STRUCTURAL AND FUNCTIONAL STUDIES OF TRPML1 AND TRPP2

Nicole Marie Benvin

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017
ABSTRACT

Structural and Functional Studies of TRPML1 and TRPP2

Nicole Marie Benvin

In recent years, the determination of several high-resolution structures of transient receptor potential (TRP) channels has led to significant progress within this field. The primary focus of this dissertation is to elucidate the structural characterization of TRPML1 and TRPP2.

Mutations in TRPML1 cause mucolipidosis type IV (MLIV), a rare neurodegenerative lysosomal storage disorder. We determined the first high-resolution crystal structures of the human TRPML1 I-II linker domain using X-ray crystallography at pH 4.5, pH 6.0, and pH 7.5. These structures revealed a tetramer with a highly electronegative central pore which plays a role in the dual Ca\(^{2+}\)/pH regulation of TRPML1. Notably, these physiologically relevant structures of the I-II linker domain harbor three MLIV-causing mutations. Our findings suggest that these pathogenic mutations destabilize not only the tetrameric structure of the I-II linker, but also the overall architecture of full-length TRPML1. In addition, TRPML1 proteins containing MLIV-causing mutations mislocalized in the cell when imaged by confocal fluorescence microscopy.

Mutations in TRPP2 cause autosomal dominant polycystic kidney disease (ADPKD). Since novel technological advances in single-particle cryo-electron microscopy have now enabled the determination of high-resolution membrane protein structures, we set out to solve the structure of TRPP2 using this technique. Our investigations offer valuable insight into the optimization of TRPP2 protein purification and sample preparation procedures necessary for structural analysis.
## Table of Contents

List of Figures and Tables ........................................................................................................ vi

Acknowledgements ................................................................................................................ ix

Dedication ............................................................................................................................... xii

### Chapter I

**Introduction to TRPML1 and TRPP2**

1.1 Transient receptor potential (TRP) channel family .......................................................... 1

1.2 Introduction to the TRPML1 protein .................................................................................. 3

   1.2.1 TRPML1 and its role in mucolipidosis type IV (MLIV) .................................................. 4

   1.2.2 Subcellular localization and channel regulation of TRPML1 ........................................ 4

   1.2.3 Trafficking of TRPML1 .................................................................................................. 6

   1.2.4 The TRPML subfamily .................................................................................................. 7

   1.2.5 Electrophysiological properties of TRPML channels ...................................................... 7

1.3 Introduction to the TRPP2 protein and ADPKD ................................................................. 9

   1.3.1 Clinical manifestations of ADPKD ................................................................................. 10

   1.3.2 Pathogenesis of ADPKD ............................................................................................. 11

   1.3.3 Cellular and subcellular localization of TRPP2 ............................................................ 12

   1.3.4 Mechanotransduction model of the PKD1/TRPP2 complex .......................................... 13
1.3.5 TRPP2 and its role in ADPKD................................................................. 14
1.3.6 Stoichiometric assembly of the PKD1/TRPP2 complex................................. 14
1.3.7 TRPP2 agonist triptolide........................................................................... 15
1.3.8 Interaction between TRPP2 and additional channels.................................... 16
1.3.9 Ion selectivity in TRPP2........................................................................... 16
1.3.10 Additional members of the TRPP subfamily............................................... 17
1.4 Structural determination of TRP channels...................................................... 18
1.5 Conclusion and thesis overview ...................................................................... 20
1.6 References...................................................................................................... 22

Chapter II

Structural and Functional Studies of the TRPML1 I-II Linker

2.1 Introduction.................................................................................................... 30
2.2 Experimental procedures for structural studies ............................................. 31
  2.2.1 Protein expression and purification of the TRPML1 I-II linker .................... 31
  2.2.2 Protein crystallization of the TRPML1 I-II linker .................................... 33
  2.2.3 Data collection and structure determination ............................................ 34
2.3 Structural features of the TRPML1 I-II linker structures.............................. 38
  2.3.1 TRPML1 I-II linker structure at pH 6.0 .................................................... 38
  2.3.2 Features of the central luminal pore ......................................................... 39
  2.3.3 TRPML1 I-II linker structures at pH 4.5 and pH 7.5 ............................... 41
2.3.4 Verification of the tetrameric I-II linker ................................................................. 43
2.3.5 Role of the I-II linker in TRPML1 assembly .......................................................... 45
2.4 Structural basis of pH and Ca\textsuperscript{2+} regulation of TRPML1 ................................................................. 47
  2.4.1 Dual regulation of TRPML1 activity by pH and Ca\textsuperscript{2+} ................................................................. 47
  2.4.2 The central pore is essential for pH and Ca\textsuperscript{2+} regulation ................................................................. 48
  2.4.3 Model of pH and Ca\textsuperscript{2+} dual regulation of TRPML1 ................................................................. 50
2.5 Characterization of MLIV-causing mutations ............................................................ 53
  2.5.1 MLIV-causing mutations in the TRPML1 I-II linker ................................................................. 53
  2.5.2 Cell culture, transfection, and confocal live-cell imaging ................................................................. 54
  2.5.3 Confocal imaging of MLIV-causing mutations ................................................................. 55
  2.5.4 Biochemical analysis and circular dichroism spectroscopy ................................................................. 58
2.6 Conclusion .................................................................................................................. 60
2.7 References .................................................................................................................. 62

Chapter III

Structural Studies of TRPP2

3.1 Introduction .................................................................................................................. 64
3.2 Screening of TRPP2 constructs using FSEC .............................................................. 65
  3.2.1 Rationale for cloning truncated constructs ................................................................... 65
  3.2.2 Transient HEK cell transfection and FSEC .............................................................. 66
  3.2.3 FSEC analysis of human TRPP2 constructs transfected in HEK293 cells ................. 67
3.3 Insect cell baculovirus-based expression system ........................................................ 69
3.3.1 Generation of recombinant baculoviruses .......................................................... 69
3.3.2 FSEC analysis of human TRPP2 constructs transfected into Sf9 cells .................... 70
3.4 FSEC of mouse TRPP2 constructs transfected in HEK293 cells ............................. 71
3.5 Purification of hTRPP2 in Hi5 cells using DDM and amphipol ................................. 73
   3.5.1 Small-scale purification in DDM ..................................................................... 74
   3.5.2 Reconstitution into amphipols ....................................................................... 75
   3.5.3 Inhibition of N-glycosylation ...................................................................... 79
   3.5.4 Large-scale purification of TRPP2 in DDM ..................................................... 82
3.6 Purification of hTRPP2 in Hi5 cells using LMNG ...................................................... 85
   3.6.1 Reconstitution of TRPP2 into amphipols ...................................................... 86
3.7 Titan Krios beta-testing phase at NYSBC .............................................................. 89
   3.7.1 Analysis of TRPP2 particles ....................................................................... 91
3.8 Structural studies of the TRPP2 F604P mutant ..................................................... 93
3.9 Titan Krios data collected at Tsinghua University .................................................. 98
3.10 Comparison of methods used for TRPP2 structural studies ................................. 101
   3.10.1 Constructs and protein purification methods .............................................. 101
   3.10.2 Parameters for cryo-EM analysis ............................................................. 110
3.11 Analysis of new constructs reconstituted into nanodiscs and amphipols .......... 113
3.12 Summary ........................................................................................................... 118
3.13 Conclusion ......................................................................................................... 120
3.14 References ....................................................................................................... 121
Chapter IV
Analysis of TRPP2 Structures

4.1 Introduction .................................................................................................................. 124
4.2 General architecture of the TRPP2 structures ......................................................... 124
4.3 Extracellular polycystin-mucolipin domain in TRPP2 ............................................. 126
   4.3.1 Structural features of the polycystin-mucolipin domain ...................................... 126
   4.3.2 Comparison of polycystin-mucolipin domains ..................................................... 129
4.4 Structural features of TRPP2\textsubscript{CL}, TRPP2\textsubscript{MI}, and TRPP2\textsubscript{SI} .......................................................... 132
4.5 Characteristics of the ion-translocation pathway in TRPP2\textsubscript{CL} ......................... 135
4.6 Ion-translocation pathways in TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI} ............................. 137
4.7 Role of glycosylation in TRPP2 ................................................................................ 140
4.8 Discussion of PKD-associated missense mutations in TRPP2 ................................ 143
4.9 Conclusion ................................................................................................................. 149
4.10 References ................................................................................................................ 150

Epilogue .............................................................................................................................. 151

Appendix A Structural basis of dual Ca\textsuperscript{2+}/pH regulation of the endolysosomal TRPML1 channel ................................................................................................................................. 154

Appendix B Vita and Publications ..................................................................................... 176
List of Figures and Tables

Chapter I

Figure 1. TRP channel subfamilies
Figure 2. Schematic representation of a TRP channel
Figure 3. Subcellular localization of TRPML1
Figure 4. Schematic representation of the PKD1/TRPP2 complex

Chapter II

Figure 1. Expression and purification of the TRPML1 I-II linker
Figure 2. Crystals of the TRPML1 I-II linker at pH 4.5 and pH 7.5
Figure 3. Crystal structure of the TRPML1 I-II linker at pH 6.0
Figure 4. Features of the central luminal pore in the TRPML1 I-II linker
Figure 5. Comparison of the TRPML1 I–II linker structures
Figure 6. Tetrameric structure of the I-II linker verified in full-length TRPML1
Figure 7. Role of intersubunit interactions in TRPML1 assembly
Figure 8. Regulation of TRPML1 by Ca$^{2+}$ and pH
Figure 9. Role of luminal pore aspartate mutations in Ca$^{2+}$ regulation
Figure 10. Schematic representation of Ca$^{2+}$/pH dual regulation of TRPML1
Figure 11. Locations of MLIV-causing missense mutations in the TRPML1 I-II linker
Figure 12. Subcellular localization of wild-type and MLIV-causing mutant TRPML1
Figure 13. Biochemical and structural analysis of the I-II linker pathogenic mutations
Table 1. Data collection and refinement statistics
Chapter III

Figure 1. FSEC traces of human TRPP2 constructs after transient transfection in HEK cells
Figure 2. FSEC traces of human TRPP2 constructs after transient transfection in Sf9 cells
Figure 3. FSEC traces of mouse TRPP2 constructs after transient transfection in HEK cells
Figure 4. SEC traces from small-scale purification of TRPP2 in DDM and amphipol A8-35
Figure 5. SEC traces from small-scale purification of TRPP2:181-927
Figure 6. Role of glycosylation in membrane protein structural determination
Figure 7. TRPP2-MBP and TRPV2 purified in DDM
Figure 8. SEC traces of TRPP2 purified in LMNG and amphipol A8-35
Figure 9. Representation of TRPP2 data collected on the Titan Krios
Figure 10. Cryo-EM particles of TRPP2 collected on the Titan Krios
Figure 11. Gel filtration profiles of TRPP2 F604P purified with LMNG and soybean lipids
Figure 12. Negative-stain and EM grids of the hTRPP2:199-927 F604P-MBP in LMNG
Figure 13. Negative-stain and cryo-EM grids of hTRPP2 F604P in amphipol A8-35
Figure 14. Cryo-EM analysis of hTRPP2:199-927 in LMNG
Figure 15. Schematic diagram displaying constructs utilized in TRPP2 structural studies
Figure 16. Comparison of gel filtration profiles of hTRPP2
Figure 17. Representative micrographs of human TRPP2
Figure 18. 2D class averages of human TRPP2 cryo-EM particles
Figure 19. Analysis of TRPP2 reconstituted in nanodiscs
Figure 20. Assessment of the homogeneity of TRPP2 F604P

Table 1. Protein purification procedures for structural studies of TRPP2
Table 2. Sample preparation and cryo-EM data acquisition parameters for TRPP2 studies
Chapter IV

Figure 1. Cryo-EM structure of human TRPP2 in lipid nanodiscs
Figure 2. Polycystin–mucolipin domain structure of TRPP2
Figure 3. Interactions in the polycystin–mucolipin domain affecting tetrameric assembly
Figure 4. Sequence alignment of the TRPML1 and TRPP2 polycystin-mucolipin domains
Figure 5. Structural comparisons of the polycystin–mucolipin domains
Figure 6. Comparison of the closed and open conformational states of TRPP2
Figure 7. Cryo-EM structures of human TRPP2 in open conformational states
Figure 8. Configuration of the TRPP2_CL pore
Figure 9. Representation of the ion permeation pathways of TRPP2_M1 and TRPP2_S1
Figure 10. Cation-binding sites along the ion permeation pore of TRPP2_M1 and TRPP2_S1
Figure 11. Presence of N-glycosylation within TRPP2_M1 and TRPP2_S1
Figure 12. Locations of pathogenic ADPKD missense mutations in TRPP2

Table 1. Glycosylated residues present in the published TRPP2 structures
Table 2. ADPKD-causing missense mutations in TRPP2
Acknowledgements

It has been a truly rewarding experience to be part of the Biological Sciences PhD program at Columbia University. I feel honored and blessed to have been given this opportunity to conduct research at this renowned university. It has been quite a journey, and I would like to acknowledge the many individuals who have helped me along the way.

During my undergraduate research experience at Cornell University, I gained an interest in the field of structural biology in Dr. Holger Sondermann’s lab. Therefore, upon entering the PhD program at Columbia, I knew I wanted to delve deeper into this remarkably visual field. I would like to express my appreciation to my thesis advisor, Dr. Jian Yang, for giving me the wonderful opportunity to work on structural biology projects involving proteins that are directly implicated in genetic diseases. Dr. Yang has provided me with valuable support, mentorship, and resources throughout my PhD and encouraged me to explore unchartered territory in the TRP channel field. The impactful nature of this research motivated me to make significant contributions to the scientific community. Attempting to solve a full-length membrane protein structure is inherently difficult and my project could not have progressed as far as it did without the advice and suggestions from many wonderful individuals.

I am sincerely grateful for having two exceptional committee members, Dr. Wayne Hendrickson and Dr. Filippo Mancia, who have a great deal of expertise in the structural biology field. For years, they have provided me with extensive support, guidance, and direction for all of my projects, which have proven to be invaluable. I am truly thankful for all of the time they have taken to offer insight and suggestions on how to move my projects forward, especially when dilemmas arose.
I would also like to extend my thanks to Dr. Reza Khayat and Dr. Chloe Bulinski for being on my thesis committee. I initiated a collaboration with Dr. Reza Khayat at the City College of New York, who has graciously dedicated time collecting and analyzing data on my protein samples. Dr. Khayat gave valuable advice for optimizing sample conditions, and his expertise in the cryo-EM field enabled the project to progress. I would also like to express my gratitude towards Dr. Chloe Bulinski for her continuous support and encouragement throughout my PhD experience. I really appreciate all of her guidance.

I would like to thank Dr. Devendra Srivastava, a postdoctoral fellow in Dr. Khayat’s lab for discussing ways to optimize protein conditions, and for testing my samples, often times late into the night. I would also like to thank Dr. Xueming Li at Tsinghua University for collaborating with us, as well as Heng Zhou and Xiaoyuan Zhou for testing protein samples. They were always willing to update me with results at any hour despite the significant time difference. I would also like to acknowledge Dr. Ravi Kalathur at NYSBC, who helped me at the initial stages of the TRPP2 project. Overall, I was fortunate enough to work collaboratively with many intelligent scientists over the years and have learned a great deal from these experiences.

I would also like to express my appreciation to current and former members of the Yang lab. In particular, I would like to extend my thanks to Dr. Minghui Li who I have been fortunate enough to work with on the TRPML1 project over the years. Minghui has provided me with great insight and advice and first introduced me to the realm of membrane protein purification. I am grateful for having the opportunity to work with Minghui and Dr. Wei Kevin Zhang, a former member of the lab. Together, the research that Minghui, Kevin, and I conducted culminated in a triple co-first author publication. In addition, I appreciate the advice and support that my colleagues Qiao Feng and Dr. Yannis Michailidis have offered me. I would also like to thank the
students that I mentored, Margarita, Gabriel, Alex, and Rukmini for aiding with experiments throughout the years.

I am grateful to the graduate students and postdocs in the department for informative discussions as well as the administrative staff for overseeing that the needs of the lab are addressed. I would like to thank all of my friends for being there for me throughout my PhD journey and for understanding my ambitious nature. In particular, I would also like to thank Philip, who understands the inherent unpredictability of structural biology projects. I appreciate his support, especially during those long days in lab.

Lastly, I am extremely grateful to my mother, father, and sister for their continuous love and guidance throughout my entire life. Their encouragement and optimism have always motivated and inspired me to believe that anything is possible if I set my mind and heart to it. I thank them for instilling in me the value of diligence, since hard work, determination, and dedication are the keys to success. They have always supported my goals and dreams, no matter how challenging they may have seemed.
For my wonderful family, for always encouraging me to pursue my dreams, be resilient, and persevere. Their unwavering love and support during the course of my life’s journey have allowed me to accomplish goals that once seemed impossible.
Chapter I

Introduction to TRPML1 and TRPP2

1.1 Transient receptor potential (TRP) channel family

Transient receptor potential (TRP) channels are a multifunctional superfamily of cation channels which consists of seven subfamilies: TRPC, TRPM, TRPA, TRPN, TRPV, TRPML, and TRPP (Figure 1) (1, 2). These subfamilies are categorized on the basis of their amino acid sequence homology and function (1, 3, 4). Interestingly, TRPN channels are not present within mammalian genomes and 27 members of the TRP superfamily are found in humans (5).

Figure 1. TRP channel subfamilies. Schematic diagram of TRP channel subfamilies illustrating the transmembrane topology. Domains or motifs at the N- and C-termini of the channels are depicted. Figure adapted from (6).
TRP channels are readily influenced by their environments, and thus play an essential role in cellular sensory perception (1). Notably, they are able to be activated by varying signals such as temperature, mechanical stimuli, and ligands (2, 5). TRP channels differ largely in their physiological functions and mechanism of actions, and are regulated by diverse intracellular signals, including calcium, lipid metabolites, and phosphoinositides (4).

TRP channels are cation channels that each contain six-transmembrane polypeptide subunits (S1-S6) and assemble as tetramers (Figure 2) (1, 5). Additionally, a pore-forming loop is present between S5 and S6. The N- and C-termini of TRP channels reside intracellularly and differ substantially between subfamilies with regard to their amino acid sequences, lengths, respective motifs, and domains (Figure 1) (1, 5). Moreover, the seven TRP subfamilies have been further divided into two subgroups based on their differences in topology and amino acid sequences (5). Group 1 TRP channels include TRPC, TRPV, TRPA, TRPM, and TRPN, while group 2 encompasses the TRPML and TRPP subfamilies (5).

The aim of this thesis is to illustrate our structural studies pertaining to TRPML1 and TRPP2, founding members of the mucolipin and polycystin TRP subfamilies (1, 5), respectively. As part of the group 2 subfamily, both TRPML1 and TRPP2 exhibit a prominent linker between their first two transmembrane segments (S1-S2) (5). The structural significance of this domain in these two proteins will be further described in Chapters II and IV. In addition, the functional characterization of the TRPML1 linker domain will be expounded upon in Chapter II.
**Figure 2. Schematic representation of a TRP channel.** (A) Illustration of a transient receptor potential protein present in the plasma membrane. The six transmembrane domains are indicated (S1-S6) and the pore loop (P) is situated between S5 and S6. (B) Tetrameric assembly of TRP channels consisting of four monomers (large circles). Of note, the pore loops of each monomer collectively form the central functional pore. Figure adapted from (7).

1.2 Introduction to the TRPML1 protein

The human TRPML1 protein is 580 amino acids in length and has a molecular mass of 65 kDa (8). Like other TRP channels, TRPML1 is comprised of six transmembrane segments (S1-S6) and its N- and C-termini face the cytoplasm (4). As aforementioned, the I-II linker between the first two transmembrane segments forms a pronounced domain. Published results from our laboratory indicate that the TRPML1 I-II linker forms a tetrameric structure (9), which will be described in detail in Chapter II.
1.2.1 TRPML1 and its role in mucolipidosis type IV (MLIV)

Mutations in the gene MCOLN1 lead to mucolipidosis type IV (MLIV) (OMIM #252650), a rare autosomal recessive lysosomal storage disorder (10, 11). MCOLN1 encodes TRPML1 (or mucolipin-1), the founding member of the TRPML channel subfamily (10). Notably, greater than 20 MLIV mutations have been identified in the TRPML1 gene (12, 13). The disease incidence of MLIV is approximately 1 in 40,000 individuals and 70-80% of patients diagnosed with MLIV are of Ashkenazi Jewish ancestry (8, 11, 13). Most commonly, MLIV becomes apparent during infancy, clinically characterized by symptoms including visual complications (corneal clouding, optic atrophy, and retinal degeneration), severe psychomotor impairments, and achlorydria (11, 14, 15). After diagnosis, the patients remain stable and little disease progression occurs for two to three decades (13, 16).

Proper diagnosis of MLIV remains challenging due to its similarity in symptoms to other neurodegenerative diseases (13). As a result, patients with MLIV have been misdiagnosed as having spastic paraplegia and cerebral palsy, among other disorders (17). Additionally, patients with milder MLIV mutations can go undiagnosed, which raises the question of the overall rarity of this disease (13, 18).

1.2.2. Subcellular localization and channel regulation of TRPML1

Regulation of TRPML1 activity is dependent upon luminal and extracellular Ca$^{2+}$ and H$^+$ (12, 19–21). TRPML1 primarily localizes to lysosomes, yet upon phagosomal biogenesis and lysosomal exocytosis it can be found in late endosomes and the plasma membrane, respectively (Figure 3) (21–24). Therefore, TRPML1 is exposed to a wide range of Ca$^{2+}$ and pH levels, depending upon where it localizes in the cell. Notably, since TRPML1 is predominantly present
within lysosomes, it is exposed to an acidic environment of pH 4.5-5 with a Ca\textsuperscript{2+} concentration of 0.5-0.6 mM \cite{21, 25, 26}. In late endosomes, it is in contact with surroundings of pH 5.0–6.0 and 0.5 mM Ca\textsuperscript{2+}, while at the plasma membrane TRPML1 exists in an environment containing a pH of 7.2–7.4 with Ca\textsuperscript{2+} levels of 1.8–2 mM \cite{21, 25, 26}. Hence, dual regulation of TRPML1 by Ca\textsuperscript{2+} and pH has been shown to vary. Of importance, the I-II linker in TRPML1 faces the lumen in lysosomes and the extracellular space when present on the plasma membrane. The regulatory mechanism involving the dual regulation of Ca\textsuperscript{2+} and pH of TRPML1 will be further elucidated in Chapter II.

**Figure 3. Subcellular localization of TRPML1.** TRPML1 mainly localizes to late endosomes and lysosomes, yet can also be present at the plasma membrane upon lysosomal exocytosis. The lysosomal membrane trafficking pathway is illustrated.
Furthermore, TRPML1 is also regulated by phosphoinositides in subcellular compartments (27, 28). TRPML1 has been shown to be activated by (PI(3,5)P$_2$) in the membranes of endosomes and lysosomes (27). Conversely, TRPML1 is inhibited by PI(4,5)P$_2$ at the plasma membrane (28). Taken together, regulation by pH, Ca$^{2+}$, and phosphoinositides ensures that TRPML1 activity is high within lysosomes and low at the plasma membrane.

1.2.3 Trafficking of TRPML1

The loss of TRPML1 due to MLIV mutations leads to the buildup of undigested lipid components and additional biomaterial in enlarged late endosomal/lysosomal compartments, as seen in other lysosomal storage disorders (23, 29). However, the agglomeration of lipid products are more heterogeneous than in other storage disorders. These lipids include cholesterol, acidic mucopolysaccharides, sphingolipids, phospholipids, and gangliosides (30–32).

