential Ankyrin 1 (TRPA1) Antagonists: Ligand-Based Discovery, Activity in a Rodent Asthma Model, and Mechanism-of-Action via Cryogenic Electron Microscopy. J Med Chem. 2021;64(7):3843-69.
24. Wang X, Li Y, Wei H, Yang Z, Luo R, Gao Y, et al. Molecular architecture and gating mechanisms of the Drosophila TRPA1 channel. Cell Discov. 2023;9(1):36.
25. Moparthi L, Survery S, Kreir M, Simonsen C, Kjellbom P, Hogestatt ED, et al. Human TRPA1 is intrinsically cold- and chemosensitive with and without its N-terminal ankyrin repeat domain. Proc Natl Acad Sci U S A. 2014;111(47):16901-6.

26. Survery S, Moparthi L, Kjellbom P, Hogestatt ED, Zygmunt PM, Johanson U. The N-terminal Ankyrin Repeat Domain Is Not Required for Electrophile and Heat Activation of the Purified Mosquito TRPA1 Receptor. J Biol Chem. 2016;291(52):26899-912.
27. Tracey WD, Wilson RI, Laurent G, Benzer S. painless, a Drosophila Gene Essential for Nociception. Cell. 2003;113(2):261-73.
28. Lee Y, Lee Y, Lee J, Bang S, Hyun S, Kang J, et al. Pyrexia is a new thermal transient receptor potential channel endowing tolerance to high temperatures in Drosophila melanogaster. Nat Genet. 2005;37(3):305-10.
29. Liu L, Li Y, Wang R, Yin C, Dong Q, Hing H, et al. Drosophila hygrosensation requires the TRP channels water witch and nanchung. Nature. 2007;450(7167):294-8.
30. Cattaneo AM, Bengtsson JM, Montagne N, Jacquin-Joly E, Rota-Stabelli O, Salvagnin U, et al. TRPA5, an Ankyrin Subfamily Insect TRP Channel, is Expressed in Antennae of Cydia pomonella (Lepidoptera: Tortricidae) in Multiple Splice Variants. J Insect Sci. 2016;16(1).
31. Basbaum AI, Bautista DM, Scherrer G, Julius D. Cellular and molecular mechanisms of pain. Cell. 2009;139(2):267-84.
32. Melo N, Capek M, Arenas OM, Afify A, Yilmaz A, Potter CJ, et al. The irritant receptor TRPA1 mediates the mosquito repellent effect of catnip. Curr Biol. 2021.
33. Shimomura K, Ino S, Tamura K, Terajima T, Tomizawa M. TRPA1-mediated repellency behavior in the red flour beetle Tribolium castaneum. Sci Rep. 2022;12(1):15270.
34. Day BLKR. Damage, Control and Management of Weevil Pests, Especially Hylobius Abietis. In: François Lieutier KRD, Andrea Battisti, Jean-Claude Grégoire, Hugh F. Evans, editor. Bark and Wood Boring Insects in Living Trees in Europe, A Synthesis: Springer; 2004. p. 414-44.
35. Galko J, Lalik M, Rell S, Nikolov C, Barta M, Pittner J, et al. Comprehensive comparison of treatments for controlling the large pine weevil (Hylobius abietis) in Central Europe. Sci Rep. 2022;12(1):9673.
36. Azeem M, Iqbal Z, Emami SN, Nordlander G, Nordenhem H, Mozūratis R, et al. Chemical composition and antifeedant activity of some aromatic plants against pine weevil (Hylobius abietis). Annals of Applied Biology. 2020;177(1):121-31.
37. Preston GM, Carroll TP, Guggino WB, Agre P. Appearance of Water Channels in Xenopus Oocytes Expressing Red Cell CHIP28 Protein. Science. 1992;256(5055):385-7.
38. Li S, Li C, Wang W. Molecular aspects of aquaporins. Vitam Horm. 2020;113:129-81.