Although most lysosomal storage disorders exhibit abnormalities in lysosomal hydrolases associated with the processing of lipids (33), the hydrolase activity within cells of MLIV patients is reported to be fairly normal (34). Rather, abnormalities that are manifested may presumably be due to a defect in the sorting and trafficking within the late endocytic pathway (34). In particular, the departure of lipids from the late endosome and lysosome to the trans-Golgi network is abnormal in ML4 cells (34, 35). Hence, TRPML1 is presumed to play an essential role in the formation of transport vesicles that traffic materials between these subcellular compartments. Additionally, it has been reported that the expression of TRPML1 increases after the induction of lysosome biogenesis (36). Thus, TRPML1 may be associated with reconfiguring lysosomes present in the late endosome–lysosome hybrid organelles (37, 38).
1.2.4 The TRPML subfamily

Structural similarities

The TRPML subfamily consists of three proteins, namely TRPML1-3, that exist in mammals (26) and exhibit 60% amino acid sequence homology to one another (16). Our research group determined the I-II linker structure of TRPML1 (9), which will be described further in Chapter II. Due to the high sequence homology among TRPML channels, TRPML2 and TRPML3 likely display similar I-II linker structures to that of TRPML1. Notably, TRPML1 is able to homo- or heteromultimerize with TRPML2 and TRPML3 (39).

Expression pattern of TRPML1-3 in tissue

TRPML1 expression exists in every tissue in mammals, yet highest expression levels occur in the kidney, brain, liver, spleen, and heart (10). Contrarily, the expression of TRPML2 and TRPML3 is limited to several organs. TRPML2 is primarily expressed in the mouse liver, heart, thymus, spleen, and kidney (40). In addition, TRPML3 is present in the kidney, lung, thymus, eye, spleen (40, 41), epithelial cells in the cochlea (42), and hair cells (42, 43). Interestingly, TRPML1-3 are all present within the kidney.

1.2.5 Electrophysiological properties of TRPML channels

Overexpressed TRPML3 has been found to exist at the plasma membrane (44). Through the whole-cell patch clamp method, TRPML3 was identified as an inwardly rectifying Ca\textsuperscript{2+} permeable channel (41, 43, 45). However, since overexpressed TRPML1 and TRPML2 primarily localize to late endosomes and lysosomes (26), it has be challenging to characterize their electrophysiological properties. A gain-of-function mutation (A419P) in mouse TRPML3 renders the channel
constitutively active and causes the varitint-waddler (Va) phenotype, characterized by pigmentation and deafness in mice (43, 45–47).

Notably, the introduction of the analogous gain-of-function (GOF) V432P mutation in TRPML1 and A396P mutation in TRPML2 results in constitutively active channels at the plasma and endolysosomal membranes (43, 48, 49). In TRPML1, the V432P mutation is situated in the lower region of S5 by the channel gate, which presumably secures the channel in an open state (27, 43). Thus, the TRPML1 V432P (TRPML1VP) mutation heightened the ability of the channel to localize to the plasma membrane, compared to that of wild-type TRPML1. Hence, this GOF mutation has essentially enabled functional characterization of the TRPML1 and TRPML2 channels through electrophysiological approaches (40, 43, 48, 49). Our findings pertaining to TRPML1VP investigations will be further discussed in Chapter II.
1.3 Introduction to the TRPP2 protein and ADPKD

TRPP2 (also referred to as polycystin-2, PC2, or PKD2) is a 968 amino acid 110 kDa membrane protein encoded by the PKD2 gene, and is part of the TRP channel superfamily (5, 50). Like other TRP channels, it contains six transmembrane segments, intracellular N- and C-termini, and a pore-forming loop between S5 and S6 (Figure 4) (1, 5, 50). TRPP2 is considered to be a non-selective calcium channel that regulates intracellular calcium levels (51–53). Of note, the C-terminal EF-hand domain binds to calcium and may aid in calcium regulation (54). Additionally, a domain distal to the EF-hand is the coiled-coil domain, which plays a role in the assembly of complexes (55, 56). The structures of the EF-hand domain and the coiled-coil domain have been solved by NMR (54, 57, 58) and X-ray crystallography (59), respectively. As aforementioned, a I-II linker exists between the first two transmembrane segments of TRPP2, which is likewise exhibited in TRPML1. The structural features of TRPP2 will be elucidated in detail in Chapter IV.

Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common monogenetic disorders in humans and is characterized by the bilateral development of fluid-filled cysts in the kidneys (60–62). Notably, mutations in TRPP2 cause approximately 15% of all ADPKD cases (50). Moreover, mutations in PKD1, a 4,302 amino acid integral membrane protein with eleven transmembrane segments (63, 64), cause 85% of the other ADPKD cases (Figure 4) (60–62). PKD1 also contains a TRP channel sequence homology domain in transmembrane segments 6-11 (Figure 4) (65). The N-terminal region of PKD1 spans ~3000 amino acids and includes several motifs and domains, while the C-terminus comprises ~ 220 amino acids (Figure 4) (64). As indicated in Figure 4, it has been reported that TRPP2 and PKD1 form a complex through the association of their coiled-coil domains (61, 66). The PKD1/TRPP2 complex has been shown to be involved in various cellular processes such as mechanosensation, cell proliferation,
cell polarity, and apoptosis (61, 67–71). The role that this complex plays in mechanosensation will be explored in greater detail in section 1.3.4.

**Figure 4. Schematic representation of the PKD1/TRPP2 complex.** Structural motifs and domains in PKD1 and TRPP2 are indicated. Of note, PKD1 and TRPP2 interact through their coiled-coil domains. The ion permeation pathway of TRPP2 is shown between the 5th and 6th transmembrane segments. Figure adapted from (65).

### 1.3.1 Clinical manifestations of ADPKD

ADPKD generally affects between 1:400 and 1:1000 individuals worldwide (72). In many cases, ADPKD can lead to the development of end-stage renal disease (ESRD) (73). The first noticeable signs of ADPKD include hypertension (74) and abdominal masses (72). Pain is reportedly the most common symptom among ADPKD patients, particularly in the abdominal regions and the lower back (75, 76). ADPKD is a late onset disorder and patients will generally experience symptoms within the third and fourth decade of life (72). Patients who have mutations
in TRPP2 exhibit a milder form of ADPKD by experiencing later onset of the disease and reach end stage renal failure more than ten years later than patients who have mutations in PKD1 (77). Hence, individuals that have mutations in PKD1 reportedly have a shorter life expectancy than those with mutations in TRPP2 (77). Moreover, effective treatments for ADPKD are yet to be determined.

1.3.2 Pathogenesis of ADPKD

The “two hit” proposed mechanism in ADPKD is believed to initiate the development of cyst formation in patients (78). According to this hypothesis, both copies of a PKD gene (PKD1 or TRPP2) need to be inactivated in order for cyst formation to occur. Thus, the inherited germ line mutation signifies the “first hit” and an additional somatic mutation that arises denotes the “second hit” (78). Therefore, the ensuing homozygous inactivation of a PKD gene is purportedly essential for the pathogenicity of ADPKD, culminating in the expansion of cysts (79). Despite evidence supporting this “two-hit” hypothesis, it remains controversial. In theory, if indeed a somatic mutation triggers the development of cysts, then all cyst cells would harbor this mutation, yet it has been shown that this phenomena occurs in only a subset of isolated cysts (80). These findings indicate that factors other than somatic mutations have been implicated in cyst development. Alternatively, it has been proposed that haploinsufficiency necessitates the formation of cysts, in which mutations in one allele may lead to cyst generation through a genedosage-dependent mechanism (81). Purportedly, if the expression level of PKD1 or TRPP2 is below a critical threshold in tissues, resulting from genetic, environmental, and stochastic factors, it could incline renal epithelial cells to initiate the development of cysts (81).
1.3.3 Cellular and subcellular localization of TRPP2

TRPP2 is primarily expressed in the kidney, yet expression levels have been found in other tissues including the heart, intestine, pancreas, bile ducts, ovary, testis, vascular smooth muscle, and placenta (50, 82, 83). Within the kidney, TRPP2 has been shown to exist in all regions of the nephron, except the glomeruli (72, 73). The intracellular localization of TRPP2 is a complex system that is highly regulated and is dependent on cell type (84). TRPP2 primarily localizes to various subcellular compartments including the endoplasmic reticulum (ER), the plasma membrane, and the primary cilium of renal epithelial cells (51, 85–87). In cells undergoing cell division, TRPP2 localizes to centrosomes and mitotic spindle poles where it is believed to regulate intracellular Ca\(^{2+}\) during or after mitosis (88, 89). Notably, several sequence motifs and binding domains within TRPP2 aid in the regulation of its trafficking between compartments (90, 91).

ER localization

Studies have indicated that native and heterologously overexpressed TRPP2 localizes to the ER membrane of tissues (51) and cultured cells (55, 86, 92). TRPP2 aids in the regulation of Ca\(^{2+}\) release and Ca\(^{2+}\) homeostasis in the ER (51, 71, 93). For instance, TRPP2 functions as a Ca\(^{2+}\) release channel in conjunction with IP3 receptors (93–95). Additionally, it has been proposed that TRPP2 is involved in apoptosis and Ca\(^{2+}\) storage (71). TRPP2 contains an endoplasmic reticulum (ER) retention motif (Glu787–Ser820) whereby phosphorylation at Ser812 by casein kinase 2 (CK2) occurs (90). Upon CK2 phosphorylation, phosphofurin acidic cluster sorting protein 1 and 2 (PACS-1 and PACS-2) subsequently associate with TRPP2, which is fundamental for the retrograde trafficking between the plasma membrane, ER, and golgi (96).
**Plasma membrane localization**

Although it is generally accepted that TRPP2 localizes to the ER and primary cilia, the localization of TRPP2 to the plasma membrane has been an area of controversy (97). It has been shown that native TRPP2 exists on the plasma membrane, but not when TRPP2 is heterologously overexpressed (52, 98). Studies reveal that the formation of the PKD1/TRPP2 complex is necessary for the targeting and retention of TRPP2 to the plasma membrane (55). Thereupon, this complex gives rise to calcium permeable non-selective cation currents (55).

**Primary cilia localization**

Studies have shown that TRPP2 localizes to the primary cilia of kidney tissues as well as cultured epithelial cells (87, 99). A fifteen amino acid conserved RVxP motif at the N-terminus of TRPP2 is necessary for proper trafficking to the primary cilia (91). Interestingly, although TRPP2 is dependent on PKD1 to traffic to the plasma membrane, TRPP2 has been shown to independently traffic to the cilia without the presence of PKD1 (91).

**1.3.4 Mechanotransduction model of the PKD1/TRPP2 complex**

PKD1 and TRPP2 have been shown to form a protein complex in primary cilia of renal epithelial cells, where they act as mechanosensors in the renal tubule (100). It has been reported that luminal fluid flow prompts the bending of cilium, whereby the PKD1/TRPP2 complex detects this movement and generates an influx of \( \text{Ca}^{2+} \) (100). Thus, this response activates a signaling cascade, leading to increased intracellular \( \text{Ca}^{2+} \) levels (51) by \( \text{Ca}^{2+} \)-induced \( \text{Ca}^{2+} \) release. This latter phenomena may occur from activation of TRPP2 in the ER (51) as well as by IP3 and ryanodine receptors (100). Additionally, TRPP2 may play a role in mechanotransduction due to its interactions with several cytoskeletal components including \( \alpha \)-actinin, troponin I, and
tropomyosin-1 \((101–103)\). Taken together, the PKD1/TRPP2 complex appears to regulate renal tubular morphology and function \((55, 104)\).

1.3.5 TRPP2 and its role in ADPKD

Aberrant calcium signaling presumably contributes to the pathogenesis of ADPKD by influencing significant signaling pathways. The loss of PKD1 or TRPP2 due to pathogenic mutations results in lower cytosolic calcium levels in cystic cells compared to normal cells \((61, 105)\). This attenuation of cytosolic calcium in part leads to increased levels of cAMP and mitogen-activated protein kinase (MAPK), prompting the proliferation of cystic cells \((104)\). Of note, in normal renal epithelial cells, cAMP stimulates anti-proliferative pathways \((61)\). Vasopressin receptor V2 (VRV2) activation induces the release of cAMP within collecting ducts of the kidney \((51)\), and a VRV2 antagonist tolvaptan has appeared to be beneficial in a number of patients with ADPKD, within a clinical trial \((106)\).

1.3.6 Stoichiometric assembly of the PKD1/TRPP2 complex

Several findings are not consistent with regard to the assembly of homomeric and heteromeric TRPP2. When TRPP2 was integrated into lipid bilayers, homomeric TRPP2 channels were reported to be tetramers \((107)\). In agreement with these findings, several near full-length and full-length structures of TRPP2 have been solved, which reveal that the channel assembles into tetramers \((108–110)\). These TRPP2 structures will be characterized in Chapter IV. Additionally, dimerization motifs have been found at the N-terminus of TRPP2 (aa 199–207) in addition to a Cys632 within the third extracellular loop \((104, 111, 112)\). These dimerization motifs are presumably involved in the tetramerization of the channel \((104, 111, 112)\).
However, previous studies conducted in our lab by a former member, Dr. Yong Yu, revealed several interesting findings regarding TRPP2 assembly. Firstly, upon being heterologously expressed in HEK293 cells and Xenopus oocytes, TRPP2 appeared to have formed homomeric trimers (59). Secondly, the crystal structure of the coiled-coil domain (aa839–873) within the C-terminus of TRPP2 also formed a trimer. Thirdly, based on photobleaching experiments, it was proposed that trimeric TRPP2 and the PKD1 protein form a complex with 3:1 stoichiometry through their coiled-coil domains (59). Taken together, as seen in the structural studies involving full-length TRPP2, a homomeric tetramer exists when concentrated at high levels during protein purification (108–110). However, at lower concentrations, TRPP2 supposedly forms trimers (59). Therefore, how TRPP2 and PKD1 stoichiometrically form a complex at the structural level remains to be elucidated. The determination of the structure of the PKD1/TRPP2 complex will reveal important insights into biochemical interactions within this complex.

1.3.7 TRPP2 agonist triptolide

The identification of chemically and naturally derived agonists that interact with TRPP2 has been lacking. However, triptolide, a diterpenoid epoxide, has been shown to presumably bind to TRPP2 and activate the opening of the channel (113). This finding resulted from a cell fractionation study in which TRPP2 appeared to be a [3H] triptolide-binding protein (113). Triptolide is a natural product derived from Tripterygium wilfordii (Thunder God Vine), which has been utilized in traditional Chinese medicine (114). Leuenroth et al. reported that triptolide can initiate the release of Ca\(^{2+}\) by a mechanism that is dependent upon TRPP2 (113). Additionally, it has been shown that triptolide can inhibit cell proliferation and decrease cyst formation in PKD1\(^{-}\) embryonic mice (115). However, Leuenroth et al. mentioned that the agonistic effect on
TRPP2 by triptolide may be indirect (115). Moreover, a recent study reported that current activation by triptolide was not seen when TRPP2 was expressed in CHO or HEK293T cells (108). Therefore, how triptolide interacts with TRPP2 remains controversial in the field.

1.3.8 Interaction between TRPP2 and additional channels

Several other channels, primarily in the TRP superfamily, have been identified to form heteromeric complexes with TRPP2, including TRPV4 (116), TRPC1 (107, 117), and TRPC4 (104, 118). TRPP2 has been reported to associate with TRPC1 and TRPC4 in mesangial cells, giving rise to complexes that influence angiotensin II (118). Additionally, TRPP2 associates with TRPC1 at the primary cilia which can be activated by G-protein-coupled receptors (117). Furthermore, it has been shown that TRPP2 and TRPV4 form a functional channel within primary cilia (116). TRPP2 also forms a complex with the cardiac ryanodine receptor (RyR2), whereby TRPP2 can lead to the inhibition of RyR2 activity (119).

1.3.9 Ion selectivity in TRPP2

Over the years, inconsistent findings regarding the biophysical properties of TRPP2 currents have been reported. As previously mentioned, overexpressed TRPP2 is retained in the ER (86). Therefore, functional characterization of TRPP2 at the plasma membrane has been challenging since channel activity cannot be readily measured (86). However, upon removal of the ER retention motif, TRPP2 is able to be studied at the plasma membrane (55, 86, 120). Additionally, it has been found that the purported F604P gain-of-function mutation in TRPP2 leads to the constitutive opening of the channel when heterologously expressed in Xenopus oocytes, producing an observable current (121). However, currents were not able to be reproduced in another study using
the transfected F604P construct in HEK293T and CHO cells (108). Therefore, TRPP2 containing the F604P mutation may display different trafficking patterns to the plasma membrane depending on the expression system utilized.

There is general agreement that TRPP2 is considered to be a Ca\(^{2+}\)-permeable non-selective cation channel (51–53, 55). However, several findings are not consistent with regard to the specific selectivity properties of Ca\(^{2+}\) in TRPP2. Of note, the differences observed in channel properties of TRPP2 may be due in part to the number of biophysical systems utilized (51–53, 55, 73). It has been reported that TRPP2 has a slightly higher selectivity of Ca\(^{2+}\) compared to K\(^+\) and Na\(^+\) (120, 122), whereas other studies suggested that TRPP2 has limited selectivity for Ca\(^{2+}\) (51, 53). In accordance with the latter findings, a recent study demonstrated that a TRPP2-TRPP3 chimera is more selective for K\(^+\) and Na\(^+\) than Ca\(^{2+}\), compared to native TRPP3 (108). Additionally, although most studies indicate that TRPP2 conducts Ca\(^{2+}\) (55), it has also been demonstrated to be blocked by Ca\(^{2+}\) (90). Evidently, many discrepancies exist with regard to the channel activity of TRPP2.

1.3.10 Additional members of the TRPP subfamily

TRPP3

TRPP3, also regarded as PKD2L1, is a homologue of TRPP2 which is composed of 805 amino acids. TRPP3 contains residues that share 50% sequence identity to TRPP2 (123). These homologous regions are apparent in the six transmembrane segments and the I-II linker domain that is present between S1-S2 (123). Like TRPP2, TRPP3 harbors a putative Ca\(^{2+}\) binding EF-hand domain followed by a coiled-coil domain at its C-terminus (124). TRPP3 is primarily expressed in the kidney, testis, non-myocyte cardiac tissue, and neurons (70, 124–126). Since TRPP3 does not contain an ER retention motif at its C-terminus (96, 127) it is thereby functionally expressed at the
plasma membrane in heterologous systems, including *Xenopus* oocytes (128). Therefore, TRPP3 can be investigated through electrophysiological methods where it has been shown to form a constitutively active Ca\(^{2+}\)-permeable non-selective cation channel (123, 128).

**TRPP5**

Another TRPP2 homologue is TRPP5, also referred to as PKD2L2. TRPP5 contains 624 amino acids and exhibits 71% sequence similarity with TRPP2, lacking homology at the N- and C-termini (127). Similar to TRPP3, TRPP5 does not harbor an ER retention motif at its C-terminus (127). TRPP5 has been shown to be expressed in mouse testis, mouse heart, and the developing mouse oocyte (126, 129, 130). Therefore, TRPP5 may physiologically function as a non-selective cation channel associated with Ca\(^{2+}\) signaling in germ line cells (70).

### 1.4 Structural determination of TRP channels

Expressing and crystallizing domains of membrane proteins, known as the divide and conquer method, is a technique that has been utilized extensively over the past ten years to obtain structural information of significant portions of membrane proteins (131). Several TRP channel domains were able to be solved using this method, including the EF-hand (54, 57, 58) and coiled-coil domains (59) of TRPP2. However, these structures provided only pieces of the overall TRP channel puzzles, and further functional information was needed in order to discern how these structures fit into the overall context of the full-length channels. Yet, the ultimate goal in the field was to determine the structures of full-length TRP channels at atomic resolution.

Notably, over the past few years great advancements were made in the field of single-particle cryo-electron microscopy (cryo-EM), primarily by the development of new direct electron
detectors and software algorithms (132, 133). These technological improvements prompted the ‘resolution revolution,’ which enabled the determination of membrane protein structures at near-atomic resolution (133). Remarkably, for the first time small ion channels could be solved at high-resolution using the cryo-EM technique, pioneered by the determination of the TRPV1 structure in 2013 (134, 135). Consequently, this groundbreaking study sent ripples not only throughout the TRP channel space, but throughout the entire field of structural biology. Of note, before the ‘resolution revolution’ X-ray crystallography was the primary method to solve protein structures at high-resolution. However, many crystallographers began to shift gears and transition to using the cryo-EM technique as well. Intriguingly, the number of cryo-EM publications have dramatically increased and led to the determination of several other high-resolution TRP channel structures, including TRPV2 (136, 137), TRPA1 (138), and most recently TRPP2 (108–110).
1.5 Conclusion and thesis overview

It has certainly been fascinating to witness the recent ‘resolution revolution’ occurring in the cryo-EM field during the duration of my PhD experience. It is amazing that within a few years the TRP channel structural field has made significant strides from determining isolated domains of TRP channels using X-ray crystallography and NMR to solving full-length structures of these channels, primarily by cryo-EM. I initially started off my PhD being trained in X-ray crystallography, only to see firsthand several years later that high-resolution membrane protein structures can be obtained with cryo-EM as well. Therefore, this thesis primarily serves to elucidate the structural investigations of TRPML1 and TRPP2 using both X-crystallography and cryo-EM methods. Functional studies of TRPML1 were performed to further support and characterize our findings.

My initial studies pertained to the structural determination of the TRPML1 I-II linker domain at two physiologically relevant pH conditions. During that time, the divide and conquer approach was being utilized extensively by many research groups to uncover structural information of TRP channels, and our rationale was to apply this method as well. Heretofore, structural knowledge of the TRPML1 channel has been lacking. Furthermore, the I-II linker structures provided a molecular basis from which to functionally assess the regulatory mechanisms of this domain. Of note, we demonstrated that these mechanisms involved the dual regulation of Ca\(^{2+}\) and pH of TRPML1. Additionally, these static structures along with functional investigations were able to provide insights into the dynamic pathogenic mechanisms of several MLIV-causing mutations. Therefore, a detailed description of our published results (9) pertaining to the TRPML1 I-II linker will be delineated in Chapter II.
The second topic of investigation that I focused on was to determine a high-resolution cryo-EM structure of TRPP2. This endeavor would provide a greater understanding of the TRPP subfamily, as well as insights into ADPKD. Since the first near-3.Å resolution cryo-EM structure of an ion channel was of a TRP channel, this major achievement significantly affected the nature of our research. Therefore, our mindset was shifted from utilizing X-ray crystallography to determine the structure of TRPP2, to using cryo-EM. Chapter III serves to elucidate the protein purification and sample preparation methods for TRPP2 structural analysis. In collaboration with the laboratory of Dr. Reza Khayat (City College of New York) and the laboratory of Dr. Xueming Li (Tsinghua University), we were able to test the purified TRPP2 protein samples by cryo-EM. We were mindful of the potential difficulties that would arise during this endeavor, as well as the competitive nature of the TRP channel structural biology field. Until recently, no structural information pertaining to full-length TRPP2 has been reported. However, several months ago, a number of groups published TRPP2 full-length or near full-length structures. A comprehensive analysis of methods utilized in these studies as well as my own will be described at length in Chapter III and the structural details of TRPP2 will be elucidated in Chapter IV.
1.6 References


55. K. Hanaoka *et al.*, Co-assembly of polycystin-1 and -2 produces unique cation-permeable


95. D. Mekahli et al., Polycystin-1 and polycystin-2 are both required to amplify inositol-trisphosphate-induced Ca\textsuperscript{2+} release. *Cell Calcium.* **51**, 452–458 (2012).


128. X. Z. Chen et al., Polycystin-L is a calcium-regulated cation channel permeable to calcium


Chapter II

Structural and Functional Studies of the TRPML1 I-II Linker

2.1 Introduction

As described previously in Chapter I, TRPML1 contains a linker domain between its first two transmembrane segments, which we refer to as the I-II linker. TRPML1 has been extensively studied in our laboratory over the years. Dr. Minghui Li, a senior research associate in our laboratory initially conducted investigations of this channel and determined the first crystal structure of the human TRPML1 I-II linker at pH 6.0. Subsequently, I then set out to determine the crystal structures of this linker at two physiologically relevant pH conditions, pH 4.5 and pH 7.5. The crystal structures of the TRPML1 I-II linker are the first high-resolution structures of a domain in the TRPML channel subfamily. Furthermore, I characterized several MLIV-causing mutations present in this domain by analyzing the subcellular localization patterns of mutant TRPML1 proteins using confocal fluorescence microscopy.

The TRPML1 I-II linker structures served as a foundation from which to investigate pH and divalent ion regulation of the channel. Dr. Wei Kevin Zhang, a former postdoctoral fellow in our lab, performed the majority of the electrophysiology experiments pertaining to this project. Taken together, our efforts culminated in a triple co-first author publication (see Appendix A) describing the studies in this chapter (1). Essentially, this chapter elucidates the structural and functional studies of the TRPML1 I-II linker, which provide insight into the Ca^{2+} and pH regulation of the channel as well as the MLIV-causing mutations in this domain.
2.2 Experimental procedures for structural studies

2.2.1 Protein expression and purification of the TRPML1 I–II linker

A region of MCOLN1 encompassing the I–II linker (Gly84-Ser296) was obtained by PCR and sub-cloned into a modified pET26b(+) vector using BamHI/XhoI restriction sites. The construct utilized for structural studies was comprised of an N-terminal maltose-binding protein (MBP) tag followed by a thrombin-recognition site, the TRPML1 I–II linker, and a C-terminal hexahistidine tag (Figure 1A). This construct was transformed into Rosetta-gami 2 (DE3) cells to enable proper disulfide-bond formation within the I-II linker. The cells were cultured in LB medium at 37°C with 50 µg/ml kanamycin and 34 µg/ml chloramphenicol in an incubator shaker at 250 r.p.m. When the optical density at 600 nm (OD$_{600}$) reached 0.8-1.0, the culture was cooled to 22°C, and the cells were induced for 12 hours with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG).

The cells were harvested by centrifugation at 3,000g for 15 min and were resuspended in a solution containing 300 mM NaCl, 50 mM sodium phosphate pH 8.0, and 2.5% (w/w) glycerol (solution A). Additionally, 0.5 mg/ml lysozyme, 25 µg/ml DNase, 2mM phenylmethylsulfonyl fluoride (PMSF), and 5 mM imidazole were added, which was followed by sonication and centrifugation at 17,000g for 30 min. The cell lysate was then incubated with Ni–NTA His-Bind resin (Novagen) for 1 hour at 4°C. Subsequently, the beads were spun down at 800g for 1 min and transferred to a disposable polyprep column. The resin was then washed with solution A containing 5 mM imidazole and the protein was eluted from the resin with solution A containing 500 mM imidazole. The eluted I-II linker protein was subsequently incubated with amylose resin (NEB) at 4°C for 2.5 hours, collected by centrifugation at 800g for 1 min, and washed with solution A in a polyprep column. Thereafter, the I-II linker protein was eluted with 20 mM maltose in solution A and thrombin (Sigma-Aldrich) was added at 4 units per milligram of protein in order to cleave the
MBP tag. Following incubation at 16°C overnight, the protein was concentrated and purified by gel filtration chromatography using a Superdex 200 column (GE Healthcare). The gel filtration buffer contained 150 mM NaCl and 10 mM HEPES pH 7.5. The peak fractions corresponding to the tetrameric I–II linker protein were collected (Figure 1B) and concentrated by ultrafiltration using the Amicon Centrifugal Filter Units (EMD Millipore) prior to crystallization. The expression and purification protocols for the TRPML1 I–II linkers at pH 4.5 and pH 7.5 were comparable to the methods described above for the I-II linker at pH 6.0 (1).

**Figure 1. Expression and purification of the TRPML1 I-II linker.** (A) Schematic representation of the construct containing the TRPML1 I-II linker in a modified pET26b (+) vector. The N-terminal MBP tag was used to increase the purity of the protein during purification, yet was cleaved off by thrombin before loading onto the gel filtration column. (B) Representative gel filtration profile of the TRPML1 I-II linker. Peaks eluted from a Superdex 200 column correspond to the I-II linker and cleaved MBP, respectively.
2.2.2 Protein crystallization of the TRPML1 I-II linker

Crystallization of the TRPML1 I-II linker at pH 6.0

Crystallization of the human TRPML1 I-II linker was conducted using the hanging drop vapor diffusion method at 16°C. The protein concentration was 4 mg/ml and the reservoir solution contained 1.38 M sodium phosphate monobasic monohydrate, 0.42 M potassium phosphate dibasic pH 6.0, and 5% pentaerythritol ethoxylate (3/4 EO/OH) (Hampton Research). The protein solution was mixed with the reservoir solution at a 1:1 ratio. For the preparation of heavy-atom-derivative crystals, crystals were soaked in a solution containing 1.53 M sodium phosphate monobasic monohydrate, 0.47 M potassium phosphate dibasic pH 6.0, and 1 mM K2Pt(CN)4 for 24 hours. Crystals were cryoprotected with Paratone-N (Hampton Research).

Crystallization of the TRPML1 I-II linker at pH 4.5 and pH 7.5

To facilitate crystallization of the human TRPML1 I-II linker at pH 4.5 and pH 7.5, a flexible region containing residues R200-E213 was removed by overlapping PCR from the Gly84-Ser296 linker and all constructs were verified by sequencing. Of note, within the pH 6.0 structure, the flexible region containing Glu199-Lys219 was unresolved. Crystallization of the TRPML1 linker at pH 4.5 and pH 7.5 was carried out utilizing the hanging drop vapor diffusion method at 20°C. Crystals for the TRPML1 I-II linker at pH 4.5 were obtained at a protein concentration of 4 mg/ml and a reservoir solution containing 200 mM magnesium sulfate, 5.3% PEG 3350, and 100 mM acetate pH 4.5 (Figure 2A). Crystals at pH 7.5 were macroseeded at a protein concentration of 4 mg/ml in a reservoir solution containing 100 mM magnesium sulfate, 4% PEG 3350, and 100 mM HEPES pH 7.5 (Figure 2B). For all conditions, the protein solution was mixed with reservoir solution at a 1:1 ratio. Crystals at pH 4.5 were cryoprotected with Paratone-N (Hampton Research).
and crystals at pH 7.5 were cryoprotected using 30% glycerol. All crystals were flash-frozen in liquid nitrogen for data collection at 100 K.

**Figure 2. Crystals of the TRPML1 I-II linker at pH 4.5 and pH 7.5.** (A-B) Crystals of the TRPML1 I-II linker domain at pH 4.5 (A) and pH 7.5 (B). Crystallization conditions are specified in the experimental procedures. TRPML1 I-II linker crystals were obtained from a modified construct, omitting 14 amino acid residues that were previously unresolved in the pH 6.0 structure.

### 2.2.3 Data collection and structure determination

**TRPML1 I-II linker at pH 6.0**

X-ray diffraction data for native and heavy atom derivative K$_2$Pt(CN)$_4$ TRPML1 I-II linker crystals at pH 6.0 were collected at 100 K on a RAXIS-IV detector with Cu K radiation ($\lambda = 1.5418$ Å) from a Rigaku RuH3R X-ray generator. Diffraction images were processed and scaled with the HKL package (2). The X-ray diffraction data set was collected at 2.3 Å and 2.6 Å resolution, for the I-II linker and the heavy atom derivative K$_2$Pt(CN)$_4$ I-II linker, respectively. The crystal belonged to the I422 space group with cell parameters of $a = b = 125.3$ Å, $c = 76.7$ Å, $\alpha = \beta = \gamma = 90.0^\circ$. The asymmetric unit contained a single monomeric subunit. The structure at pH 6.0 was determined with the single isomorphous replacement with anomalous scattering (SIRAS) method.
using native and platinum-derivative data sets. Shake and Bake (3) was used to identify two platinum sites which were used in SOLVE (4) to determine the initial phases (1).

Additionally, density modification was carried out with RESOLVE (4) in order to ameliorate phase accuracy. The quality of the density map was reasonably sufficient to locate the majority of the residues and the initial atomic model of the I-II linker was built into the electron density map using COOT (5). Structural refinement with CNS (6) and manual rebuilding were carried out iteratively prior to deposition in the Protein Data Bank (accession code: PDB 5TJA). Since the density was missing for Glu199-Lys219, the final model contained amino acid residues 84-198 and 220-293. Data collection and refinement statistics are summarized in Table 1.

**TRPML1 I-II linker at pH 4.5**

X-ray diffraction data for the TRPML1 I-II linker at pH 4.5 were collected at 100 K on a RAXIS-IV detector with Cu K radiation (λ = 1.5418 Å) from a Rigaku RuH3R X-ray generator. All diffraction images were processed and scaled with the HKL package (2). The X-ray diffraction data set was collected at 2.4 Å resolution, and the crystal belonged to the P42_12 space group with cell parameters of a = b = 94.4 Å, c = 50.7 Å, α = β = γ = 90.0°. The structure of the I–II linker at pH 4.5 was solved by molecular replacement with the program PHASER (7) using the structure at pH 6.0 as the search model. The atomic model was built using COOT and structural refinement was carried out using CNS and deposited into the Protein Data Bank (accession code: 5TJB). Crystalloigraphic statistics are summarized in Table 1.
TRPML1 I-II linker at pH 7.5

X-ray diffraction data at pH 7.5 were collected at the National Synchrotron Light Source (NSLS) beamline X29 at Brookhaven National Laboratory. Diffraction images were processed and scaled with the HKL package (2). The X-ray diffraction data set was collected at 2.4 Å resolution, and the crystal belonged to the F432 space group with cell parameters of \( a = b = c = 182.8 \, \text{Å}, \, \alpha = \beta = \gamma = 90.0^\circ \). The structure of the I–II linker at pH 7.5 was also solved by molecular replacement with the program PHASER (7) using the structure at pH 6.0 as the search model. The atomic model was built using COOT and cycles of structural refinement were carried out using CNS and deposited into the Protein Data Bank (accession code: 5TJC). Data collection and refinement statistics are summarized in Table 1.
Table 1. Data collection and refinement statistics. Published in (1).

<table>
<thead>
<tr>
<th></th>
<th>TRPML1 I–II linker, pH 6.0 (PDB 5TJA)</th>
<th>TRPML1 I–II linker, pH 6.0, K$_2$Pt(CN)$_6$ derivative (PDB 5TJB)</th>
<th>TRPML1 I–II linker, pH 4.5 (PDB 5TJC)</th>
<th>TRPML1 I–II linker, pH 7.5 (PDB 5TJD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>I422</td>
<td>I422</td>
<td>P42$_2$2</td>
<td>F432</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>125.3, 125.3, 76.7</td>
<td>125.3, 125.3, 76.5</td>
<td>94.4, 94.4, 50.7</td>
<td>182.8, 182.8, 182.8</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.5418</td>
<td>1.5418</td>
<td>1.5418</td>
<td>1.0750</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.0–2.3 (2.38–2.30)$^a$</td>
<td>20.0–2.6 (2.69–2.60)$^a$</td>
<td>20.0–2.4 (2.49–2.40)$^a$</td>
<td>20.0–2.4 (2.49–2.40)$^a$</td>
</tr>
<tr>
<td>R$_{sym}$ (%)</td>
<td>9.9 (37.7)</td>
<td>13.8 (64.4)</td>
<td>6.8 (43.4)</td>
<td>5.9 (34.0)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>20.0 (4.6)</td>
<td>14.6 (2.8)</td>
<td>19.0 (2.8)</td>
<td>22.3 (3.7)</td>
</tr>
<tr>
<td>CC$_{1/2}$</td>
<td>0.982 (0.944)</td>
<td>0.961 (0.866)</td>
<td>0.968 (0.872)</td>
<td>0.969 (0.882)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (99.6)</td>
<td>99.9 (99.8)</td>
<td>99.3 (98.3)</td>
<td>98.0 (98.6)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>7.2 (6.1)</td>
<td>6.9 (6.4)</td>
<td>4.7 (3.8)</td>
<td>4.2 (4.1)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.0–2.3</td>
<td></td>
<td>20.0–2.4</td>
<td>20.0–2.4</td>
</tr>
<tr>
<td>No. reflections</td>
<td>13,298 (1,337)</td>
<td></td>
<td>8,777 (927)</td>
<td>9,088 (1,033)</td>
</tr>
<tr>
<td>R$<em>{free}$/R$</em>{free}$</td>
<td>0.228 / 0.253</td>
<td></td>
<td>0.212 / 0.252</td>
<td>0.216 / 0.253</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>1,518</td>
<td></td>
<td>1,268</td>
<td>1,426</td>
</tr>
<tr>
<td>Water</td>
<td>111</td>
<td></td>
<td>104</td>
<td>66</td>
</tr>
<tr>
<td>B factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>33.0</td>
<td></td>
<td>42.1</td>
<td>50.3</td>
</tr>
<tr>
<td>Water</td>
<td>37.2</td>
<td></td>
<td>45.5</td>
<td>47.6</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.010</td>
<td></td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.7</td>
<td></td>
<td>1.5</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Each data set was collected from a single crystal. $^a$Values in parentheses are for the highest-resolution shell.
2.3 Structural features of the TRPML1 I-II linker structures

2.3.1 TRPML1 I-II linker structure at pH 6.0

We first determined the 2.3 Å X-ray crystal structure of the TRPML1 I-II linker (Gly84–Ser296) at pH 6.0 (Figure 3 and Table 1), mimicking the pH of an endosome. As previously mentioned, residues Glu199–Lys219, comprising a region of the loop between β4 and β5 were unresolved, most likely due to its flexibility. A prominent feature of the TRPML1 linker structure is that a monomer exists in the asymmetric unit, which forms a tetramer through a four-fold rotational crystallographic symmetry axis (Figures 3A, B). The structure has a height of 30 Å and the diameter at the widest point is 90 Å (Figures 3A, B).

Figure 3. Crystal structure of the TRPML1 I-II linker at pH 6.0. (A) Ribbon representation of the I-II linker structure. Monomeric subunits are highlighted with different colors. (B) Electrostatic potential surface representation of the tetrameric structure. (A-B) Upper panels illustrate views looking down the four-fold symmetry axis from the extracellular surface of the membrane. Lower panels indicate side views displayed parallel to the plane of the membrane. Published in (J).
Each subunit contains 213 amino acids and is comprised of 2 long α-helices (α1 and α2), two short α-helices (α3 and α4), and eight β-strands (Figures 3A, 4A). In addition, three loops are present between α1 and α2, α3 and β1, and β4 and β5. The four α-helices are arranged in a compact manner against a central five-stranded antiparallel β-sheet consisting of β1, β4, β5, β6, and β7. The first 22 amino acids at the N-terminus of the I-II linker form α1 (Figure 3A), which presumably links to the first transmembrane segment in the full-length channel. In addition α2 constructs a level surface atop the domain (Figure 3A) (I). Of significance, an intrasubunit disulfide bond exists between Cys253 in β6 and Cys284 in the short loop between β7 and β8 (I).

Notably, domain-swapping is evident, as a 3-stranded beta-sheet is formed between β2 and β3 (green) within one monomer and β8 from an adjacent monomer (Figure 4A). Similarly, β8 (purple) from one monomer forms another 3-stranded beta-sheet with β2 and β3 in another adjacent monomer (Figure 4A). The formation of this antiparallel three-stranded β-sheet between two adjacent subunits is also illustrated in Figure 3A. The overall architecture of the TRPML1 I-II linker is similar to that of the I-II linker in TRPP2, which likewise forms a tetrameric structure (8–10). The structure of TRPP2, as well as the comparison between the linker domains in TRPP2 and TRPML1 will be elucidated in Chapter IV.

2.3.2 Features of the central luminal pore

At the center of the tetramer exists a pore, which we refer to as the luminal pore. This region is 14Å at its narrowest point (Figure 3A), which would enable hydrated K⁺, Na⁺, and Ca²⁺ ions to translocate through the channel. Intriguingly, the central luminal pore has several distinctive features. Firstly, a 16–amino acid (L106–T121) luminal pore loop links α1 and α2 (Figures 4A, B). Secondly, 12 aspartates in the tetrameric structure (Asp111, Asp114, and Asp115 in each
subunit) are situated in the luminal pore loops (Figure 4B). Thirdly, at neutral pH, these aspartates generate a highly electronegative region (Figure 4C). The role that this luminal pore plays in the Ca\(^{2+}\)/pH regulation of TRPML1 will be detailed in section 2.4.

Figure 4. Features of the central luminal pore in the TRPML1 I-II linker. (A) Stereo view of the monomeric I-II linker. (B) Enlarged stereo side view representation of the 16 residue luminal pore loop (L106–T121). (C) View looking down the central pore of the I-II linker, illustrating the negative charges (black circles) of the aspartate residues. Modified figure published in (1).
2.3.3 TRPML1 I-II linker structures at pH 4.5 and pH 7.5

After solving the I-II linker structure at pH 6.0, we then set out to ascertain whether TRPML1 I-II linker structures at additional physiologically relevant pH values would differ from one another. Since TRPML1 channels primarily localize to lysosomes and the I-II linker faces the lysosomal lumen, we solved the crystal structure of the I-II linker at pH 4.5, mimicking its native physiological state. In addition, we determined the I-II linker crystal structure at pH 7.5, simulating the pH at the plasma membrane where TRPML1 can be present during exocytosis (11). At the plasma membrane, the I-II linker would be surrounded by the extracellular fluid.

Remarkably, despite the changes in pH, the structures at pH 4.5 and pH 7.5 were nearly identical to the original pH 6.0 structure when superimposed (Figure 5A). After closer analysis, we detected local minor conformational changes among the different pH structures near the periphery. Notably, the structure at pH 6.0 has the longest α-helix near the N-terminus, while the pH 4.5 structure has the shortest (Figure 5B). The first nine amino acids of α1 within the pH 4.5 structure were therefore unresolved. Moreover β2, β3, and β8 were not able to be fully resolved in the pH 4.5 structure. We presume that these differences can be attributed to the distinct crystal packing at the various pH conditions. Intriguingly, even the orientation of the aspartate side chains within the luminal pore loop region (Figure 5C) were nearly identical. Taken together, the I–II linker appears to be in a highly stable state since significant conformational changes were not observed in the structures at different pH conditions (1).
Figure 5. Comparison of the TRPML1 I–II linker structures. (A) Superposition of the I–II linker crystal structures obtained at pH 4.5 (magenta), pH 6.0 (yellow) and pH 7.5 (cyan). Orientation of the structures are illustrated as seen from the extracellular side of the membrane. (B) A close-up view indicating slight differences among the overlaid structures. (C) Enlarged view of the luminal pore in the superimposed structures.
2.3.4 Verification of the tetrameric I-II linker

In order to determine whether the tetrameric structure of the TRPML1 I-II linker is indeed present within full-length TRPML1, our collaborators in the laboratory of Dr. Xueming Li (Tsinghua University) performed cryo-EM analysis of purified protein to visualize the linker. Thus, the cryo-EM density of full-length wild-type C. elegans TRPML1 was visible for the extracellular domain, however, the density was indeterminate for the transmembrane domains. A masking technique was then implemented to partition the densities of the extracellular and transmembrane domains. The identity of the densities of these regions were decipherable based on their dimension as well as the resemblance of the transmembrane domain of TRPML1 to that of TRPV1 (1).

Subsequently, a cryo-EM density map of the extracellular domain was obtained at an overall resolution of 5.28 Å with $C_4$ symmetry (Figure 6A). Despite the low-resolution electron density map, approximately 79% of the density comprised the I-II linker. The extra density could be attributed to the E199-K219 loop and/or glycosylation. The overall density revealed a tetramer with a large central pore, hence verifying that the I-II linker is present as a tetramer both within the isolated domain as well as within the context of the overall full-length structure. Moreover, the TRPML1 I-II linker crystal structure fitted quite reasonably into the cryo-EM density map (Figure 6B). Since we could not obtain an electron density that would yield a high-resolution structure, an atomic model was not built (1).
Figure 6. Tetrameric structure of the I-II linker verified in full-length TRPML1. (A-B) Cryo-EM reconstructions of the TRPML1 I-II linker present within full-length TRPML1. The gray envelopes signify cryo-EM reconstructions of the TRPML1 extracellular domain at 5.28 Å resolution (A) with the crystal structure of the I-II linker docked within the cryo-EM density (B). Upper panels indicate views looking down the four-fold rotational crystallographic symmetry axis as seen from the extracellular/luminal side of the membrane. Lower panels denote views parallel to the membrane. Published in (I).
2.3.5 Role of the I-II linker in TRPML1 assembly

Next we wanted to investigate the role of the I-II linker structure in influencing the tetramerization of full-length TRPML1. We proceeded to simultaneously mutate two residues that are involved in intersubunit interactions, Leu144 and Arg146 (Figure 7A), to L144K and R146S in both the I-II linker structure as well as in the full-length TRPML1 channel. Dr. Minghui Li then assessed their oligomeric states by native gel electrophoresis. Upon examination of full-length TRPML1 with and without the L144K R146S mutation in a nondenaturing blue native gel, a band corresponding to a tetramer was present for full-length wild-type TRPML1 but not for the mutant channel. Instead, full-length TRPML1 with the L144K R146S mutation resulted in aggregated oligomers. Consistent with these findings, the I-II linker alone containing the L144K R146S mutation migrated at a low molecular weight in a nondenaturing gel, indicative of a monomer, while the wild-type I-II linker migrated as an apparent tetramer. Taken together, these results signify that the L144K R146S mutation appeared to have interfered with the tetrameric assembly of both the I-II linker domain as well as the full-length TRPML1 channel (1).

Subcellular localization of TRPML1-LR

Consequently, I performed confocal fluorescence imaging in order to investigate whether the subcellular localization pattern of the full-length TRPML1-LR channel, containing the L144K R146S mutation, was altered compared to that of wild-type. Detailed methods for cell culture and live-cell confocal imaging are provided in section 2.5.2. Briefly, Hela cells were transfected with wild-type or mutant (L144K R146S) full-length TRPML1 tagged with EGFP on the N-terminus. To visualize the lysosomes, Hela cells were loaded with the lysosomal marker LysoTracker® Red DND-99 (Invitrogen). Consistent with previous imaging studies performed by other groups, wild-
type TRPML1 localized to the lysosomes (12–14). Conversely, TRPML1-LR did not localize to lysosomes and instead was distributed throughout the cell (Figure 7B). These results further verified that the tetramerization of the I-II linker is essential for the association of the full-length tetrameric structure and revealed that the intact linker is necessary for proper subcellular localization (I).