39. Agre P, Kozono D. Aquaporin water channels: molecular mechanisms for human diseases. FEBS Lett. 2003;555(1):72-8.
40. Azad AK, Raihan T, Ahmed J, Hakim A, Emon TH, Chowdhury PA. Human Aquaporins: Functional Diversity and Potential Roles in Infectious and Non-infectious Diseases. Front Genet. 2021;12:654865.
41. Pauling L, Corey RB, Branson HR. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A. 1951;37(4):205-11.
42. Berg JM, Gatto GJ, Jr., Hines JK, Beneken Heller J, Tymoczko JL, Stryer L. Biochemistry. Tenth edition, international edition. ed: Macmillan Learning; 2023. p. 33-67.
43. Puente-Lelievre C, Malik AJ, Douglas J, Ascher D, Baker M, Allison J, et al. Tertiary-interaction characters enable fast, model-based structural phylogenetics beyond
the twilight zone. bioRxiv. 2024;Preprint.
44. Malik AJ, Langer D, Verma CS, Poole AM, Allison JR. Structome: a tool for the rapid assembly of datasets for structural phylogenetics. Bioinform Adv. 2023;3(1):vbad134.
45. Durairaj J, Waterhouse AM, Mets T, Brodiazhenko T, Abdullah M, Studer G, et al. Uncovering new families and folds in the natural protein universe. Nature. 2023;622(7983):646-53.
46. Batool M, Ahmad B, Choi S. A Structure-Based Drug Discovery Paradigm. Int J Mol Sci. 2019;20(11).
47. Eriksson UK, Fischer G, Friemann R, Enkavi G, Tajkhorshid E, Neutze R. Subangstrom resolution X-ray structure details aquaporin-water interactions. Science. 2013;340(6138):1346-9.
48. Brooks-Bartlett JC, Garman EF. The Nobel Science: One Hundred Years of Crystallography. Interdisciplinary Science Reviews. 2015;40(3):244-64.
49. Miles AJ, Wallace BA. Circular dichroism spectroscopy of membrane proteins. Chem Soc Rev. 2016;45(18):4859-72.
50. Chua EYD, Mendez JH, Rapp M, Ilca SL, Tan YZ, Maruthi K, et al. Better, Faster, Cheaper: Recent Advances in Cryo-Electron Microscopy. Annu Rev Biochem. 2022;91:1-32.
51. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.
52. Freigassner M, Pichler H, Glieder A. Tuning microbial hosts for membrane protein production. Microb Cell Fact. 2009;8:69.

53. Carlesso A, Delgado R, Ruiz Isant O, Uwangue O, Valli D, Bill RM, et al. Yeast as a tool for membrane protein production and structure determination. FEMS Yeast Res. 2022;22(1).
54. EasySelect™ Pichia Expression Kit: For Expression of Recombinant Proteins Using pPICZ and pPICZα in Pichia pastoris. Carlsbad, California: Invitrogen; 2010.
55. Nordén K, Agemark M, Danielson J, Alexandersson E, Kjellbom P, Johanson U. Increasing gene dosage greatly enhances recombinant expression of aquaporins in Pichia pastoris. BMC Biotechnol. 2011;11:47.
56. Lee HJ, Lee HS, Youn T, Byrne B, Chae PS. Impact of novel detergents on membrane protein studies. Chem. 2022;8(4):980-1013.
57. Stetsenko A, Guskov A. An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. Crystals. 2017;7(7).
58. Sadaf A, Cho KH, Byrne B, Chae PS. Amphipathic agents for membrane protein study. Methods Enzymol. 2015;557:57-94.
59. Autzen HE, Julius D, Cheng Y. Membrane mimetic systems in CryoEM: keeping membrane proteins in their native environment. Curr Opin Struct Biol. 2019;58:259-68.
60. Shimomura O, Johnson FH, Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J Cell Comp Physiol. 1962;59:223-39.
61. Tsien RY. The Green Fluorescent Protein. Annual Review of Biochemistry. 1998;67(1):509-44.
62. Arutyunova E, Lysyk L, Morrison M, Brooks C, Joanne Lemieux M. Expression and Purification of Human Mitochondrial Intramembrane Protease PARL. In: Schmidt-Krey I, Gumbart JC, editors. Structure and Function of Membrane Proteins. New York, NY: Springer US; 2021. p. 1-20.
63. Drew D, Kim H. Screening for high-yielding Saccharomyces cerevisiae clones: using a green fluorescent protein fusion strategy in the production of membrane proteins. Methods Mol Biol. 2012;866:75-86.
64. Drew D, Newstead S, Sonoda Y, Kim H, von Heijne G, Iwata S. GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. Nat Protoc. 2008;3(5):784-98.
65. Shapiro HM. Microbial Cytometry: What It Was, Is, and May Be. In: Wilkinson MG, editor. Flow Cytometry in Microbiology: Technology and Applications. United Kingdom: Caister Academic Press; 2015. p. 1-16.

66. López-Amorós R, Comas J, Vives-Rego J. Flow cytometric assessment of Escherichia coli and Salmonella typhimurium starvation-survival in seawater using rhodamine 123, propidium iodide, and oxonol. Appl Environ Microbiol. 1995;61(7):2521-6.
67. Rines DR, Thomann D, Dorn JF, Goodwin P, Sorger PK. Live cell imaging of yeast. Cold Spring Harb Protoc. 2011;2011(9).
68. Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proceedings of the National Academy of Sciences. 2009;106(6):1760-5.
69. Kapust RB, Waugh DS. Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr Purif. 2000;19(2):312-8.
70. Zubcevic L, Le S, Yang H, Lee SY. Conformational plasticity in the selectivity filter of the TRPV2 ion channel. Nat Struct Mol Biol. 2018;25(5):405-15.
71. Huang J, Korsunsky A, Yazdani M, Chen J. Targeting TRP channels: recent advances in structure, ligand binding, and molecular mechanisms. Front Mol Neurosci. 2024;16:1334370.
72. Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. Nature. 2013;504(7478):107-12.
73. Wu X, Rapoport TA. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). Proc Natl Acad Sci U S A. 2021;118(41).
74. Nogales E, Scheres SH. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. Mol Cell. 2015;58(4):677-89.
75. Bandell M, Macpherson LJ, Patapoutian A. From chills to chilis: mechanisms for thermosensation and chemesthesis via thermoTRPs. Curr Opin Neurobiol. 2007;17(4):490-7.
76. Venkatachalam K, Montell C. TRP channels. Annu Rev Biochem. 2007;76:387-417.
77. Hofmann L, Wang H, Zheng W, Philipp SE, Hidalgo P, Cavalié A, et al. The S4---S5 linker - gearbox of TRP channel gating. Cell Calcium. 2017;67:156-65.
78. Taberner FJ, Fernandez-Ballester G, Fernandez-Carvajal A, Ferrer-Montiel A. TRP channels interaction with lipids and its implications in disease. Biochim Biophys Acta. 2015;1848(9):1818-27.
79. Young JW. Recent advances in membrane mimetics for membrane protein research. Biochem Soc Trans. 2023;51(3):1405-16.
80. Schuler MA, Denisov IG, Sligar SG. Nanodiscs as a new tool to examine lipid-protein interactions. Methods Mol Biol. 2013;974:415-33.

81. Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, et al. Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs. Liposomes, Part F. Methods in Enzymology2009. p. 211-31.
82. Suo Y, Lee SY. Sample preparation of the human TRPA1 ion channel for cryo-EM studies. Methods Enzymol. 2021;653:75-87.
83. Yeo H, Mehta V, Gulati A, Drew D. Structure and electromechanical coupling of a voltage-gated Na(+)/H(+) exchanger. Nature. 2023;623(7985):193-201.
84. Huynh KW, Cohen MR, Moiseenkova-Bell VY. Application of amphipols for structure-functional analysis of TRP channels. J Membr Biol. 2014;247(9-10):843-51.
85. Postis V, Rawson S, Mitchell JK, Lee SC, Parslow RA, Dafforn TR, et al. The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. Biochim Biophys Acta. 2015;1848(2):496-501.
86. Frauenfeld J, Löving R, Armache J-P, Sonnen AFP, Guettou F, Moberg P, et al. A saposin-lipoprotein nanoparticle system for membrane proteins. Nature Methods. 2016;13(4):345-51.
87. Gallagher JR, Kim AJ, Gulati NM, Harris AK. Negative-Stain Transmission Electron Microscopy of Molecular Complexes for Image Analysis by 2D Class Averaging. Curr Protoc Microbiol. 2019;54(1):e90.
88. Lyumkis D. Challenges and opportunities in cryo-EM single-particle analysis. J Biol Chem. 2019;294(13):5181-97.
89. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290-6.
90. Wang F, Gong H, Liu G, Li M, Yan C, Xia T, et al. DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. J Struct Biol. 2016;195(3):325-36.
91. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022;19(6):679-82.
92. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-multimer. 2022.
93. Greenfield NJ. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat Protoc. 2006;1(6):2527-35.
94. Wallace BA. The role of circular dichroism spectroscopy in the era of integrative structural biology. Curr Opin Struct Biol. 2019;58:191-6.

95. Deen PM, Verdijk MA, Knoers NV, Wieringa B, Monnens LA, van Os CH, et al. Requirement of human renal water channel aquaporin-2 for vasopressin-dependent concentration of urine. Science. 1994;264(5155):92-5.
96. Frick A, Eriksson UK, de Mattia F, Oberg F, Hedfalk K, Neutze R, et al. X-ray structure of human aquaporin 2 and its implications for nephrogenic diabetes insipidus and trafficking. Proc Natl Acad Sci U S A. 2014;111(17):6305-10.
97. Kitchen P, Oberg F, Sjohamn J, Hedfalk K, Bill RM, Conner AC, et al. Plasma Membrane Abundance of Human Aquaporin 5 Is Dynamically Regulated by Multiple Pathways. PLoS One. 2015;10(11):e0143027.
98. Li W, Jin WW, Tsuji K, Chen Y, Nomura N, Su L, et al. Ezrin directly interacts with AQP2 and promotes its endocytosis. J Cell Sci. 2017;130(17):2914-25.
99. Chivasso C, Hagstromer CJ, Rose KL, Lhotellerie F, Leblanc L, Wang Z, et al. Ezrin Is a Novel Protein Partner of Aquaporin-5 in Human Salivary Glands and Shows Altered Expression and Cellular Localization in Sjogren's Syndrome. Int J Mol Sci. 2021;22(17).
100. Magnez R, Bailly C, Thuru X. Microscale Thermophoresis as a Tool to Study Protein Interactions and Their Implication in Human Diseases. Int J Mol Sci. 2022;23(14).
101. Zhang K, Julius D, Cheng Y. Structural snapshots of TRPV1 reveal mechanism of polymodal functionality. Cell. 2021;184(20):5138-50 e12.
102. Cordero-Morales JF, Gracheva EO, Julius D. Cytoplasmic ankyrin repeats of transient receptor potential A1 (TRPA1) dictate sensitivity to thermal and chemical stimuli. Proc Natl Acad Sci U S A. 2011;108(46):E1184-91.
103. Hekkelman ML, de Vries I, Joosten RP, Perrakis A. AlphaFill: enriching AlphaFold models with ligands and cofactors. Nat Methods. 2023;20(2):205-13.