Figure 7. Role of intersubunit interactions in TRPML1 assembly. (A) Intersubunit interactions within the TRPML1 I-II linker. Schematic representation of an intersubunit interface within the TRPML1 I-II linker, highlighting intersubunit interactions. Residues Leu144 and Arg146 are denoted in red boxes, alongside their interacting partners. (B) Confocal images of Hela cells expressing either GFP-tagged full-length wild-type TRPML1 (top panel) or full-length TRPML1 containing the L144K R146S mutation (bottom panel). Red illustrates lysosomes loaded with LysoTracker Red DND-99 and yellow indicates colocalization of GFP-tagged TRPML1 and lysosomes. Only wild-type TRPML1 colocalized with the lysosomal marker. The indicated scale bar represents 5µm and is applicable to all panels. Published in (I).
2.4 Structural basis of pH and Ca\(^{2+}\) regulation of TRPML1

2.4.1 Dual regulation of TRPML1 activity by pH and Ca\(^{2+}\)

In order to measure human TRPML1 currents, a well-characterized constitutively active mutant channel, TRPML1\(^{VP}\), was utilized in HEK293T cells. As described in Chapter I, this mutation allows TRPML1 to traffic to the plasma membrane in order for electrophysiological studies to be carried out (15, 16). TRPML1\(^{VP}\) containing an N-terminal EGFP tag produced a strong inward-rectifying current in a nominal divalent-ion-free (NDF) extracellular/luminal solution at pH 7.4 upon investigation with whole-cell patch-clamp recordings. The current was inhibited in a dose-dependent manner upon addition of Ca\(^{2+}\) to the extracellular/luminal solution. Furthermore, a dose–response curve exhibited an IC\(_{50}\) (calcium concentration producing half maximal inhibition) of 0.27 mM, with a Hill coefficient of 1 (Figure 8A). Thus, these results indicate a one-to-one Ca\(^{2+}\) blocking mechanism (1).

Moreover, Ca\(^{2+}\) blockage of the TRPML1 channel was greatly dependent on the extracellular or lysosomal luminal pH (7.4 and 4.6, respectively), and decreased upon reduction of pH from 7.4 to 4.6. As illustrated by Figure 8A, at pH 4.6 the dose–response curve was shifted to the right, increasing the apparent IC\(_{50}\) by 14-fold, from .27 mM to 3.8 mM. Notably, the shape of the dose–response curve was altered and the Hill coefficient was changed from 1 to 0.5. These findings therefore alluded to the presence of negative cooperativity between extracellular/luminal Ca\(^{2+}\) and H\(^{+}\). Given that Ca\(^{2+}\) inhibition at pH 4.6 decreased, TRPML1\(^{VP}\) currents increased when the extracellular/luminal pH was reduced from 7.4 to 4.6 (Figure 8B). Of note, since the V432P mutation was implemented in these studies, resulting in a constitutively open channel, the consequences of the dual regulation by luminal Ca\(^{2+}\) and pH on gating may not be detected. Dr. Wei Kevin Zhang in our lab performed the electrophysiology experiments (1).
2.4.2 The central pore is essential for pH and Ca\textsuperscript{2+} regulation

As previously mentioned, the central pore of TRPML1 is highly electronegative, which likely intensifies Ca\textsuperscript{2+} conduction by attracting and concentrating Ca\textsuperscript{2+} through an intricate and pH dependent process. Additionally, upon Ca\textsuperscript{2+} binding, the central pore in the I-II linker can be obstructed, resulting in reduced Ca\textsuperscript{2+} conduction. A model for pH and Ca\textsuperscript{2+} regulation will be detailed in section 2.4.3.

In order to further investigate what function the 12 aspartate residues have in the central pore, we mutated all of them concurrently to glutamine within the TRPML1\textsuperscript{VP} construct, hence creating the TRPML1\textsuperscript{VP}-3DQ channel. At pH 7.4, upon addition of Ca\textsuperscript{2+}, the dose-response curve for

Figure 8. Regulation of TRPML1 by Ca\textsuperscript{2+} and pH. (A) Dose-response curves indicating Ca\textsuperscript{2+} inhibition of TRPML1\textsuperscript{VP} at pH 7.4 and pH 4.6. Error bars represent the standard error of the mean (SEM). The number of recordings are displayed in parentheses. Curves are fits to the Hill equation. (B) Time course of TRPML1\textsuperscript{VP} currents at pH 7.4 and pH 4.6 with 1 mM Ca\textsuperscript{2+} and either Na\textsuperscript{+} or NMDG\textsuperscript{+} as the charge carrier. Published in (1).
TRPML1\textsuperscript{VP} exhibited an IC\textsubscript{50} of .27 mM (Figure 9A). However, upon examination of TRPML1\textsuperscript{VP}-3DQ, the dose-response curve shifted to the right, significantly increasing the apparent IC\textsubscript{50} by 20-fold to 5.5 mM, maintaining the Hill coefficient at 1 (Figure 9A). Therefore, TRPML1\textsuperscript{VP}-3DQ greatly led to the reduction in Ca\textsuperscript{2+} inhibition at pH 7.4. As described earlier, Ca\textsuperscript{2+} inhibition of TRPML1\textsuperscript{VP} (Figures 8A, B) was reduced when the pH was lowered from 7.4 to 4.6. Interestingly, the TRPML1\textsuperscript{VP}-3DQ channel led to additional reduction of Ca\textsuperscript{2+} inhibition at pH 4.6, thus increasing the apparent IC\textsubscript{50} from 3.8 mM to >10 mM (Figure 9B). Taken together, these results suggest that the 12 aspartate residues in the central luminal pore are essential for Ca\textsuperscript{2+} and pH regulation of TRPML1 (1).

Figure 9. Role of luminal pore aspartate mutations in Ca\textsuperscript{2+} regulation. (A-B) Dose-response curves of Ca\textsuperscript{2+} inhibition of TRPML1\textsuperscript{VP} or TRPML1\textsuperscript{VP}-3DQ at pH 7.4 (A) and pH 4.6 (B). Measurements were taken at a potential of -80 mV. Error bars indicate the standard error of the mean (SEM). The number of recordings are indicated in parentheses and the curves are fits to the Hill equation. Published in (1).
We further investigated Ca$^{2+}$ conduction within the luminal pore by performing Ca$^{2+}$ imaging to determine Ca$^{2+}$ influx into HEK293T cells. These experiments were carried out by Deyuan Su at the Kunming Institute of Zoology. Upon cellular expression with TRPML1$^{VP}$, the extracellular solution utilized was altered from 0 to 3 mM Ca$^{2+}$, and an increase in concentration of intracellular Ca$^{2+}$ was subsequently observed over time. Notably, the rise in intracellular Ca$^{2+}$ levels occurred more rapidly at pH 4.6 than at pH 7.4. This finding likely signified quicker Ca$^{2+}$ conduction at pH 4.6 resulting from weakened Ca$^{2+}$ blockage as well as reduced Ca$^{2+}$ accretion at low pH.

Subsequently, TRPML1$^{VP-3DQ}$ was expressed and rise time in intracellular Ca$^{2+}$ concentration was comparable at pH 4.6 and pH 7.4. The increase in intracellular Ca$^{2+}$ in TRPML1$^{VP-3DQ}$ was comparable to that of TRPML1$^{VP}$ at pH 7.4, leading to strong Ca$^{2+}$ blockage. Conversely, at pH 4.6, TRPML1$^{VP-3DQ}$ exhibited a slower increase in Ca$^{2+}$ levels, possibly due to reduced Ca$^{2+}$ accumulation at the pore. In summary, these findings indicate that the 12 aspartates present within the central luminal pore contribute to Ca$^{2+}$ conduction within TRPML1 ($I$).

2.4.3 Model of pH and Ca$^{2+}$ dual regulation of TRPML1

Based on our structural and functional findings, we suggested a model for the dual Ca$^{2+}$ and pH regulation of TRPML1. At pH 7.4, the aspartates in the luminal-pore are generally negatively charged (Figure 10A). Hence, these negative charges through electrostatic interactions can attract, accumulate, and bind to extracellular/luminal Ca$^{2+}$. Therefore, the Ca$^{2+}$ ions can essentially block the luminal pore, resulting in decreased Ca$^{2+}$ conduction (Figure 10A).

In contrast, at pH 4.6 protonation of the aspartates occur, thus reducing the number of negative charges within this region, leading to decreased Ca$^{2+}$ block and increased Ca$^{2+}$ conduction (Figure 10B). Of note, protonation presumably occurs in acidic environments since pH 4.6 is similar to the
pKₐ of the aspartates. Yet, not all aspartates may become protonated at pH 4.6, which can then attract Ca²⁺ and still contribute to a low level of Ca²⁺ block. Our proposed model is further supported by the demonstration of negative cooperativity between Ca²⁺ and H⁺ at pH 4.6 (Figure 8A). When the aspartates are mutated to glutamines, the negative charges in the luminal pore are eliminated, thus resulting in an even greater reduction of the Ca²⁺ block which subsequently leads to increased Ca²⁺ conduction (Figure 10C).

In summary, with regard to Ca²⁺ conductivity, the acidic environment within the lysosome facilitates Ca²⁺ conductance within this organelle by decreasing the inhibitory effect of luminal Ca²⁺ on TRPML1. Alternatively, at the plasma membrane the neutral extracellular pH results in low levels of Ca²⁺ conductivity by intensifying the inhibitory effect of luminal Ca²⁺ on TRPML1 (I).
Figure 10. Schematic representation of Ca$^{2+}$/pH dual regulation of TRPML1. Depiction of the I-II linker structure, looking down the four-fold symmetry axis. Negative charges are indicated by orange dashes and Ca$^{2+}$ ions are signified by purple circles. (A) Illustration of the highly electronegative luminal pore at pH 7.4. Negative charges attract and obstruct the luminal pore upon accumulation of Ca$^{2+}$, thereby decreasing Ca$^{2+}$ conduction. (B) A reduction in pH induces decreased Ca$^{2+}$ hindrance, leading to increased Ca$^{2+}$ conduction. (C) All twelve aspartates are mutated to glutamines, hence eliminating negative charges in the luminal pore, resulting in increased Ca$^{2+}$ conduction.
2.5 Characterization of MLIV-causing mutations

2.5.1 MLIV-causing mutations in the TRPML1 I-II linker

The elucidation of the TRPML1 I-II linker structures provided us with a framework from which to investigate and characterize MLIV-associated missense mutations. Interestingly, three missense mutations, namely, L106P, C166F, and T232P, have been identified in human MLIV patients (17–21) and are located in the I-II linker domain (Figure 11). Wakabayashi et al. reported that these mutations manifest as milder symptoms compared to other MLIV truncation or frame-shift mutations, which typically lead to the complete loss of the TRPML1 protein (17). However, based on other studies, it appears that the severity of the disease stemming from the same mutation may be distinctly different from one individual to another. For instance, the T232P mutation led to a severe MLIV phenotype in one patient (18) yet manifested milder MLIV symptoms in another patient (22).

Intriguingly, the L106P mutant is present not only in the putative serine lipase consensus motif (L104-A113) (20, 23) located in the luminal pore loop, but also at the intersection between the pore loop and α1 (Figures 11A,B). Furthermore, T232P resides within the central β-sheet on β5 (Figures 11A, C) and C166F is located on β1 (Figures 11A, D). Cys166 forms an intrasubunit disulfide bond with Cys192 in β4. Notably, C166F disrupts this interaction and mutates a small amino acid to a substantially larger hydrophobic residue, while the other two MLIV mutations give rise to prolines (L106P and T232P). Additionally, Dong et al. determined that TRPML1VP with the T232P MLIV-causing mutation inhibited the constitutive activity of TRPML1VP (16). Taken together, it was plausible that these mutations could be detrimental to the folding, localization, function, and assembly of the channel.
Figure 11. Locations of MLIV-causing missense mutations in the TRPML1 I-II linker. (A) Schematic ribbon representation of the I-II linker tetrameric structure highlighting the locations of the MLIV-causing missense mutations L106P, T232P, and C166F, indicated by colored spheres. (B-D) Monomeric structures of the I-II linker depicting the specific locations of the missense mutations L106P (B), T232P (C), and C166F (D), denoted by red spheres. Published in (I).

2.5.2 Cell culture, transfection, and confocal live-cell imaging

Hela cells were maintained in DMEM (Gibco) containing 10% FBS (Atlanta Biologicals) and 100 µg/mL of penicillin/streptomycin (Invitrogen) at 37°C in a 5% CO₂ humid atmosphere. Hela cells were cultured and plated onto poly-D-lysine hydrobromide (Sigma) coated coverslips 18 to 24 hours prior to transfection. The cells were grown on the coverslips in DMEM plus 10% FBS supplemented with 100 µg/mL of penicillin/streptomycin at 37°C until they became about 80% confluent. Hela cells were then transfected with wild-type full-length TRPML1 as well as full-length TRPML1 containing mutations L144K R146S (LR), T232P, L106P, or C166F linked to EGFP on the N-terminus. Cells were transfected with 1µg/ml of DNA using LipoD293 In Vitro
DNA Transfection Reagent Ver II. (SignaGen Laboratories) following the manufacturer’s instructions. After 17 hours post-transfection, Hela cells were loaded with 100 nM of the lysosomal marker LysoTracker® Red DND-99 (Invitrogen) for 50 minutes at 37°C to visualize lysosomes in the cells. Cells were then gently washed three times with PBS (Gibco) and live-cell images were acquired in PBS at 20°C (1).

Confocal images were acquired from random live cells in an unbiased manner. The microscope used was a Nikon Eclipse TE2000-S equipped with a 60x oil-immersion objective, a Hamamatsu C9100-13 back-thinned EM-CCD digital camera (Hamamatsu Photonics Ltd. Instruments Pentamax), and a spinning disc head (Yokogawa). Z-stacks were acquired at an optical slice thickness of 0.3 μm using a Piezo Focus Drive (Ludl) and data were analyzed using the Volocity software (PerkinElmer). All z-stacks were deconvolved utilizing the iterative restoration function in Volocity with the iteration limit set to 40 and the confidence limit set to 99%. All images illustrated are derived from deconvolved z planes (1).

2.5.3 Confocal imaging of MLIV-causing mutations

Since MLIV is a lysosomal storage disorder, I investigated the subcellular localization of wild-type TRPML1 and TRPML1 containing three MLIV-causing mutations (L106P, T232P, and C166F) using confocal fluorescence microscopy. Since full-length wild-type and mutant TRPML1 constructs contained EGFP, their expression and locations in Hela cells could be readily visualized. Additionally, the use of LysoTracker® Red DND-99 clearly identified the lysosomes in the cells. Live-cell confocal imaging revealed a distinct difference in localization patterns between the cells transfected with wild-type and mutant TRPML1 constructs, as illustrated in Figure 12.
The wild-type TRPML1 channels localized to the lysosomes (Figure 12), as expected based on previously reported studies (12, 13). However, as shown in the confocal images, all three of the disease-causing mutations did not localize to the lysosomes (Figure 12), yet rather appeared to have been retained in the ER (1). These results are consistent with formerly reported imaging results of TRPML1 mutant proteins (12, 14). Therefore, these mutations presumably cause MLIV since they are not targeted to the lysosomes and retention of these mutants in the endoplasmic reticulum is presumably due to incorrectly folded proteins (12, 13). This plausible explanation is supported by our circular dichroism findings and biochemical analysis of these mutants (detailed in 2.5.4). Hence, residues Leu106, Thr232, and Cys166 may play a role in the correct transport of TRPML1 through the biosynthetic pathway, and missense mutations resulting in misfolded channels may lead to the rerouting of these proteins within the cell (12).
Figure 12. Subcellular localization of wild-type and MLIV-causing mutant TRPML1 channels. Confocal images of live Hela cells expressing full-length GFP-tagged wild-type TRPML1 or full-length TRPML1 harboring pathogenic mutations L106P, T232P, and C166F. Red denotes lysosomes labeled with LysoTracker® Red DND-99 and yellow signifies colocalization of GFP-tagged TRPML1 and lysosomes. Only wild-type TRPML1 localized to the lysosomes. The scale bar represents 5µm in all panels. Published in (1).
2.5.4 Biochemical analysis and circular dichroism spectroscopy

Experimental Procedures

TRPML1 I–II linkers containing wild-type or MLIV-causing mutations (L106P, T232P, and C166F) were expressed through viral infection in Hi5 insect cells (*Trichoplusia ni*) (Expression Systems) and subsequently purified by Dr. Minghui Li. Fluorescence-detection size-exclusion chromatography (FSEC) was performed with a spectrofluorometric detector RF-10AXL (Shimadzu) for fluorescence detection and for size-exclusion chromatography, a Superose 6 10/300 column (GE Healthcare) was utilized. Detailed methods regarding insect cell expression, purification, and FSEC analysis will be provided in Chapter III.

Circular dichroism spectroscopy

Circular dichroism spectroscopy, CD, was performed with a J-815 CD spectrometer (Jasco). TRPML1 I–II linker proteins at a concentration of 1 mg/ml were loaded into a quartz cuvette with a path length of 0.01 cm for measurement. A sample containing only buffer was used for subtraction of the baseline signal. Measurements were obtained from 185 to 250 nm with a 0.1-nm interval, 1-nm bandwidth, and scanning speed of 50 nm/min. For each TRPML1 I–II linker protein, three samples were measured, and three accumulations were measured for each sample and averaged. The MRE (mean residue ellipticity) was computed from the raw CD signal, θ, using the equation \[ [\theta]_{MRE} = \theta/(10 \times Cr \times l) \], where \( Cr \) indicates the protein concentration (M × residue number) and \( l \) specifies the cuvette path length (cm) (I).
Analysis of the TRPML1 I-II linker mutant proteins

The cell lysates of three TRPML1 I-II linker proteins containing MLIV-causing mutations L106P, T232P, or C166F in addition to the wild-type TRPML1 I-II linker were subjected to FSEC. Upon analysis of these samples, it was evident that not only were these pathogenic channels expressed at a relatively low yield, but they also exhibited a fair amount of protein aggregation (Figure 13A). Therefore, we used circular dichroism to examine secondary structural changes caused by these mutations. Circular dichroism (CD) is a technique that rapidly determines protein characteristics including the folding and secondary structure of proteins (24). It is widely used to ascertain whether a purified protein is folded and if a particular mutation affects the stability or conformation of the protein (24).

CD analysis of these three purified mutant linker proteins demonstrated significant changes in their secondary structures compared to that of wild-type TRPML1 (Figure 13B). Additionally, full-length wild-type TRPML1 in a blue native gel migrated as a band consistent with that of a tetramer, while the full-length pathogenic proteins appeared to have migrated as aggregated oligomers. Moreover, in HEK293T cells, these pathogenic mutants in the TRPML1\textsuperscript{VP} channel background produced little or no current. Consistent with the confocal imaging results, biochemical and structural analyses indicate that these mutations all appear to affect the structural assembly of the tetramer in isolation as well as in the full-length channel (I).
Figure 13. Biochemical and structural analysis of the I-II linker pathogenic mutations. (A) Representative FSEC profiles of the wild-type and mutant TRPML1 I-II linker proteins tagged with GFP. Peaks correspond to the fluorescence of aggregated protein, TRPML1 tagged with GFP, and free GFP, respectively. (B) Illustrative CD spectra of TRPML1 wild-type and mutant I–II linker proteins. \([\theta]_{\text{MRE}}\) denotes mean residue ellipticity. Published in (1).

2.6 Conclusion

In summary, our studies of the TRPML1 I-II linker provided us with a greater understanding of the structure and function of this domain within the full-length TRPML1 channel. The crystal structures that we determined of the TRPML1 I-II linker are the first reported high-resolution structures of a TRPML channel domain. The structure demonstrates that the I-II linker forms a tetramer with four-fold rotational symmetry. Interestingly, the I–II linker structures at pH 4.5, 6.0, and 7.5 seem to exist in a highly stable state since significant conformational changes were not detected, despite changes in pH conditions. Ultimately, determining the full-length structure of TRPML1 will provide valuable information and is an area to explore in future research.
To expand upon our structural findings, functional studies were performed, revealing that the central luminal pore plays a significant role in pH and Ca\(^{2+}\) regulation of TRPML1. Taken together, our studies indicate that the dual Ca\(^{2+}\)/pH regulation involves the intricate interrelationship between the Ca\(^{2+}\) blockade of the luminal pore and the protonation of aspartates. This salient finding can further be supported by future NMR studies that ascertain protonation states at various pH conditions.

Importantly, our TRPML1 I-II linker structures served as a basis from which to analyze and thereby characterize several MLIV-causing mutations that reside in this domain, namely L106P, C166F, and T232P. The localization patterns of these MLIV-causing mutations in the cell were investigated using confocal fluorescence microscopy, demonstrating that mutations in TRPML1 lead to the mislocalization of these proteins. In addition to our confocal imaging results, biochemical and CD studies support that the MLIV-causing mutations in the I-II linker can significantly influence the tetrameric assembly of the channel. In conclusion, our findings have revealed novel insights into the structure and function of TRPML1 as well as how disease-causing mutations in TRPML1 lead to MLIV.
2.7 References


Chapter III

Structural Studies of TRPP2

3.1 Introduction

As described in Chapter I, TRPP2 is a member of the TRP channel family of proteins and mutations in this channel lead to autosomal dominant polycystic kidney disease (ADPKD) (1, 2). Due to the clinical relevance of this protein, investigations pertaining to TRPP2 are becoming increasingly more prevalent. Initially, several structural studies on isolated cytoplasmic domains of TRPP2 were published, including the crystal structure of the coiled-coil domain, which was solved in our lab (3), as well as the NMR structures of the EF-hand domain (4–6). In order to gain insight into the structural effects of ADPKD mutations as well as the mechanisms underlying channel regulation, I set out to determine a high-resolution structure of near full-length TRPP2.

Prior to this endeavor, no published data on the crystal or cryo-EM structure of a full-length TRP channel were reported. Since then, several full-length TRP channel structures were solved using cryo-EM, due to recent technological advances in this field. As mentioned in Chapter I, major improvements in classification software as well as direct electron detectors have occurred, enabling structures of membrane proteins to be solved at atomic resolution. Prominently, this recent ‘resolution revolution’ (7) was pioneered by the elucidation of the TRPV1 channel (8–10). Subsequently, high-resolution cryo-EM structures of other TRP channels including TRPA1 (11), and TRPV2 (12, 13) were determined. Additionally, the high-resolution crystal structure of TRPV6 was solved (14).

Until quite recently, no structures have been solved containing the transmembrane region of TRPP2. Several months ago, Shen et al. reported the first cryo-EM structures of a closed state of
TRPP2 that included the transmembrane domains (15). Thereafter, Grieben et al. and Wilkes et al. published similar TRPP2 structures of closed and open states, respectively (16, 17). These structures will be analyzed in detail in Chapter IV. The aim of this chapter is to highlight the methods we implemented in order to successfully express and purify the human TRPP2 channel, as well as the cryo-EM analysis that we performed. An additional topic of discussion will address the comparison of methods utilized in this study as well as those applied by other groups that solved TRPP2 structures.

3.2 Screening of TRPP2 constructs using FSEC

3.2.1 Rationale for cloning truncated constructs

Based on my preliminary studies, TRPP2 seemed to be a promising protein for structural determination. Since working on a full-length membrane protein was a new endeavor, it was important to create and analyze a wide range of constructs to assess which had suitable yield and purity for structural studies. Initially, my goal was to identify constructs that could be used for crystallographic analysis. Many variables appear to affect the expression level, homogeneity, and ability to crystallize membrane proteins. Some of these variables include truncations, detergent types (18), lipids (19), homologues used (20), types of affinity tags, as well as the location of tags in the construct (21, 22). However, after the emergence of the first high-resolution cryo-EM structure of TRPV1 (8, 9), we applied cryo-EM methodology in order to evaluate TRPP2 protein samples for structural determination.

Before generating truncations within TRPP2, a secondary structure prediction search on human TRPP2 using PSIPRED (23) and Jpred (24) was performed in order to get a general sense of where alpha-helices and beta-sheets were located in the structure. Subsequently, numerous truncations
on the N- and C-termini of TRPP2 were created, removing flexible and disordered regions of these proteins, and cloned into a pEGFP vector (Clontech). Consequently, each construct contained an EGFP tag on the N-terminus of the TRPP2 protein. Enhanced green fluorescence protein (EGFP) tags were utilized in order to detect the protein expression levels, homogeneity, and aggregation of particular constructs through a method known as fluorescence-detection size-exclusion chromatography (FSEC) (22). Notably, the S1-S6 transmembrane segments of the protein were included in all of the constructs. According to studies performed by Dr. Yong Yu, a previous member of our lab, the coiled-coil domain is thought to play a key role in the assembly of homomeric TRPP2 (3). Therefore, all of the constructs that I created contained the coiled-coil domain at the C-terminus.

3.2.2 Transient HEK cell transfection and FSEC

TRPP2 constructs were then transiently transfected into human embryonic kidney 293 (HEK 293T) cells (ATCC) using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Of note, HEK293T cells were used instead of the glycosylation-deficient GnTI−/− HEK cells (HEK293S) used by other groups for FSEC analysis (22). Since the expressed protein had an EGFP tag on its N-terminus, the transfection efficiency was able to be examined under the fluorescent microscope. Forty-eight hours after transfection, cells were pelleted and resuspended in a solution containing 50 mM NaHPO₄, 300 mM NaCl, 2.5% (w/w) glycerol pH 8.0, and Halt protease-inhibitor cocktail (Thermo Fisher Scientific). The nonionic detergent n-Dodecyl-β-D-maltoside (DDM; Anatrace) was added at a final concentration of 10% and lysates were incubated at 4°C for 1 hour. The solution was then clarified by centrifugation at 16,000 r.p.m. for 10 min. FSEC was subsequently performed by injecting the supernatant onto a Superose 6 10/300 column.
(GE Healthcare), for size-exclusion chromatography (SEC), connected to a spectrofluorometric detector RF-10AXL (Shimadzu) which measured the fluorescence. The gel filtration buffer contained .05% DDM, 50 mM NaHPO₄, 300 mM NaCl, and 2.5% glycerol pH 8.0.

3.2.3 FSEC analysis of human TRPP2 constructs transfected in HEK293 cells

FSEC is a valuable screening technique that can assess the monodispersity and expression level of a protein tagged with green fluorescent protein (GFP) by analyzing the elution trace of GFP fluorescence (20). By subjecting proteins fused to GFP to SEC methods, FSEC can assess in a simple manner the stability, approximate molecular size, as well as the yield of membrane proteins, without the need for purification (21, 22). When analyzing FSEC traces, it is important to examine the extent to which proteins are present in the void volume peak or the included volume peak. Proteins that migrate in the void volume are likely misfolded and insoluble, while proteins in the included volume peak are presumed to be correctly folded (20, 22).

In addition, the shape of the included peak is a way to gauge the quality of the protein. A monodispersed sharp included peak is highly preferable for crystallization and other structural studies over a polydispersed included peak that contains multiple oligomeric or conformational states (20–22). Through this method, I was able to detect which constructs produced a greater monodispersed included peak with low levels of aggregation that could be used further for structural analysis. The FSEC traces for the most promising human TRPP2 (hTRPP2) constructs expressed in HEK293T cells, hTRPP2:181-927 and hTRPP2:191-927, are illustrated in Figure 1. These FSEC profiles represent a subset of the total number of constructs analyzed.
Figure 1. FSEC traces of human TRPP2 constructs after transient transfection in HEK cells. 
(A-B) FSEC traces of human TRPP2:181-927 (A) and TRPP2:199-927 (B). The void, TRPP2-GFP, and ‘free’ GFP peaks are illustrated.

Upon FSEC analysis, the traces of both hTRPP2:181-927 and hTRPP2:191-927 were very similar. Notably, they exhibited relatively monodispersed peaks with regard to their included volume peaks containing GFP-TRPP2, characteristic of well-folded proteins. Therefore, these truncations appeared to be relatively soluble in DDM because only a small amount of the proteins aggregated and migrated in the void peak. Since these findings were very promising, hereafter, I primarily focused on further investigating these two constructs for structural analysis.

HEK cells provide near native translocation machinery, post-translational modifications, and a lipid environment that can be very important for the production of human membrane proteins (22). Although it appeared highly favorable to express human TRPP2 in its native environment of human kidney cells, due to limited resources we could not pursue large-scale purification from HEK cells since we did not have the proper equipment to carry out this procedure. However, our
lab was fully equipped with the necessary materials needed to scale-up purification of TRRP2 from insect cells. Hence, I started to utilize the baculovirus-based expression system in insect cells to test the expression levels of promising constructs including TRPP2:181-927 and TRPP2:199-927.

3.3 Insect cell baculovirus-based expression system

The insect cell baculovirus-based expression system has been previously utilized to express mammalian membrane proteins since it uses similar protein post-translational modifications and sorting machinery to that found in mammalian cells (22). An advantage of using the baculovirus system is that the virus can easily be propagated in insect cells to high titers (21, 25). Additionally, this system often creates properly folded membrane proteins (21, 25), an essential characteristic for structural studies.

3.3.1 Generation of recombinant baculoviruses

The Bac-to-Bac® Baculovirus Expression System (Invitrogen) protocol was implemented according to manufacturer’s instructions in order to efficiently create recombinant baculoviruses for TRPP2 structural studies. The human TRPP2 gene containing various N- and C-terminal truncations initially linked to GFP were subsequently cloned into a pFastBac™ vector. After confirming the correct truncations through sequencing, the pFastBac™ constructs were transformed into DH10Bac™ competent E. coli cells (Invitrogen) to produce recombinant bacmids. Since DH10Bac™ cells have a baculovirus shuttle vector (bacmid), upon transformation transposition occurs between the pFastBac™ vector and the bacmid resulting in recombinant bacmids. After transformation, the DH10Bac™ cells were transferred to LB agar plates containing 50 μg/ml kanamycin, 7 μg/ml gentamicin, 10 μg/ml tetracycline, 100 μg/ml Bluo-gal, and 40 μg/ml
IPTG. Blue/white selection was carried out and white colonies were chosen, which contained the recombinant bacmid in a background of blue colonies that carry the non-recombinant bacmid.

Subsequently, I purified the recombinant bacmid DNA by performing minipreps. The isolated bacmid DNA was then transfected into *Spodoptera frugiperda* (Sf9) insect cells (Invitrogen) in order to create recombinant baculovirus stocks at 27°C. Several days after transfection, I checked the fluorescence of the insect cells using a fluorescent microscope to assess the transfection efficiency, whereby most of the cells were infected by the virus. Subsequently, I collected the media in each well from the transfection plate, which contained the P1 virus. I then amplified this virus in Sf9 cells and collected the virus generated, known as the P2 virus. The P2 virus contained amplified baculoviral stock that was later utilized to infect Hi5 (*Trichoplusia ni*) cells (Expression Systems) for small and large-scale expression of TRPP2.

### 3.3.2 FSEC analysis of human TRPP2 constructs transfected into Sf9 cells

For FSEC analysis, I then harvested 500μl of infected insect cells by centrifugation. FSEC analysis was performed using the same procedure described in section 3.2.2 for HEK293 cells. FSEC was performed on a number of samples, and the two most promising constructs yet again were human TRPP2:181-927 and TRPP2:199-927. The FSEC profiles of these constructs resulted in generally monodispered main protein peaks, suggestive of homogenously folded proteins (Figure 2). These results were in agreement with the FSEC profiles of these constructs from HEK
cells. Subsequently, I continued to perform small-scale and large-scale purification of these truncated proteins, which will be described in detail in later sections.

**Figure 2.** FSEC traces of human TRPP2 constructs after transient transfection in Sf9 cells. (A-B) FSEC traces of hTRPP2:181-927 (A) and hTRPP2:199-927 (B) transiently expressed in Sf9 cells and solubilized in DDM. The TRPP2-GFP and ‘free’ GFP peaks are highlighted.

### 3.4 FSEC of mouse TRPP2 constructs transfected in HEK293 cells

Screening through additional homologues of TRPP2 is a robust approach to determine which proteins are the most stable in detergents and have increased expression levels (26). Therefore, we made various mouse TRPP2 (mTRPP2) constructs with truncations at the N- and C-termini of the protein. Both the six transmembrane segments as well as the coiled-coil domain in mTRPP2 were present in the constructs. Dr. Ravi Kalathur, a scientist at the New York Structural Biology Center (NYSBC), helped to expedite the cloning and screening of these constructs. The mTRPP2 truncations were cloned into pTriEx 1.1 vectors modified for ligation independent cloning [LIC] containing N-terminal tags [His10-Flag-GFP-TEV-mTRPP2] or C-terminal tags [mTRPP2-TEV-
GFP-Flag-His10]. The protocols for the solubilization of whole cell lysates in DDM and subsequent FSEC analysis were the same as those described for the human TRPP2 constructs. Representative FSEC profiles of a subset of samples are indicated in Figure 3.

Figure 3. FSEC traces of mouse TRPP2 constructs after transient transfection in HEK cells. (A-D) FSEC traces of mTRPP2:25-916 (A), mTRPP2:35-923 (B), mTRPP2:18-938 (C), and mTRPP2:68-938 (D). Of note, (A, C) had GFP linked to the C-terminus of the constructs, while (B, D) contained GFP at the N-terminus.
Notably, all of the constructs exhibited quite high levels of aggregation in the void volume peaks, indicative of insoluble and misfolded proteins. In addition, despite making various truncations and placing the GFP tag on the N- and C-termini, the FSEC traces overall looked very similar in all four samples. Therefore, based on these preliminary studies it appeared that the truncations and the positions of the tags did not play a major role in the overall expression level and homogeneity of the samples. Thus, since the FSEC profiles from the transfected mouse TRPP2 constructs did not look promising, I primarily focused my efforts on optimizing purification conditions for the human TRPP2 constructs.

3.5 Purification of hTRPP2 in Hi5 cells using DDM and amphipol

Since both hTRPP2:199-927 and hTRPP2:181-927 exhibited the most promising FSEC profiles in both HEK cells and insect cells, I then proceeded to express and purify these constructs at a small-scale from insect cells. It has been shown that Hi5 cell lines demonstrate higher (tenfold) susceptibility to baculovirus infection as well as higher protein yield compared to Sf9 cells (27). Additionally, since it has been reported that Hi5 cells are more favorable for the expression of glycosylated proteins (28, 29), hereafter, Hi5 cells were utilized in all purifications described. However, amplification of the P1 virus was carried out in Sf9 cells.

Generation of TRPP2 constructs with C-terminal MBP

The human TRPP2 gene containing the truncations 199-927 and 181-927 was cloned into a modified pFastBac vector with SbfI and Ascl restriction sites. The resulting construct contained hTRPP2 followed by a tobacco etch virus (TEV) protease recognition sequence site, a maltose binding protein (MBP) tag, and an octa-histidine tag at the C-terminus. The purified proteins that
were further investigated were derived from the expression of this vector. The TRPP2 baculoviral stock was generated according to the manufacturer’s instructions (Invitrogen, Bac-to-Bac), as previously mentioned.

3.5.1 Small-scale purification in DDM

A 200 ml culture of Hi5 insect cells was infected with the hTRPP2:199-927 virus when the cell density reached 2×10^6 cells/mL at 27 °C. Forty-eight hours after infection, cells were harvested by centrifugation at 4°C. Cell pellets were suspended in a buffer containing 250 mM NaCl, 10% glycerol, and 20 mM Hepes pH 8.0 (buffer A) in the presence of a protease inhibitor cocktail (Roche) and homogenized with a glass dounce homogenizer. The cells were then disrupted by sonication, and were subjected to ultracentrifugation at 45,000 r.p.m for 1 hr at 4°C. The resulting pelleted cell membrane was resuspended in buffer A containing the Roche protease inhibitor cocktail and homogenized once more with a glass dounce homogenizer. The TRPP2 protein was then extracted with 1% DDM for 1 hr at 4°C.

Subsequently, the solubilized membrane was clarified from cell debris by ultracentrifugation at 45,000 r.p.m for 30 min at 4°C. The supernatant was then incubated with amylose resin (NEB) with gentle agitation for 1 hr at 4°C. The resin was isolated by centrifugation at 3,000 r.p.m and transferred to a gravity column. The resin was subsequently washed with 15 column volumes of buffer A containing .05% DDM. TRPP2 was eluted with the same buffer containing .05% DDM and 20 mM maltose. The protein was concentrated by ultrafiltration utilizing the Amicon Centrifugal Filter Units (EMD Millipore) and concentrations were measured using a NanoDrop (Thermo-Scientific). TRPP2 was further purified by size-exclusion chromatography (SEC) with a Superose 10/300 GL 6 column equilibrated with SEC buffer containing .05 % DDM, 150 mM
NaCl, and 20 mM Hepes pH 7.4. The peak fractions containing TRPP2 were then collected and examined on an SDS-PAGE gel.

3.5.2 Reconstitution into amphipols

Apart from detergents, amphipathic polymers (amphipols) are an alternative approach to stabilize the transmembrane domains of membrane proteins after solubilization and purification with detergent (30). These amphipols are able to mimic the lipid environment of the membrane and firmly attach to the transmembrane domains of membrane proteins (31, 32). In fact, amphipols have been extensively used in order to determine several other TRP channel cryo-EM structures including TRPV1 (8, 9), TRPV2 (13), and TRPA1 (11).

The methods described previously were utilized to purify the TRPP2 protein. Upon elution of TRPP2 in .05% DDM, TEV protease (40 μg per 1 mg TRPP2) was added in order to cleave the MBP tag. Simultaneously, the protein was then mixed with amphipol A8-35 (Anatrace) at a 1:3 (w/w) ratio and incubated overnight at 4°C with gentle agitation. Afterwards, the sample was incubated with Bio-Beads SM-2 (Bio-Rad) (15 mg per 1 ml channel/detergent/amphipol mixture) for 8 hr in order to remove detergent, according to published methods (9, 11) with modifications. Bio-Beads were subsequently removed over a disposable polyprep column. The protein was then concentrated and purified by SEC in a buffer containing 150 mM NaCl and 20 mM Hepes pH 7.4.

Size-exclusion chromatography (SEC), like FSEC, is a highly effective method for assessing the monodispersity and stability of a protein. The presence of a single symmetrical Gaussian peak is indicative of a properly folded protein that is monodisperse, while multiple asymmetric peaks likely suggest an unfolded and unstable protein that is polydisperse (20, 22). As shown in Figures 4A and 4C, the SEC profiles of the hTRPP2:181-927 and hTRPP2:199-927 eluted in .05% DDM.
are similar with relatively monodispersed TRPP2 peaks, indicating that the proteins are correctly folded and homogenous. However, since the proteins were expressed at a small-scale, the SEC profiles demonstrated relatively poor expression. Of note, the critical micelle concentration (CMC) of DDM is 0.0087% and the concentration used in the gel filtration buffer was significantly higher than the CMC. The critical micelle concentration (CMC) of a detergent is the threshold concentration above which micelles are produced.

Interestingly, upon exchange of DDM with amphipol A8-35 the protein peaks were no longer detected (Figures 4B, D). The protein appeared to have aggregated and visible precipitation was present after centrifugation prior to loading onto the gel filtration column. At first, since TEV and amphipol A8-35 were added at the same time, it was difficult to determine what caused the aggregation. Consistent with this observation, it has been reported that in some cases upon amphipol reconstitution, trace amounts of aggregates were identified (32, 33). Presumably, the Bio-Beads were absorbing DDM too quickly, without allowing amphipol A8-35 to properly stabilize the protein, thus leading to protein aggregates. Of significance, it has been demonstrated that adding Bio-Beads slowly in steps to the protein-detergent mixture can aid in the reduction of aggregated protein (34). Therefore, in subsequent protein preps I added Bio-Beads in a stepwise manner to remove DDM.

Additionally, I also altered the amphipol reconstitution methods by initially adding TEV protease followed by mixing the protein with amphipol A8-35 at a higher ratio, 1:6 or 1:8 (w/w), rather than 1:3. Presumably, a higher ratio of amphipol present would aid in the stabilization of the protein. As a result, a small protein peak corresponding to TRPP2:181-927 was visible (Figure 5C). However, this peak was significantly smaller than the corresponding TRPP2 peak in the DDM sample (Figure 5A). Importantly, when I began performing large-scale preps with another
detergent, Lauryl Maltose Neopentyl Glycol (LMNG), and subsequently reconstituted the protein in amphipol, I began to obtain a higher protein yield with less aggregation (Figure 8). Therefore, by adjusting parameters in the amphipol reconstitution protocol, I was able to solve the amphipol dilemma that I encountered.
Figure 4. SEC traces from small-scale purification of TRPP2 in DDM and amphipol A8-35. (A) SEC trace of human TRPP2:181-927 in a buffer containing 0.05% DDM, 150 mM NaCl, and 20 mM Hepes pH 7.4. The peaks correspond to TRPP2 and MBP, respectively. (B) SEC trace of human TRPP2:181-927 after amphipol reconstitution in a buffer containing 150 mM NaCl and 20 mM Hepes pH 7.4. The TRPP2 peak was not present. (C) SEC trace of human TRPP2:199-927 purified in the same buffer used in (A). The peaks depict TRPP2 and MBP. (D) SEC trace of human TRPP2:199-927 after amphipol reconstitution in the same buffer used in (B). Notably, the TRPP2 peak was not exhibited. (E) SDS-PAGE of TRPP2 samples stained with Coomassie Blue to visualize the protein bands. The lanes are as indicated: 1, 181-927 eluted from amylose resin; 2, 181-927 after TEV cutting; 3, 181-927 gel filtration profile peak (A); 4, 199-927 before TEV cutting; 5, 199-927 gel filtration profile peak (C); 6, 181-927 eluted from amylose resin; 7, 181-927 after TEV cutting and amphipol reconstitution; 8, 199-927 before TEV cutting and amphipol reconstitution; 9, 199-927 after TEV cutting and amphipol reconstitution; 10, 199-927 profile peak after amphipol reconstitution (D). All of the proteins indicated were purified in parallel.
3.5.3 Inhibition of N-glycosylation

Post-translational modifications such as glycosylation play a role in the proper folding and overall stability of membrane proteins. Although modifying these predicted glycosylation sites can lead to more homogeneous samples, it could also be deleterious to the expression of the protein since some proteins need native glycosylation (35). Initially, I was purifying TRPP2 with the goal of obtaining its crystal structure. Since TRPP2 contains five glycosylation sites, (Asn 299, Asn305, Asn328, Asn362, and Asn375) (36), the glycosylated protein would presumably lead to heterogeneity in the protein sample if used for crystallization. Therefore, I added an antibiotic tunicamycin to the insect cells upon viral infection.

In theory, tunicamycin should inhibit the addition of glucosamine into N-glycosidically linked glycoproteins (37). As indicated in the SDS-PAGE gel in Figure 5, it appears that the TRPP2 protein was not glycosylated after tunicamycin was introduced, since the band in lane 4 migrated at a slightly lower molecular weight than the glycosylated samples (lanes 1-3) (Figure 5D). In addition, upon cutting with TEV the TRPP2 band in lane 6 also had a lower molecular weight compared to the glycosylated protein in lane 7, indicating that indeed glycosylation was inhibited (Figure 5D). However, the protein peak of TRPP2 with tunicamycin seemed broader (Figure 5B) than that seen in the DDM sample without the antibiotic (Figure 5A).
Figure 5. SEC traces from small-scale purification of TRPP2:181-927. (A) SEC trace of human TRPP2:181-927 in a buffer containing .05 % DDM, 150 mM NaCl, and 20 mM Hepes pH 7.4. Peaks corresponding to TRPP2 and MBP+TEV are denoted. (B) SEC trace of human TRPP2:181-927 with the addition of tunicamycin purified in the same gel filtration buffer utilized in (A). (C) SEC trace of human TRPP2:181-927 after amphipol A8-35 reconstitution in a buffer containing 150mM NaCl and 20 mM Hepes pH 7.4. (D) SDS-PAGE of TRPP2 samples stained with Coomassie Blue. The lanes are as indicated: 1-3, 181-927 elute from amylose resin (.05% DDM); 4, 181-927 elute from amylose resin (tunicamycin sample); 5, 181-927 after TEV cutting (.05% DDM sample); 6, 181-927 after TEV cutting (tunicamycin sample); 7, 181-927 after TEV cutting (sample reconstituted in amphipol). All of the proteins indicated were initially solubilized in DDM and subsequently purified in parallel.
Since we later decided to use cryo-EM instead of X-ray crystallography to determine the structure of TRPP2, we assumed that glycosylation would not hinder structural determination using cryo-EM. As indicated in Figure 6, although protein engineering is often required for protein crystallization in order to remove glycosylation and flexible sites, glycosylated membrane proteins can be determined by cryo-EM and presumably require less protein engineering (38). Therefore, we did not see a need to use tunicamycin extensively thereafter, and instead purified glycosylated TRPP2.

Figure 6. Role of glycosylation in membrane protein structural determination. Protein engineering is often required for successful protein crystallization in order to eliminate glycosylation and flexible sites. Alternatively, glycosylated membrane proteins can be determined by cryo-EM with less need for modifying regions that will likely cause heterogeneity in the samples. Figure adapted from (38).
3.5.4 Large-scale purification of TRPP2 in DDM

Since recent advancements occurred in the field of single-particle cryo-EM, we also wanted to utilize this structural approach to obtain a high-resolution structure of TRPP2. Thus, we initiated a collaboration with the laboratory of Dr. Reza Khayat (City College of New York) and began testing TRPP2 samples that had relatively monodispersed peaks. Dr. Devendra Srivastava, a postdoctoral fellow in the Khayat lab, prepared the negative-stain and EM grids and Dr. Khayat along with Dr. Srivastava both tested the samples at the Simons Electron Microscopy Center at NYSBC.

I subsequently started performing large-scale preps with hTRPP2:199-927 to obtain sufficient protein yield for cryo-EM analysis. In order to achieve a higher yield, I infected 3-4 L of Hi5 insect cells with virus followed by incubation for 72 hr. The purification procedure was the same as previously denoted, with slight modifications. Notably, the MBP tag was not cleaved and 2 mM TCEP (tris (2-carboxyethyl) phosphine) was added to the buffer at every step of the purification procedure. An SDS-PAGE gel was used to evaluate the purity of the protein sample and TRPP2-MBP appeared to be pure (Figure 7C).

The protein yield was significantly higher (Figure 7A) than previously observed, yet the sample was slightly heterogeneous. Nonetheless, only fractions within the main peak were collected. A few microliters of the TRPP2 protein at 0.84 mg/ml in 0.05% DDM were applied to EM protochip C-flat real gold grids and plunge-frozen in liquid ethane cooled by liquid-nitrogen to immobilize the protein particles present within the thin layer of vitreous ice. However, detergent micelles were apparent and the protein was not easily detectable (Figure 7B).
Figure 7. TRPP2-MBP and TRPV2 purified in DDM. (A) SEC trace of human TRPP2:199-927-MBP in a buffer containing .05 % DDM, 150 mM NaCl, 20mM Hepes, and 2 mM TCEP pH 7.4. (B) Representative electron micrograph of TRPP2 in .05% DDM embedded in vitreous ice. (C) SDS-PAGE of TRPP2 samples. The lanes are as indicated: 1, TRPP2 eluted in .05% DDM; 2, SEC fraction E10; 3, SEC fraction E11; 4, SEC fraction E12; 5, SEC fraction F12; 6, SEC fraction F11; 7, SEC fraction F10; 8, SEC fraction F9. (D-E) SEC profiles of TRPP2:199-927-MBP purified in .05% DDM two days (D) and nine days (E) at 4°C after elution from amylose resin. (F) SEC trace of TRPV2 solubilized and purified in DDM (30). (G) Negative-stain EM micrograph of TRPV2 in DDM (30). Of note, (F-G) are adapted from (30).

Similar EM grids were observed using protein from subsequent preps. Aliquots of TRPP2 were then tested two and nine days after elution from amylose resin to assess the long-term stability of the protein at 4°C (Figures 7D, E). Interestingly, although hTRPP2:199-927 samples appeared to be stable at 4°C for more than one week and had relatively monodispersed peaks, the samples did not look promising when tested on negative-stain or EM grids. However, my findings suggest that small-scale preps yield more monodispersed protein peaks than do larger preps. At smaller yields, the main TRPP2 protein peak appeared to be less polydispersed and therefore more homogeneous.

Perhaps the intrinsic nature of DDM led to polydispersity in the samples. Notably, as illustrated in Figures 7F and 7G, the solubilization and purification of TRPV2 in DDM led to heterogeneity in the sample as indicated by the SEC profile and negative-stained EM micrograph (30). Interestingly, both the TRPP2 and TRPV2 SEC profiles appear very similar, as do their respective cryo and negative-stained micrographs. Moreover, the gel filtration profiles of these proteins exhibited shoulders to the left of the main protein peaks. Taken together, these results further confirmed that DDM does not sustain a stable environment for TRPV2 (30) or TRPP2.
3.6 Purification of hTRPP2 in Hi5 cells using LMNG

I then proceeded to test how well TRPP2 behaves in the maltose-neopentyl glycol (MNG) class of detergents, which was used for the structural determination of TRPV2 by cryo-EM (12). This class of detergents is relatively new and has been reported to aid in the stabilization of membrane proteins for structural studies (30). MNG detergents have been reported to have a milder effect when stabilizing membrane proteins compared to DDM (39). Depending on the size of the prep, .5-4 liters of Hi5 insect cells were infected with the hTRPP2:199-927 virus when the cell density reached $2 \times 10^6$ cells/mL at 27°C. Seventy-two hours after infection, cells were harvested by centrifugation at 4°C. Cell pellets were suspended in a buffer containing 200 mM NaCl, 10% glycerol, and 50 mM Hepes pH 8.0 (buffer A) in the presence of the SIGMAFAST protease inhibitor cocktail (Sigma-Aldrich) and homogenized with a glass dounce homogenizer. The cells were then sonicated and clarified by ultracentrifugation at 45,000 r.p.m for 1 hr at 4°C. The pelleted cell membrane was resuspended in buffer A containing the SIGMAFAST protease inhibitor cocktail and further homogenized. The TRPP2 protein was then extracted with 1% Lauryl Maltose Neopentyl Glycol (LMNG, Anatrace) for 1 hr at 4°C. Subsequently, the solubilized membrane was subjected to ultracentrifugation at 45,000 r.p.m for 30 min at 4°C. The resulting supernatant was then incubated with amylose resin for 1 hr at 4°C. The resin was collected by centrifugation at 3,000 r.p.m, transferred to a gravity column, and washed with 15 column volumes of buffer A containing .003% LMNG. TRPP2 was eluted with the same buffer containing .003% LMNG and 40 mM maltose. The protein was then concentrated and further purified by SEC with a Superose 6 column equilibrated with SEC buffer containing .003% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4.
3.6.1 Reconstitution of TRPP2 into amphipols

Briefly, the eluted hTRPP2 protein was cleaved with TEV protease for 8 hr and the protein was then mixed with amphipol A8-35 at a 1:6 (w/w) ratio. After rotating the sample with amphipol for 6 hr, I added Bio-Beads sequentially every 1.5 hr using 15 mg per 1 ml channel/detergent/amphipol mixture. After 10 hr, all the Bio-Beads were added and the sample rotated overnight at 4°C. The protein was then concentrated and subjected to SEC, utilizing a gel filtration buffer containing 150 mM NaCl and 20 mM Hepes pH 7.4. The protein appeared to have existed in multiple conformations and/or oligomeric states, resulting in a polydisperse protein peak (Figure 8A). As seen in the preceding DDM SEC profiles, a prominent shoulder was present to the immediate left of the TRPP2 main SEC peak after reconstitution into amphipol (Figure 8A). According to an SDS-PAGE gel, the main peak fractions also contained TEV protease.

In a subsequent prep, upon cleavage with TEV I tried to separate TEV and MBP from TRPP2 using cobalt beads. Theoretically, since both TEV and MBP are his-tagged they should bind to the cobalt beads after cleavage, and TRPP2 should not. However, TRPP2 also bound to the cobalt beads as well and was only eluted from the resin when imidazole was added. Therefore, it appeared that a strong interaction with TRPP2 and MBP was present, although they were able to be separated by gel filtration. Additionally, since a linker region containing the TEV protease-cleavage site was created right after the TRPP2 sequence in the expression vector, it is plausible that TEV protease was still able to bind quite tightly to this linker region of TRPP2, thus being present in the protein peak and subsequent gel.

Thereafter, since TEV could not be easily separated from TRPP2, MBP was not cleaved in the majority of subsequent preps. In Figure 8B, TRPP2-MBP was purified in a gel filtration buffer containing .0008% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4. Of note, the CMC of LMNG
is 0.001% and the protein was stable after purification due to the presence of the relatively monodispersed peak. Since negative-staining of TRPP2 samples in 0.003% LMNG (3x above the CMC) generated bad ice, we subsequently tested how the protein behaved below the CMC of LMNG.

When TRPP2:199-927-MBP was reconstituted in amphipol A8-35 (Figure 8C), polydispersity in the main SEC peak containing TRPP2 was less evident than seen in the cleaved protein sample (Figure 8A), although the peak was still quite broad. Evidently, the presence of polydispersity in the TRPP2 main peak varied from one prep to another. However, despite the improvements in the quality of the protein, when the amphipol A8-35 samples were applied to negative-stain and EM grids the protein complex appeared to dissociate, exhibiting broken particles and high levels of aggregation. Similar results were seen when TRPP2 was solubilized in DDM and transferred to amphipol A8-35.
Figure 8. SEC traces of TRPP2 purified in LMNG and amphipol A8-35. (A) SEC trace of human TRPP2:199-927 after amphipol reconstitution in a buffer containing 150 mM NaCl and 20 mM Hepes pH 7.4. TEV was also present in the main TRPP2 peak. (B) TRPP2:199-927-MBP purified in .0008% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4. (C) SEC trace of human TRPP2:199-927-MBP after amphipol reconstitution purified in the same gel filtration buffer used in (A). MBP was not cleaved from TRPP2. From the left, the peaks correspond to tetrameric TRPP2 followed by excess amphipol.
3.7 Titan Krios beta-testing phase at NYSBC

To produce suitable samples for cryo-EM analysis on the newly installed Titan Krios at NYSBC, LMNG was utilized during extraction and purification of TRPP2. The purification procedure for the TRPP2:199-927 protein was primarily the same as previously denoted, with slight modifications. The gel filtration buffer contained .0008% LMNG, 150 mM NaCl, 20 mM Hepes pH 7.4, and 2 mM TCEP. The protein peak fractions F11 and F10 were collected, avoiding the polydispersed region from E11 and E12 (Figure 9A). The protein eluted from the column at .6mg/ml in F10. Based on previous negative-stain experiments, the protein in the F10 fraction appeared to be more homogenous than in the F11 fraction.

The protein was then brought to our collaborators in the Khayat lab. Dr. Srivastava made negative-stain and cryo-EM grids to utilize during the beta-testing phase on the Titan Krios. Briefly, several microliters of the human TRPP2 protein (.45mg/ml) were loaded onto Protochip c-flat gold grids 1.2/1.3 and blotted for 3.0 s using a blot force of 1 at 100% relative humidity and 4°C using Vitrobot Mark IV. The grids were then immediately plunged into liquid ethane cooled by liquid nitrogen. Micrographs were acquired by a Titan Krios microscope equipped with an FEI Falcon 2 direct electron detector. Dr. Khayat along with Dr. Srivastava both collected cryo-EM data on the microscope. Figure 9B illustrates a representative micrograph from the Titan Krios session.
Figure 9. Representation of TRPP2 data collected on the Titan Krios. (A) Gel filtration profile of purified TRPP2 in .0008% LMNG. (B) Representative EM micrograph of human TRPP2:199-927. Particles were selected by semi-automatic particle picking and subjected to 2D screening. (C) 2D class averages of TRPP2 cryo-EM particles. (D) Enlarged 2D class average from (C) with the most promising features.
Additionally, we also set up EM grids with the presumable TRPP2 agonist triptolide, mentioned in Chapter I. TRPP2 protein was incubated overnight with 10 µM of triptolide and EM grids were subsequently prepared and tested on the Titan Krios. It appeared that the particles with and without triptolide (Figures 10A, B) looked very similar, generating 2D class averages comparable to those seen in Figure 9.

Figure 10. Cryo-EM particles of TRPP2 collected on the Titan Krios. (A-B) Illustrative micrographs of human TRPP2:199-927 with the presence (A) and absence (B) of triptolide.

3.7.1 Analysis of TRPP2 particles

Upon analysis of the micrographs, it appeared that the particles exhibited a strongly preferred orientation, highlighting the top view of TM domain cross-sections (Figures 9C, D). 2D class averages are indicated in Figure 9C, while Figure 9D highlights the 2D class average with the most promising attributes. For the first time, 2D class averages revealed a distinct tetrameric structure exhibiting four-fold symmetry (Figure 9D). Moreover, subtle structural features within the tetramer were able to be visualized.
Routinely, 2D projections can subsequently be combined to generate a 3D reconstruction, only if enough particles in various orientations are available (38). Although the data we collected resulted in reasonable 2D class averages, 3D class averages and reconstructions were not able to be produced due to the lack of additional orientations. One factor that may have led to preferred particle orientation is that adsorption of TRPP2 particles to the air-water interface could have occurred (40). Within the time between blotting and freezing of the sample, Brownian motion could result in the colliding of particles at the air-water interface (40, 41). Since some particles tend to attach to the air-water interface in a preferred orientation, this phenomenon may have taken place in our samples.

The ice thickness may have also contributed to the preferred orientation of TRPP2. Although thinner ice is conducive to a better signal-to-noise (SNR) ratio, (42) it may also promote preferred orientations. When ice is thin on the EM grid, particles may become compressed between air-water interfaces when the layer of water encompassing the particle is too thin (40). Thicker ice, on the other hand, would affect the data collection by leading to higher background, limiting the ability to reach high-resolution. However, thicker ice could provide a greater range of orientations yet lower SNR due to increased ice contrast (42). Additionally, we had trouble seeing particles on EM grids that had thicker ice.

Furthermore, we primarily used protein in a detergent that was at or below the CMC, not only for EM grids used during the Krios session, but in general. Our rationale for staying below the CMC of the detergents was to prevent excess detergent particles from interfering with data collection. Moreover, the particles appeared to be of better quality when the detergent used was below the CMC, and the protein seemed stable after being run through the gel filtration column. However, staying below the CMC may have also played a role in how the air/water interface
affected the particles. For instance, the CMC of LMNG is .001% and the EM grids that we tested during the Krios session were below the CMC, at 0.0008 %.

Although it was a wonderful opportunity to test our samples on the Titan Krios at NYSBC, unfortunately the microscope was still in its initial beta-testing phase when we were in need of it most. Soon after, we initiated another collaboration with the laboratory of Dr. Xueming Li (Tsinghua University) in order to expedite the process of optimizing grid conditions to overcome the preferred orientation dilemma. Graduate students Heng Zhou and Xiaoyuan Zhou prepared the negative-stain and EM grids as well as collected and analyzed the data. Additionally, Dr. Minghui Li then became involved in the project for several months and aided with protein purification.

3.8 Structural studies of the TRPP2 F604P mutant

As described in Chapter I, Arif et al. reported that F604P, a point mutation in the S5 transmembrane domain of TRPP2, resulted in a constitutively open gain-of-function (GOF) channel when expressed in Xenopus oocytes (43). This finding was particularly of interest to us since we wanted to solve the structure of TRPP2 in an open state. Therefore, if indeed the protein was locked in an open state it could aid in the reduction of conformational heterogeneity, possibly enabling us to overcome the preferred particle orientation dilemma we were facing.

Subsequently, I introduced F604P by mutagenesis into the 199-927 truncation present within the pFastBac™ vector. This construct also contained hTRPP2 followed by a TEV protease recognition sequence site, an MBP tag, and an octa-histidine tag at the C-terminus. Additionally, the TRPP2 F604P protein was expressed and purified using the aforementioned protocol, with slight modifications. Initially, the gel filtration buffer used to purify the TRPP2 F604P-MBP protein included .005% LMNG, 150 mM NaCl, 20 mM Hepes pH 7.4, and 2 mM TCEP.
Figure 11A displays a similar SEC profile to that of the wild-type TRPP2-MBP protein (Figure 9). This mutant protein also exhibited a slight shoulder adjacent to the main TRPP2 peak, indicating the possible presence of multiple oligomeric protein states. However, upon addition of soybean lipids (Avanti polar lipids), the level of polydispersity decreased (Figure 11B). Perhaps the soybean lipid aided in stabilizing the protein, yet the SEC profile was still quite broad. Soybean lipid was added to the elution buffer at a 1:5 ratio (soybean lipid to detergent) when the protein was eluted from the amylose resin. The gel filtration buffer contained .001% soybean lipid, .005% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4.

In order to test the stability of TRPP2 in the presence of soybean lipids, two 500µl aliquots of protein were left at 4°C for 5 days after elution from the amylose resin. One of the samples was then run on the gel filtration column (Figure 11C) and the other was heated to 37°C to test its thermostability before being loaded onto the column (Figure 11D). The gel filtration buffer used for both these runs included the same components used to purify the protein illustrated in Figure 11B. Notably, the proteins maintained their homogenous behavior and displayed monodispersed peaks as seen in the larger scale purification of TRPP2 in Figure 11B. Remarkably, the TRPP2 protein was stable even after being heated to 37°C. Not only did this experiment indicate the stabilizing effect soybean lipids had on the TRPP2 F604P-MBP protein, but it also reassured us that the protein sample would most likely be stable during shipment to our collaborators in China.
Figure 11. Gel filtration profiles of TRPP2 F604P purified with LMNG and soybean lipids. (A) Gel filtration profile of human TRPP2:199-927 F604P-MBP purified in a buffer containing .005% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4. (B) Gel filtration profile of human TRPP2:199-927 F604P-MBP purified in a buffer containing .005% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4, with the addition of .001% soybean lipids. (C) Gel filtration profile of human TRPP2:199-927 F604P-MBP purified in the buffer indicated in (B). After elution from the amylose resin, the sample was stored at 4°C for 5 days before being run on the gel filtration column. (D) Same procedure utilized in (C), however, the protein was heated to 37°C before being injected onto the gel filtration column.
Therefore, since the SEC profiles looked promising for TRPP2:199-927 F604P-MBP in LMNG, especially in the presence of soybean lipids, I purified another batch of this protein and sent it to our collaborators at Tsinghua University. Of note, all negative-stain and grid preparations that will be discussed in the rest of this chapter were performed by the Li group at Tsinghua University. Additionally, the Li group ran the shipped protein samples through the gel filtration column so that the grids could be made with fresh protein. The samples were run on the size-exclusion column in a buffer containing .0008% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4, with the addition of .0002% soybean lipids. However, although the SEC profiles of TRPP2:199-927 F604P-MBP in LMNG looked promising, the negative-stain and EM grids did not. These grids exhibited a fair amount of protein aggregates and the protein was not detectable in the EM grids (Figure 12), presumably due in part to the dissociation of the protein.

Figure 12. Negative-stain and EM grids of the hTRPP2:199-927 F604P-MBP in LMNG (A-B) Representative negative-stain image (A) and EM micrograph (B) of human TRPP2:199-927 F604P-MBP
Similarly, upon analysis of the TRPP2 F604P-MBP protein reconstituted in amphipol A8-35, both the negative-stained and EM grids exhibited a large amount of broken or aggregated particles (Figures 13A, B). Within some grids, the ice appeared to be too thin, which may have exacerbated the particles, causing them to dissociate to a greater extent. However, although the protein displayed relatively monodispersed SEC profiles when reconstituted in amphipol, it appeared that the TRPP2 truncated proteins could not form stable tetramers and thus were not able to be used for structural analysis. We previously observed similar results from grids containing wild-type TRPP2 reconstituted in amphipol after being solubilized in DDM or LMNG. It remains to be seen why our samples reconstituted in amphipol appeared to have exhibited dissociated particles when both negative-stain and EM grids were made. Interestingly, prior to setting up grids, the addition of detergent to the amphipol sample resulted in the re-association of the channel and the quality of particles present in the grids increased considerably. Therefore, for the 199-927 TRPP2 truncation, it appeared that amphipol samples were not suitable for data collection.

Figure 13. Negative-stain and cryo-EM grids of hTRPP2 F604P in amphipol A8-35. (A-B) Illustrative negative-stain EM (A) and cryo-EM image (B) of hTRPP2 F604P-MBP after reconstitution in amphipol. Aggregation and dissociation of protein particles are apparent.
Nonetheless, it is important to consider that analysis of proteins by size-exclusion chromatography and SDS-PAGE does not necessarily predict how samples will behave when subjected to EM analysis. Therefore, the most effective way to assess the quality of a protein sample is to test it on negative-stain grids (42). Since images of particles are averaged together in order to obtain a 3D reconstruction, it is important for protein samples to be homogenous, especially to attain high-resolution. Yet preparing homogeneous samples for cryo-EM analysis remains challenging (41).

### 3.9 Titan Krios data collected at Tsinghua University

Our collaborators at Tsinghua University next tested hTRPP2:199-927 samples purified in 0.0008% LMNG, 150 mM NaCl, and 20 mM Hepes pH 7.4. Our goal was to try to obtain sharper 2D class averages as well as to solve the particle orientation dilemma that we previously encountered. Just prior to making the grids, .2% glycerol was added to the purified protein sample in order to minimize particle aggregation. For creating EM grids, 4µl of the human TRPP2 protein (.5mg/ml) were loaded onto glow-discharged Quantifoil R1.2/1.3 holey carbon 400 mesh copper grids. The Vitrobot (FEI) was set at 100% humidity at 8°C and blotted for 3.5 s, blot force 0. The grids were then plunged into liquid ethane cooled by liquid nitrogen. Micrographs were subsequently acquired by a Titan Krios TEM equipped with a K2 Summit direct electron detector (Gatan).

2D classification of the particles was then carried out and 2D class averages (Figure 14C) highlighted common features of the particles. Based on these results, it was evident that the preferred orientation dilemma was still present, since we could not see any apparent side views. However, from 2D classification of the particles we were able to obtain the same view as
previously observed (Figure 9D), yet this view of TRPP2 demonstrated an improvement in overall particle quality. Notably, we were able to discern prominent features including the S1-S4 helices of the voltage sensing domain (outer periphery) as well as the central pore within TRPP2 (Figure 14C). Collectively, our 2D class averages represent the TM domain cross-sections of TRPP2 in great detail and are very similar to a view seen in the solved TRPP2 structure as reported by Grieben et al. (16).

Besides adding glycerol to the samples, we used other additives upon grid preparation in order to overcome the preferred particle orientation dilemma. Of note, in a previous study reporting the ryanodine receptor (RyR1) in its closed state 0.2% of fluorinated octyl-maltoside, a non-solubilizing detergent, was applied to the protein buffer prior to setting up the EM grids (44). Fluorinated octyl-maltoside can presumably enhance protein distribution within vitreous ice (44). Therefore, we also used this detergent to assess whether it too could improve the distribution of TRPP2 particles. However, this did not ameliorate the quality of the particles to a great extent. Hence, the preferred particle orientation dilemma continued to persist.
Figure 14. Cryo-EM analysis of hTRPP2:199-927 in LMNG. (A-B) Representative negative-stain image (A) and electron micrograph (B) of human TRPP2:199-927 in .0008% LMNG. Ice particles were also present within the EM grids. Protein particles were selected and subjected to 2D screening. (C) Enlarged views of representative TRPP2 2D class averages, highlighting detailed features of the tetrameric channel. (D) Ribbon representation of the TRPP2 tetramer as reported by Grieben et al. (16). Voltage sensing domain S1-S4 helices are depicted in light blue, and the pore domains (S5, pore helix 1, pore helix 2, S6) are denoted in dark blue. The linker domain between S1 and S2 is displayed in orange and yellow. Glycosylation sites are indicated by sticks. Notably the S1-S4 helices are clearly visible within (C), resembling the architecture of the solved TRPP2 structure (D). Of note, (D) was adapted from (16).
3.10 Comparison of methods used for TRPP2 structural studies

Initially, Shen et al. (15) reported the first TRPP2 cryo-EM structures at 3Å resolution in a closed conformational state from three N- and C-terminally truncated constructs reconstituted into amphipols and nanodiscs (Figure 15). Thereafter, Grieben et al. (16) also published a closed state TRPP2 cryo-EM structure at 4.2 Å of a truncated construct (Figure 15). Of note, Grieben et al. (16) also collected X-ray crystallography data of TRPP2, yet at a lower resolution, 5.5 Å. Additionally, Wilkes et al. also reported TRPP2 cryo-EM structures at 4.2 Å and 4.3 Å, purportedly displaying open conformational states of the full-length channel. Thus, this section serves to analyze the differences in methods implemented by these three independent research groups as well as our group for structural determination of TRPP2.

3.10.1 Constructs and protein purification methods

Cytoplasmic domains of TRPP2 are not necessary for structural stability

After analysis of these publications, what we found to be most striking is that the human TRPP2 structures remained stable without their C-terminal cytoplasmic domains, especially the coiled-coil domain, and the structures were solved without their presence. Like all other TRP channels, TRPP2 formed a tetrameric structure (15–17). Yet, as previously mentioned, TRPP2 has been found to form trimers when expressed in Xenopus oocytes in HEK293 cells (3). Additionally, the coiled-coil domain was shown to be necessary for the homomeric assembly of TRPP2 (3). Therefore, I always created truncations that included this domain in order to presumably stabilize the protein (Figure 15).
As indicated in Figure 15, Shen et al. made use of a hTRPP2:198-703 construct, which did not include the EF-hand domain as well as two other constructs TRPP2:198-792 and TRPP2:53-792, which did contain these domains. Interestingly, all three of the reported TRPP2 structures determined by Shen et al. (15) were superimposable and very similar with no additional densities present for the EF-hands in any of the structures. These findings suggest that the intracellular domains of TRPP2 are unstructured and flexible.

![Figure 15. Schematic diagram displaying constructs utilized in TRPP2 structural studies.](image)

hTRPP2:1-968 was used by Wilkes et al. (17); hTRPP2:53-792, 198-792, and 198-703 were used by Shen et al. (15); hTRPP2:185-723 was utilized by Grieben et al. (16); and hTRPP2:199-927 was used in the Yang lab. Diagram modified from (15).

Notably, the densities of the termini within the structures reported by all three groups could not be elucidated. Of note, the hTRPP2:199-927 construct that I utilized included the coiled-coil domain. Thus, this region was likely to be flexible as well, since Wilkes et al. utilized the full-length construct of TRPP2 containing the coiled-coil domain and could not resolve this region.
Additionally, the hTRPP2:199-927 truncation that I utilized was very similar at the N-terminus to the other truncations used for structural determination.

Furthermore, it has been reported that conformational changes within NMR structures of the EF-hand domain occurred upon binding of calcium (4). Taken together, although the C-terminal domains were not able to be resolved in the various TRPP2 structures due to their flexibility, it remains to be seen what role the N- and C-terminal domains play in the overall TRPP2 channel. Moreover, it is apparent that the presence of these domains did not produce significant heterogeneity in the protein samples to hinder structural analysis.

Table 1 indicates the methods implemented in various TRPP2 studies to express, purify, and analyze protein samples using cryo-EM. However, it is important to note that many parameters during the purification and sample preparation procedures vary from one group to another.
Table 1. Protein purification procedures for structural studies of TRPP2

<table>
<thead>
<tr>
<th></th>
<th>Shen et al., 2016</th>
<th>Grisben et al., 2017</th>
<th>Wilkes et al., 2017</th>
<th>Yang and Li, Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>3 Å</td>
<td>4.2 Å</td>
<td>4.2 Å (TRPP2α)</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.3 Å (TRPP2β)</td>
<td></td>
</tr>
<tr>
<td>Protein truncation</td>
<td>Human TRPP2: 196-703, 196-792, 53-792</td>
<td>Human TRPP2: 185-723</td>
<td>Human TRPP2: Full-length</td>
<td>Human TRPP2: 199-927</td>
</tr>
<tr>
<td>Tag</td>
<td>N-terminal MBP Removed with TEV</td>
<td>C-terminal Flag tag Removed with TEV</td>
<td>Strep-II tag Not removed</td>
<td>C-terminal MBP Not removed</td>
</tr>
<tr>
<td>Expression system</td>
<td>HEK293S GoT&lt;sup&gt;T&lt;/sup&gt; cells</td>
<td>SF9 insect cells</td>
<td>HEK293S GoT&lt;sup&gt;T&lt;/sup&gt; cells</td>
<td>Hi5 insect cells</td>
</tr>
<tr>
<td>Litters of cultured cells for purification</td>
<td>Not available</td>
<td>24 L</td>
<td>4 L</td>
<td>3.5 L</td>
</tr>
<tr>
<td>Detergent solubilization</td>
<td>DDM Buffer not specified</td>
<td>1% DDM, 150 mM NaCl, 5% glycerol, 50 mM HEPES, 20 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, pH 7.5</td>
<td>1% 0.4% LMNG/CHS 150 mM NaCl, 10% glycerol, 25 mM HEPES, 2.5 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, pH 7.5</td>
<td>1% LMNG in buffer, 200 mM NaCl, 10% glycerol, 50 mM HEPES pH 8.0</td>
</tr>
<tr>
<td>Protein elution</td>
<td>0.5 mM DDM, 20 mM maltose, 150 mM NaCl, 50 mM HEPES, 2 mM TCEP, 2 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, 0.1 mg/ml soybean lipids pH 7.0</td>
<td>0.033% UDM, 400 mM imidazole, 150 mM NaCl, 5% glycerol, 50 mM HEPES, 20 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, pH 7.5</td>
<td>5 mM desthiobiotin Buffer not specified</td>
<td>0.01% LMNG, 40 mM maltose, 200 mM NaCl, 10% glycerol, 50 mM HEPES pH 8.0</td>
</tr>
<tr>
<td>After protein elution</td>
<td>Nanodisc, PKD2-MSFP2N2:soybean lipids 1:1/200 molar ratio, TEV protease (40 µg/mg TRPP2), 30 µg Bio-Beads per 1 ml reconstitution mixture</td>
<td>Buffer exchange, 0.033% UDM, 150 mM NaCl, 5% glycerol, 20 mM MES, 20 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, pH 6</td>
<td>Amphipol A8-35: added in three-fold excess (w/v), 15 mg Bio-Beads added per 1 ml TRPP2-detergent-ampiphilic mixture</td>
<td>.0005% LMNG, Amphipol A8-35: added to TRPP2 in a six-fold excess (w/v), 15 mg Bio-Beads added per 1 ml TRPP2-detergent-ampiphilic mixture</td>
</tr>
<tr>
<td>Gel filtration buffer</td>
<td>150 mM NaCl, 10 mM HEPES, 2 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, pH 7.4</td>
<td>0.033% UDM, 200 mM NaCl, 20 mM HEPES, 20 mM CaCl&lt;sub&gt;2&lt;/sub&gt;, pH 7.5</td>
<td>Buffer not specified</td>
<td>.0005% LMNG, 150 mM NaCl, 20 mM HEPES pH 7.4 or 150 mM NaCl 20 mM HEPES pH 7.4</td>
</tr>
</tbody>
</table>

References  (15)  (16)  (17)
<table>
<thead>
<tr>
<th>Table 1 (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shen et al., 2016</strong></td>
</tr>
<tr>
<td>Glycosylation reduced</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Presence of glycosylated residues in solved structures</td>
</tr>
<tr>
<td>Lipids added during purification</td>
</tr>
<tr>
<td>Presence of lipids in solved structures</td>
</tr>
</tbody>
</table>

References  
(15)  (16)  (17)
Expression systems and glycosylation

It appears that in the case of purifying TRPP2, HEK293S GnTI− and HEK293S GnTI− cells are the most preferable expression systems, followed by Sf9 cells, and Hi5 cells. In agreement with our findings, Grieben et al. (16) also demonstrated that it is not necessary to purify TRPP2 in HEK cells, yet, since this system provides a more native environment for the protein, it appears to be beneficial. However, we did not have the capability or resources to perform large-scale purification from HEK cells in our lab. Interestingly, Grieben et al. (16) purified protein from 24 liters of Sf9 cells instead of Hi5 cells, even though Hi5 cells have been shown to produce protein at a higher yield (27). Thus, I primarily used Hi5 cells for protein expression and could obtain good protein yield from a few liters of culture.

It has been reported that glycosylation of TRPP2 is key for its stability and trafficking (36). As previously mentioned, five N-glycosylation sites are present within TRPP2 (Asn299, Asn305, Asn328, Asn362, and Asn375) which reside within the linker domain present between the first two transmembrane segments of the channel (36). The structural implications of the presence of glycosylation sites in TRPP2 will be further elucidated in Chapter IV. While Shen et al. (15) and Wilkes et al. (17) expressed TRPP2 in HEK293S GnTI− and HEK293S GnTI− cells, respectively, Grieben et al. (16) performed extensive enzymatic deglycosylation using α-L-fucosidase and PNGase. However, residual oligosaccharides were present within the structures (15–17). Of note, the protein that I primarily purified was in a glycosylated state.
Detergents and lipids

The choice and use of detergents during protein purification and EM sample preparation appeared to have played an important role in the overall success of this project. Since the MNG family of detergents are relatively new, the use of LMNG for cryo-EM analysis has not been extensively studied. Although Wilkes et al. solubilized TRPP2 in LMNG, the protein was later reconstituted in amphipol and EM grids were not made with LMNG (17). Alternatively, Grieben et al. solubilized TRPP2 in DDM, yet set up EM grids using the UDM (n-Undecyl-β-D-maltopyranoside) detergent (16). Of note, UDM has one less carbon than DDM in its chemical structure. Presumably, Grieben et al. screened through various detergents to find the most promising one for structural studies.

Most of our efforts were geared towards optimizing the detergent concentration conditions used during protein purification as well as grid preparation. Since samples solubilized in DDM and LMNG had promising SEC profiles, we primarily focused on testing these two detergents. Although we conducted additional experiments to determine the stability of TRPP2 when solubilized in additional detergents including OG, CHAPS, and DMNG, cryo-EM analysis of these samples were not extensively tested. However, our findings indicated that DDM in general was not a suitable detergent for cryo-EM analysis and LMNG may not be conducive to the proper distribution of particles. Therefore, these findings may demonstrate why DDM and LMNG detergents were not used during EM grid preparation by other groups. Thus, these observations underscore the need to test a wide variety of detergents in order to assess which are most suitable for structural studies of a particular protein.
Notably, lipids were visible in the TRPP2 structures described by Shen et al. (15) and Wilkes et al. (17). I recently started incorporating soybean lipids into my purification protocols, which presumably helped to stabilize the channel. Upon the addition of soybean lipids, although the gel filtration profiles of both the wild-type and mutant (F604P) TRPP2 constructs demonstrated greater monodispersity than without the presence of lipids, the protein particles seemed to be similar when tested on cryo grids.

**Gel filtration profiles of TRPP2**

Upon observation of the gel filtration profiles of TRPP2 reported by Shen et al. (15), Wilkes et al. (17), and our research group, several patterns are apparent (Figure 16). Notably, the gel filtration profile from the TRPP2:199-927 protein reconstituted in amphipol A8-35 that I purified (Figure 16E) is similar to the profiles reported by the other groups (Figures 16A, D). However, the amphipol A8-35 profiles displayed by Shen et al. and Wilkes et al. had peaks that appeared to be more monodispersed, indicating samples with greater homogeneity. Additionally, the gel filtration profile for the TRPP2:199-927 protein in .0008% LMNG (Figure 16F), our most promising sample for structural analysis, seemed to be broader than when reconstituted in amphipol A8-35 (Figure 16E) and had the broadest peak overall. Furthermore, it appears that the hTRPP2 proteins reconstituted in lipid nanodiscs (Figures 16B, C) resulted in a larger amount of aggregated protein compared to hTRPP2 samples reconstituted in amphipol A8-35 (Figures 16A, D).
Figure 16. Comparison of gel filtration profiles of hTRPP2. (A-C) Gel filtration profiles reported by Shen et al. (15) of hTRPP2:53-792 in amphipol A8-35 (A), hTRPP2:198-792 in lipid nanodiscs (B), and hTRPP2:198-703 in lipid nanodiscs (C). (D) Gel filtration profile of full-length hTRPP2:1-968 in amphipol A8-35 shown by Wilkes et al. (17). (E) Gel filtration profile of hTRPP2:199-927 in amphipol A8-35 purified in our lab. (F) Gel filtration profile of hTRPP2:199-927 purified in our lab in .005% LMNG and .001% soybean lipid.
## 3.10.2 Parameters for cryo-EM analysis

### Table 2. Sample preparation and cryo-EM data acquisition parameters for TRPP2 studies

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Shea et al., 2016</th>
<th>Grieben et al., 2017</th>
<th>Wilkes et al., 2017</th>
<th>Yang and Li Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EM grid preparation: concentration</strong></td>
<td>3.5 μl of TRPP2 concentrated to 1-3.5 mg/ml for nanodiscs or 1 mg/ml TRPP2 in amphipol A8-35</td>
<td>3 μl of TRPP2 in UDM concentrated to 4.5 mg/ml</td>
<td>3 μl of TRPP2 in amphipol A8-35 concentrated to 1 mg/ml</td>
<td>4 μl of TRPP2 in .0008% LMNG concentrated to .5 mg/ml or 1-2 mg/ml TRPP2 in amphipol A8-35</td>
</tr>
<tr>
<td><strong>EM grid preparation parameters</strong></td>
<td>-Glow-discharged Quantifoil 1.2/1.3 holey carbon 400 mesh copper grids</td>
<td>-Glow-discharged holey carbon grids 200 mesh Protocips</td>
<td>-Glow-discharged holey carbon grids (Quantifoil R2/2)</td>
<td>-Glow-discharged Quantifoil R1.2/1.3 holey carbon 400 mesh copper grids</td>
</tr>
<tr>
<td></td>
<td>-Vitrobot Mark III (FEI), 75% humidity, 4°C, 20 s wait time, -1 mm offset, blotting time 7 s</td>
<td>-Vitrobot Mark IV, (FEI), 80–100% humidity, blotting time 3 s</td>
<td>-Vitrobot IV (FEI), 100% humidity, 10°C, blotting time 9 s</td>
<td>-Vitrobot (FEI), 100% humidity, 8°C, blotting time 3.5 s, blot force 9</td>
</tr>
<tr>
<td></td>
<td>-Plunge frozen in liquid ethane</td>
<td>-Plunge frozen in liquid ethane</td>
<td>-Plunge frozen in liquid ethane</td>
<td>-Plunge frozen in liquid ethane</td>
</tr>
<tr>
<td><strong>Data acquisition</strong></td>
<td>Tecnai TF20 TEM (FEI), 300 kV, K2 Summit direct detector (Gatan) or Tecnai TF30 Polara TEM (FEI), 300 kV, K2 Summit direct detector</td>
<td>Tecnai F30 Polara TEM (FEI), 300 kV, K2 Summit direct detector (Gatan)</td>
<td>JEOL 3200 FSC TEM, 300 kV, K2 Summit direct detector (Gatan)</td>
<td>Titan Krios TEM, 300 kV, K2 Summit direct detector (Gatan)</td>
</tr>
</tbody>
</table>

References: (15) (16) (17)

Table 2 illustrates the parameters used for sample preparation and cryo-EM data acquisition, after the proteins were purified using size-exclusion chromatography. Of note, the Yang and Li group column refers to the parameters used by our collaborator at Tsinghua University. Markedly, the concentrations that we used to prepare EM grids of TRPP2 in LMNG were generally lower.
than the concentrations used by other groups. This may have played a role in the overall quality of the grids that we observed.

Representative EM micrographs from the various research groups are illustrated in Figure 17. Notably, our micrograph denoted in Figure 17D contains particles that look similar to those present in the other micrographs.

**Figure 17.** Representative micrographs of human TRPP2. (A-D) Representative EM micrographs of TRPP2 samples from various research groups. (A) hTRPP2:198-703 in lipid nanodiscs reported by Shen et al. (15). (B) hTRPP2:1-968 in amphipol revealed by Wilkes et al. (17). (C) hTRPP2:185-723 in 0.035% UDM depicted by Grieben et al. (16). (D) hTRPP2:199-927 in 0.0008% LMNG purified from our lab and tested by the Li group.
After data collection from the micrographs, TRPP2 particles were subjected to analysis and 2D class averages were generated (Figure 18). Although other groups obtained 2D class averages exhibiting particles in multiple random orientations (Figures 18A-E) as previously mentioned, the particles from our sample only displayed a top view displaying TM domain cross-sections (Figure 18F).

Figure 18. 2D class averages of human TRPP2 cryo-EM particles. (A-C) 2D class averages reported by Shen et al. (15) of hTRPP2:198-703 in lipid nanodiscs (A), hTRPP2:198-792 in lipid nanodiscs (B), and hTRPP2:53-792 in amphipol A8-35 (C). (D) Representative 2D class averages of hTRPP2:185-723 in 0.035% UDM reported by Grieben et al. (16). (E) 2D class averages of hTRPP2:1-968 in amphipol A8-35 revealed by Wilkes et al. (17). (F) Representative 2D class averages of hTRPP2:199-927 in .0008% LMNG, purified in our lab, illustrating detailed features of the preferred orientation.
Essentially, the top view seen in our sample (Figure 18F) seemed to have the highest quality features for this particular orientation, while the other 2D class averages with the same view appeared blurry. Notably, the four-fold symmetrical axis in the tetramer is clearly visible from the 2D class averages of hTRPP2:199-927, highlighted by distinctive S1-S4 helices and the central pore.

3.11 Analysis of new constructs reconstituted into nanodiscs and amphipols

I was still optimizing methods in order to overcome the preferred orientation dilemma for the TRPP2 F604P protein when the first TRPP2 structure was published. I then made additional truncations removing the C-terminal coiled-coil domain, yet keeping the EF-hand domain. The TRPP2 F604P-MBP:181-787 construct was expressed and purified in DDM as previously described, with modifications. Consequently, the protocols for the reconstitution of TRPP2 in nanodiscs were adapted from Shen et al. and Gao et al. (10, 15). In brief, soybean polar lipid extract was rehydrated by sonication in the buffer utilized to elute the protein, producing a clear lipid stock at 10 mM concentration. The MSP2N2 membrane scaffold protein was expressed and purified from E. coli as previously reported (45). Purified TRPP2 channel (1 mg/ml) solubilized in 0.5 mM DDM was then combined with MSP2N2 (170 - 250 mM) and the soybean lipid stock at a 1:1:200 molar ratio. This mixture was then incubated on ice for 30 min. Subsequently, Bio-Beads were added to remove DDM at 1 hr time intervals for a total of four batches. The sample was gently rotated at 4°C.

TEV protease was then added in combination with the final batch of Bio-Beads followed by sample incubation overnight at 4°C. Bio-Beads were then removed and the nanodisc reconstitution mixture was clarified by centrifugation before loading onto a Superose 6 column in a buffer
containing 150 mM NaCl, 10 mM HEPES, and 2 mM CaCl$_2$ pH 7.4. Of note, calcium chloride was added in all buffers during purification, adapted from Shen et al. (15), since it was suggested that calcium can stabilize the EF-hand domain (4, 6).

However, upon SEC analysis of the nanodisc sample, the main included peak corresponding to tetrameric TRPP2 was quite broad and not monodispersed, suggestive of heterogeneity in the sample (Figure 19A). Nonetheless, this peak was further assessed by preparing and testing negative-stain and cryo-EM grids (Figures 19B, C). The cryo grids were generated with 1.9 mg/ml of TRPP2 sample. Yet, there appeared to be a fair amount of aggregation in the EM grids (Figure 19C). Therefore, the parameters we used to purify and analyze the sample needed to be further optimized to produce suitable particles.

**Figure 19. Analysis of TRPP2 reconstituted in nanodiscs.** (A) Gel filtration profile of hTRPP2 F604P:181-787 after reconstitution in lipid nanodiscs indicates a heterogeneous sample. Demonstrative negative-stain (B) and cryo-EM (C) micrographs of TRPP2 illustrating aggregated protein particles. Ice particles are also present in (C).
Additionally, the hTRPP2 F604P:185-723 construct was also generated based on the truncation used by Grieben et al. (16). This protein was purified in parallel with either DDM or LMNG as previously mentioned. The hTRPP2 F604P:185-723 truncation contained the EF-hand domain but lacked the coiled-coil domain. The eluted TRPP2 protein was mixed with amphipol A8-35 at a 1:6 (w/w) ratio and subsequently incubated overnight with gentle agitation at 4 °C. To enable the substitution from detergent to amphipol to occur, the sample was incubated with Bio-Beads for 8 hr. After removal of Bio-Beads, TEV protease was added to the protein sample and incubated at 4°C overnight. TRPP2 proteins were then concentrated by centrifugation prior to loading onto the Superose 6 gel filtration column. This method for reconstitution of TRPP2 in amphipols was adapted from Shen et al.(15).

The hTRPP2 F604P:185-723 samples that were solubilized in LMNG and reconstituted in amphipol appeared to have had greater monodispersity in their gel filtration profiles (Figures 20D, G) rather than when TRPP2 was first purified in DDM and then transferred to amphipol (Figure 20A). The presence or removal of the MBP tag did not appear to play a major role with regard to the monodispersity of the samples (Figures 20D, G). However, removal of the MBP tag (Figures 20G-I) in samples that were solubilized in LMNG and transferred to amphipol led to protein particles that were less aggregated and dissociated (Figure 20I), compared to samples when the tag was not removed (Figure 20F). Nonetheless, the cryo-EM samples in amphipol were not as suitable for analysis as were those in LMNG, despite the greater level of monodispersity in the amphipol samples.
Figure 20. Assessment of the homogeneity of TRPP2 F604P. (A) Gel filtration profile of hTRPP2 F604P:185-723 after being purified in DDM, reconstituted in amphipol, and cleaved with TEV. Negative-stain (B) and EM micrographs (C) of hTRPP2 F604P:185-723 purified in (A) are indicated. The protein concentration used to make the EM grids was .96 mg/ml. (D) Gel filtration profile of hTRPP2 F604P-MBP:185-723 after being purified in LMNG and reconstituted in amphipol, without being cleaved with TEV. Notably, the TRPP2 tetrameric peak is monodisperse and therefore homogeneous. Negative-stain (E) and cryo-EM images (F) of hTRPP2 F604P-MBP:185-723 purified in (D). 1mg/ml of protein was used to generate the EM grids. (G) Gel filtration profile of purified hTRPP2 F604P:185-723 after exchange from LMNG to amphipol, and being cleaved with TEV. The TRPP2 protein peak is monodisperse and purified to homogeneity. Negative-stain (H) and electron micrographs (I) of hTRPP2 F604P-MBP:185-723 purified in (G). 2mg/ml of protein sample was used to make cryo-EM grids for data collection.

Interestingly, less aggregation was present in the negative-stain micrographs of purified hTRPP2 F604P:185-723 in amphipol (Figures 20B, E, H) compared to the negative-stain images of purified hTRPP2 F604P-MBP:199-927 in amphipol (Figure 13A). The 185-723 truncation may have played a role in stabilizing the protein in amphipol and creating greater homogeneity in the sample, as seen in both the gel filtration profiles and the negative-stain grids (Figure 20). However, the EM grids of the two different truncated proteins looked similar.
3.12 Summary

Structural determination of membrane proteins has and remains to be inherently challenging (46). Due to the intrinsic nature of membrane proteins, various barriers exist including obtaining a sufficient amount of protein as well as finding conditions that stabilize the protein in a correctly folded state. Although we were able to overcome these barriers, the preferred orientation exhibited by the TRPP2 particles was by far the greatest obstacle we had to surmount in the TRPP2 project.

Heterogeneity is a main problem that all structural biologists must overcome with their samples. It became apparent to us that heterogeneity may not only significantly hinder the crystallization of a protein, but also cryo-EM analysis of a sample as well. Therefore, it is important to have homogeneity in the sample, although some heterogeneity can prove to be beneficial to provide alternative conformational states. The homogeneity of the TRPP2 protein samples varied from one protein prep to another and was assessed by a variety of techniques including FSEC, SEC, and SDS-PAGE prior to preparing negative-stain and EM grids. Hence, we pursued structural analysis with the samples that looked most promising.

Our most significant finding was obtaining a view of hTRPP2-MBP:199-927 that contained highly detailed structural features. Remarkably, looking down the four-fold symmetrical axis in the 2D class averages, we were able to clearly see the central pore as well as the S1-S4 helices within the tetramer. However, the distribution of the particles appeared to be too homogeneous in our case. Hence, the particles exhibited a strongly preferred orientation impeding us from moving forward to create a reasonable 3D reconstruction due to the lack of particles in additional orientations.

Since protein purification and sample preparation requires a series of steps and includes many variables, it is challenging to determine which conditions specifically contributed to the particle
orientation dilemma that we had to overcome. Over time, the protein preps were optimized by altering a number of parameters including the protein truncations, detergents, buffer solutions, and lipid concentrations. One of the primary factors that likely contributed to the presence of the preferred particle orientation was the type of detergent utilized during protein purification and sample preparation, primarily DDM and LMNG. Of note, during small-scale preps with DDM or LMNG, the SEC TRPP2 peaks appeared more monodispersed compared to protein purified in large-scale preps. Therefore, an increase in protein concentration led to greater heterogeneity of the samples. Hence, although TRPP2 was stable below the CMC of the detergents, upon SEC analysis, it appeared that DDM or LMNG detergents were not very suitable for cryo-EM analysis. EM grids that contained DDM did not exhibit good particles, while EM grids containing LMNG displayed particles with a particular particle orientation. Studies to further investigate how to address the preferred orientation conundrum in the TRPP2 F604P samples are currently underway.

Moreover, additional factors may have resulted in the heterogeneous samples we were obtaining. Protein truncations, glycosylation, and the expression systems implemented could also play a role in the proper folding of TRPP2 and may have contributed to the heterogeneity of the sample. Of note, the amphipol dilemma I originally encountered during purification, whereby TRPP2 was precipitating out of solution upon amphipol reconstitution, was resolved. The amphipol reconstitution protocol was optimized by altering the concentration of amphipol utilized and adding Bio-Beads in batches to stabilize the protein during reconstitution. However, TRPP2 protein particles in amphipol did not look promising.
3.13 Conclusion

New technological advances in the field of single-particle cryo-EM provide an excellent opportunity for high-resolution structural determination of membrane proteins. Overall, the structural investigations described in this chapter provide valuable insights into the optimization of the TRPP2 protein purification and sample preparation procedures necessary to obtain a high-resolution cryo-EM structure. However, although we made great strides towards producing human TRPP2 protein with suitable yield and purity for cryo-EM analysis, a major obstacle that we needed to overcome in order to obtain a 3D model of the protein was the preferred orientation that TRPP2 was fixed in. It remains to be determined what factors specifically caused the preferred particle orientation. Despite my ambitious efforts to obtain a cryo-EM structure of human TRPP2, additional research groups published high-resolution structures of the same protein. A detailed understanding of the solved TRPP2 channel structures will be described in Chapter IV. These structures provide a basis from which to explore and elucidate a wide array of additional TRPP2 structures.
3.14 References


4.1 Introduction

As aforementioned, three research groups recently solved cryo-EM structures of TRPP2. The methods utilized to purify these proteins were discussed in Chapter III. The aim of this chapter is to provide a detailed understanding of the structural features of TRPP2. Furthermore, comparisons between the solved TRPP2 structures as well as their respective ion-permeation pathways will be elucidated.

4.2 General architecture of the TRPP2 structures

Like other TRP channel structures (1–3), the structure of TRPP2 (4–6) is tetrameric and contains six transmembrane helices (Figures 1A, C). The transmembrane region of each TRPP2 monomer contains a voltage-sensing domain (VSD) which includes the S1 to S4 helices, and a pore domain consisting of the S5 and S6 helices (Figure 1B) (4). In addition, the S4-S5 linker joins the VSD and pore domains (4). Moreover, a P-loop structure exists between S5 and S6, which comprises pore helix 1, a pore filter (Leu641-Gly642) (Figure 8B), and pore helix 2 (5). Within the tetramer, each VSD domain in one monomer participates in ‘domain swapping’ with the pore domain of the adjacent monomer (4). Although all three determined TRPP2 structures share general features, unique characteristics and properties will be expounded upon in subsequent sections.
Figure 1. Cryo-EM structure of human TRPP2 in lipid nanodiscs. (A) View of TRPP2 in a closed state conformation from the side (left) or bottom (right). Each monomeric subunit within the tetramer is color coded. Densities of the nanodisc (gray) and of lipids (purple) are also indicated. (B-C) Ribbon representations of the TRPP2 monomeric (B) and tetrameric structures illustrating each monomeric subunit with a different color as in (A). Figure adapted from (4).
4.3 Extracellular polycystin-mucolipin domain in TRPP2

Notably, an extracellular linker is present within TRPP2, which is also a characteristic feature of TRPML channels (7). Shen et al. erroneously mentioned that this domain is limited to the PKD2s, PKD1, and PKD1-like proteins (4, 8). In fact, all three research groups that solved the TRPP2 structures (4–6) failed to acknowledge that the TRPML subfamily exhibits this domain. The research groups referred to this linker as the polycystin domain (4) or the tetragonal opening for polycystin (TOP) domain (5, 6). Instead, we hereby propose that the I–II linker domain in TRPP2 be regarded as the ‘polycystin–mucolipin domain’ (9) due to its structural similarity in both TRPP2 and TRPML1 (Figures 4, 5). Henceforth, the I-II linker domain in TRPP2 and TRPML1 will be referred to as the polycystin–mucolipin domain in this chapter.

4.3.1 Structural features of the polycystin-mucolipin domain

The polycystin-mucolipin domain in TRPP2 comprises 219 amino acid residues (Ser244-Leu462) positioned in three α helices (α1–α3) and a central five-stranded anti-parallel β-sheet (4, 5). Of significance, the polycystin-mucolipin domain in TRPP2 is a tetramer, similar to that of the corresponding domain in TRPML1 (9). Additionally, three β-hairpins are situated at ~ 90° intervals (Figures 2A, B) (5). Since these structural elements are reminiscent of a three-leaf clover (TLC), these β-hairpins can be referred to as TLC 1-3 (Figures 2A, B) (5). Furthermore, TLC 1-3 interacts with adjacent polycystin-mucolipin domains, the S3–S4 extension, and the pore, enabling intersubunit cooperativity (5).
Figure 2. Polycystin–mucolipin domain structure of TRPP2. (A) Schematic demonstrating the VSLD (voltage-sensor-like domain) in light blue, PD (pore domains) in dark blue, and polycystin–mucolipin domain (orange and yellow). The helices, β-strands, and glycosylation sites in the polycystin–mucolipin domain are depicted by cylinders, arrows, and hexagons, respectively. (B) Perpendicular orientations of the polycystin–mucolipin domain. The five-stranded β-sheet, β1-5, is indicated in yellow while the three α-helices (α1–3) are denoted in orange. The TLC extension (dark red) includes a long loop between β1 and β2 generating three β-hairpins (TLC1–3). Glycosylation sites, NAG 3–5, are indicated in pink. Figure adapted from (5).

Moreover, extensive inter- and intra-subunit interactions are established within the polycystin–mucolipin domain that aid in stabilizing the structure. For instance, TLC 1 and the β4-β5 turn of one subunit interacts with the β3-β4 turn and the α1 helix from the polycystin–mucolipin domain of an adjacent subunit (Figure 3A) (4). Furthermore, the β4-β5 turn also interacts with α1 of an adjoining domain through hydrogen bonding between Thr448 and Tyr248 (Figure 3A). Evidently,
the polycystin-mucolipin domain appears to play a role in channel assembly, due to the presence of extensive associations with the polycystin-mucolipin domain in adjacent subunits (4). In addition, the polycystin-mucolipin domain associates with the TRPP2 transmembrane core at the helix turn and pre-S6 loop (Figure 3B) (4).

**Figure 3. Interactions in the polycystin–mucolipin domain affecting tetrameric assembly.** (A) Left, TLC 1 and the β4-β5 turn of one subunit (blue surface representation) interacts with the β3-β4 turn and the H1 helix (α1), indicated in yellow, from the polycystin-mucolipin domain of an adjacent subunit (ribbon). Right, enlarged view of interactions. (B) The polycystin–mucolipin domain (blue surface representation) associates with transmembrane segments (ribbons) at the pre-S6 loop and the helix turn indicated by the magenta boxes. Figure adapted from (4).
4.3.2 Comparison of polycystin–mucolipin domains

The structure of the polycystin-mucolipin domain in TRPP2 is similar to the overall architecture of the TRPML1 I-II linker (Figure 5A) (9), as detailed in Chapter II. Although the polycystin–mucolipin domains of TRPML1 and TRPP2 include varying primary sequences (Figure 4), in general, they exhibit very similar tetrameric structures (Figure 5A). Notably, in TRPML1, the luminal pore loop projects downward toward the ion-selectivity filter, while this loop in TRPP2 is positioned upward in the vicinity of the luminal/extracellular environment (Figure 5C) (9). When the TRPML1 I-II linker was docked onto the full-length TRPP2 structure (Figure 5E), it appeared that a continuous α-helix was created from α1 of TRPML1 and S1 of TRPP2. In addition, both polycystin-mucolipin domains in TRPML1 and TRPP2 contain several N-linked glycosylation sites. These sites are present in TRPML1 at residues 159, 179, 220, and 230 (10), and in TRPP2 at residues 299, 305, 328, 362, and 375 (11).
Figure 4. Sequence alignment of the TRPML1 and TRPP2 polycystin-mucolipin domains.
The amino acid sequences of the TRPML1 and TRPP2 polycystin-mucolipin domains are illustrated. Secondary structural elements are indicated in orange and green corresponding to TRPML1 and TRPP2, respectively. Red residues highlight luminal pore-loop aspartates. Identical residues in both sequences are marked in bold black. Figure published in (9).
Figure 5. Structural comparisons of the polycystin–mucolipin domains. (A-C) Overlay of the tetrameric polycystin–mucolipin domain structures of human TRPML1 at pH 6.0 (orange) and human TRPP2 (green) illustrated from above (A) or parallel (B-C) to the membrane surface. The α1 helix is prominently displayed in (B) and (C). (C) Superimposed peripheral regions from the tetrameric structures in (B) indicating varying orientations of the luminal pore-loop. Structures viewed from the membrane surface contain a luminal pore-loop directed upward (in TRPP2) or downward (in TRPML1). (D) Overlay of the monomeric structures of the polycystin–mucolipin domain in TRPML1 and TRPP2, denoting the alignment of α-helices and β-sheets. (E) Docking of the TRPML1 polycystin–mucolipin domain (PDB code:5TJA) onto the TRPP2 structure (PDB code:5T4D). Ribbon representations highlight the side views parallel to the surface of the membrane. (A-D) were produced using PyMOL and reported by Li et al. (9) and (E) was generated with Chimera (12).
4.4 Structural features of TRPP2\textsubscript{CL}, TRPP2\textsubscript{MI}, and TRPP2\textsubscript{SI}

The structures of the closed conformational state of TRPP2, determined by Shen et al. (4) and Grieben et al. (6) will hereafter be referred to as TRPP2\textsubscript{CL}. The cryo-EM structure of the closed state is indicated in Figures 1 and 6. Additionally, the full-length human TRPP2 open state structures, solved by Wilkes et al. (6) will be regarded as TRPP2\textsubscript{SI} (single-ion state) and TRPP2\textsubscript{MI} (multi-ion state). As indicated by Figure 6, conformational differences exist between the TRPP2\textsubscript{CL} and TRPP2\textsubscript{SI} structures, primarily in S1-S2, S5-S6, and the polycystin–mucolipin domain (6).

**Figure 6. Comparison of the closed and open conformational states of TRPP2.** (A-B) Ribbon representations illustrating the superposition of the TRPP2\textsubscript{CL} (4) (pink) and TRPP2\textsubscript{SI} (6) (purple) tetrameric structures. Orientations of the structures are illustrated parallel to the membrane surface (A) as well as from the extracellular side of the membrane looking down the luminal pore (B). Cations positioned along the translocation pathway are specified by green spheres. Illustrative representations of the structures were generated using Chimera (12) with PDB codes: 5T4D (TRPP2\textsubscript{CL}) and 5MKE (TRPP2\textsubscript{SI}).
Moreover, both the open TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI} structures (6) are in complex with cations and are distinctly different from each other as well as from the closed state TRPP2\textsubscript{CL} structure, particularly in the S4-S5 linker (4–6). Notably, the structures of TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI} differ from one another in various sections including the polycystin–mucolipin domain, in a loop between S2 and S3, and in a linker between S4 and S5 (Figure 7) (6). The S4-S5 linker within TRPP2\textsubscript{SI} likely associates with \( \alpha 4 \) residues located in the EF-hand, consequently leading to a conformational change at the lower gates.
Figure 7. Cryo-EM structures of human TRPP2 in open conformational states. (A) Ribbon representations demonstrating the tetrameric structures of TRPP2\textsubscript{SI} (6) (A) and TRPP2\textsubscript{MI} (6) (B). The TRPP2\textsubscript{SI} structure is denoted in light blue and the TRPP2\textsubscript{MI} structure is indicated in orange. The structures are oriented parallel to the membrane surface. (C) Overlay of TRPP2\textsubscript{SI} and TRPP2\textsubscript{MI} structures. Cations located along the translocation pathway are illustrated by green spheres. Structural representations were created utilizing Chimera (12) with PDB codes 5MKE and 5MKF for TRPP2\textsubscript{SI} and TRPP2\textsubscript{MI}, respectively.
4.5 Characteristics of the ion-translocation pathway in TRPP2}_{\text{CL}}

In both the TRPP2}_{\text{CL}} structures, it is apparent that the two gates, present at the selectivity filter and the S6 region, are captured in a closed non-conductive state (Figures 1, 6) (4, 5). Primarily, cations present in the ER lumen or extracellular environment are able to flow into the upper vestibule of TRPP2 through the 9 Å wide opening at the center of the polycystin-mucolipin domain (5). Moreover, four lateral openings serve as additional entry points for cations (Figure 8A) (5). Furthermore, eight acidic residues from each monomer are present in this vestibule rendering this region highly negatively charged (Figure 8D) (4, 5). These negatively charged side chains likely aid in repelling anions while attracting and accumulating cations to improve ion conduction (Figure 8D) (4, 5).

The selectivity filter between pore helices 1 and 2, which encompasses Leu641, Gly642, and Asp643, is one gate that impedes the movement of hydrated cations in the closed state (Figures 8B, C) (4). Notably, Leu677 and Asn681 establish the narrowest region of the lower gate (Figures 8B, C), obstructing the conduction of hydrated cations (4). Interestingly, the upper gate at the selectivity filter as well as the lower gate are presumably allosterically regulated by a network of interactions (Figure 8E) (4). As indicated in Figure 8E, these associations are likely to be formed by residues of one monomer at the selectivity filter and pore helix 1 (Leu641 and Arg638, respectively) and residues from an adjacent monomer in pore helix 2 and the S6 helix (Phe661 and Val665, respectively) (4).
Figure 8. Configuration of the TRPP2 CL pore. (A) Views of sliced molecular surfaces colored by electrostatic potential from the membrane surface (left) and down the four-fold symmetry axis. Pore openings, vestibules, and gates are indicated by numbers 1-6. Yellow dashed lines highlight the openings leading to the pore (5). (B) Ion permeation pathway within the pore. Notable residues at the selectivity filter and lower gate are depicted as sticks. (C) Calculations of the pore radius. (D) Presence of negatively charged residues colored by surface electrostatic potential at the outer pore region (left). Negatively charged aspartate and glutamate residues are displayed as sticks (right). (E) Interactions coupling the selectivity filter and pore helix 1 of one monomer with S6 and pore helix 2 of an adjacent monomer. Figure adapted from Grieben et al. (A) (5) and Shen et al. (B-E) (4).
4.6 Ion-translocation pathways in TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI}

TRPP2\textsubscript{MI} contains a restricted entrance gate from the extracellular surface of the membrane at Thr265 and Lys267 within the Cat0 site (Figures 9C, 10A, 10C). In TRPP2\textsubscript{MI}, the entrance gate has a radius of 3.0 Å compared to the wider gates in TRPP2\textsubscript{CL} (4.4 Å) (5) and TRPP2\textsubscript{SI} (4.7 Å) (Figure 9C) (5, 6). Figures 10A and 10B further illustrate the differences between TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI} at the entrance gate. Due to its narrow entrance gate, TRPP2\textsubscript{MI} appears to function as a preselectivity filter, limiting the access of hydrated ions yet allowing dehydrated ions to enter the gate (6). In addition, TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI} are significantly broader than TRPP2\textsubscript{CL} at the selectivity filter and the lower gates (Figure 9C) (6). Intriguingly, a number of Ca\textsuperscript{2+} and Na\textsuperscript{+} ions exist in the ion transit pathways of TRPP2\textsubscript{MI} and TRPP2\textsubscript{SI} (Figures 9D, E and 10A, B) (6). Of note, both Ca\textsuperscript{2+} and Na\textsuperscript{+} were incorporated into the protein purification procedure.

With a pore radius of 1.7 Å at Leu641, TRPP2\textsubscript{SI} has the widest selectivity filter compared to that of TRPP2\textsubscript{MI} (1.4 Å) and TRPP2\textsubscript{CL} (1.0 Å) (Figure 9C). Wilkes et al. identified alterations in the conformation of Asp643, and its interaction with Arg638 is changed, impacting the arrangement of the pore radius (6). In TRPP2\textsubscript{MI} (Figure 10C), the carbonyl oxygens of Leu641 at the Ca2 site are capable of coordinating Ca\textsuperscript{2+} or Na\textsuperscript{+}, while in TRPP2\textsubscript{SI} (Figure 10D), the Cat2\textsuperscript{*} site can contain two cations (6). Further along the translocation pathway, Phe 669, located midway between the Cat3 and Cat4 sites, constricts the channel (Figure 10C) (6). Lastly, Asn681 and Ser685 associate with a density at the Cat5 site in TRPP2\textsubscript{MI} (Figure 10C) and TRPP2\textsubscript{SI} (Figure 10D) located at the channel exit (6).
Figure 9. Representation of the ion permeation pathways of TRPP2MI and TRPP2SI. (A-B) Solvent-accessible pathways along the pores in structures TRPP2MI (A) and TRPP2SI (B). Two diagonally opposed subunits are illustrated. A number of residues along the ion permeation pathway are depicted as sticks. (C) Radius of the pore in different regions along the translocation pathway in the open states, TRPP2MI and TRPP2SI (6), and closed state TRPP2CL (5). (D, E) Sliced molecular surface representation of TRPP2MI (D) and TRPP2SI (E). Left, viewed down the pore axis of the tetramer. Right, viewed parallel to the membrane. Representations are colored by electrostatic potential, red indicating highly negatively charged regions. Ca2+ (green spheres) and additional cations (blue spheres) are depicted. Yellow arrows denote exit pathways for ions. Figure adapted from (6).
Figure 10. Cation-binding sites along the ion permeation pore of TRPP2MI and TRPP2SI. (A-B) View looking down the four-fold symmetry axis of TRPP2MI (A) and TRPP2SI (B). Each monomer is illustrated in a separate color. (C-D) Movement of cations Ca1 and Cat2-5 in TRPP2MI (C) or Cat2* and Cat5* in TRPP2SI (D). Densities along the transit pathway are represented by black mesh. A number of pronounced side chains are indicated. Figure adapted from Wilkes et al. (6).
4.7 Role of glycosylation in TRPP2

As mentioned previously, TRPP2 contains five potential N-glycosylation sites ([11]) (Asn299, Asn305, Asn328, Asn362, and Asn375) situated in TLC2 or TLC3 within the polycystin–mucolipin domain. It has been found that glycosylation of this domain is key for trafficking and stability of TRPP2 ([11]). However, all three of the research groups attempted to reduce the glycosylation levels present in the protein during expression and purification prior to solving the TRPP2 structures ([4–6]). Yet, in the TRPP2 structure solved by Grieben et al. NAG, N-acetylglucosamine, components at three N-glycosylation sites (Asn328, Asn362, and Asn375) were determined (Figure 2B, Table 1) ([5]). Two additional N-glycosylation sites, Asn299 and Asn305, were unable to be resolved due to a disordered region, Ser298–Ala303 ([5]).

Furthermore, TRPP2\textsubscript{MI} contains four N-glycosylated residues (Asn299 in TLC2; Asn328, Asn362, and Asn375 in TLC3), while TRPP2\textsubscript{SI} exhibits three sites (Asn305, Asn362 and Asn375) that are N-glycosylated (Table 1) ([6]). However, it appears that the large glycan group at Asn305 in TRPP2\textsubscript{SI} would inhibit the intrasubunit interactions that exist in TRPP2\textsubscript{MI}. Notably, either Asn305 or Asn328 can be glycosylated at any given time. If one residue is glycosylated, the other site would not be available for glycosylation to occur ([6]). Intriguingly, all of the reported TRPP2\textsubscript{CL} structures have N-glycosylation sites at Asn328, Asn362, and Asn375, as observed in the TRPP2\textsubscript{MI} structure, with the exception of Asn299 ([4, 5]). Perhaps Asn299 is playing a critical role in the open state TRPP2\textsubscript{MI} structure. Additionally, Asn328 (which is not glycosylated in TRPP2\textsubscript{SI}) interacts with Glu343 ([6]) while Asn375 interacts with S3b-S4 ([5]).
Table 1. Glycosylated residues present in the published TRPP2 structures

<table>
<thead>
<tr>
<th>Construct</th>
<th>N-glycosylated sites in the structure</th>
<th>Proposed conformational state</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPP2:53–792</td>
<td>Asn328, Asn362, and Asn375</td>
<td>closed</td>
<td>Shen et al. (4)</td>
</tr>
<tr>
<td>TRPP2:185–723</td>
<td>Asn328, Asn362, and Asn375</td>
<td>closed</td>
<td>Grieben et al. (5)</td>
</tr>
<tr>
<td>TRPP2:1–968 full-length</td>
<td>Asn299, Asn328, Asn362, and Asn375</td>
<td>open (TRPP2\textsubscript{MI})</td>
<td>Wilkes et al. (6)</td>
</tr>
<tr>
<td>TRPP2:1–968 mustached</td>
<td>Asn305, Asn362, and Asn375</td>
<td>open (TRPP2\textsubscript{SI})</td>
<td>Wilkes et al. (6)</td>
</tr>
</tbody>
</table>

It remains to be seen whether N-glycosylation induces conformational changes in TRPP2\textsubscript{SI} and TRPP2\textsubscript{MI} or if the structures themselves direct the arrangement of N-glycosylation (6). Furthermore, due to glycosylation, the two structures differ with regard to the interactions between the polycystin-mucolipin domain in one subunit and the P loop of the adjacent subunit (Figure 11) (6). Notably, the variability in N-glycosylation present within TRPP2\textsubscript{SI} or TRPP2\textsubscript{MI} affect the conformational changes in β5 and α1 (Figure 11) (6). In TRPP2\textsubscript{MI}, Gln456 associates with Gln622 of pore helix 1 and Glu651 interacts with Trp455 in β5 (Figure 11A) (6). However, in TRPP2\textsubscript{SI}, Glu651 in pore helix 2 associates with the polycystin-mucolipin domain through Lys271, while Gln622 interacts with Ser254 within α1 (Figure 11B). Interestingly, pore helix 1 and pore helix 2 are arranged closer together in TRPP2\textsubscript{MI} (Figure 11A) than in TRPP2\textsubscript{SI} (Figure 11B) (6).
Figure 11. Presence of N-glycosylation within TRPP2MI and TRPP2SI. (A-B) Enlarged views of alterations in the structural conformations among β5 and α1 of the polycystin-mucolipin domain in one subunit and the P loop (pore helix 1 and 2) of the adjacent subunit. Detailed interactions are demonstrated within TRPP2MI (A) and TRPP2SI (B). Pore helix 1 and 2, part of the P loop between S5 and S6, are indicated in blue. β5 and α1 are highlighted in yellow. Notable residues are displayed as sticks and density maps are represented by mesh. Black lines illustrate interactions. Left, monomeric structure of TRPP2MI denoting the region featured in (A-B). Figure adapted from (6).
4.8 Discussion of ADPKD-associated missense mutations in TRPP2

As described in Chapter I, mutations in TRPP2 account for 15% of all autosomal dominant polycystic kidney disease (ADPKD) cases (13). Although 278 variations in TRPP2 have been identified in patients with ADPKD, according to the ADPKD database at the Mayo Clinic (http://pkdb.mayo.edu/), not all of these variants have been documented to be pathogenic (5). It is presumed that 26 germline missense mutations, situated within the main architecture of the channel, appear to play a role in the development of ADPKD (4, 5).

The structures of TRPP2 now provide us with a framework from which to consider the roles that these ADPKD-associated mutations play in channel formation (5). ADPKD missense mutations were mapped onto these structures and their relative locations are indicated in Table 2 and Figure 12. Notably, many ADPKD missense mutations create extensive interactions with one another due to their close proximity, especially within the polycystin-mucolipin domain. Intriguingly, the TRPP2 polycystin-mucolipin domain contains approximately 60% of the pathogenic missense mutations in this channel (6). The structural implications of various pathogenic mutations present in the polycystin-mucolipin, VSD, and PD domains will subsequently be elucidated.
Table 2. ADPKD-causing missense mutations in TRPP2

<table>
<thead>
<tr>
<th>ADPKD Mutation</th>
<th>Domain</th>
<th>Structural Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>W280R</td>
<td>polycystin-mucolipin</td>
<td>α2 helix</td>
</tr>
<tr>
<td>Y292C</td>
<td>polycystin-mucolipin</td>
<td>α2 helix</td>
</tr>
<tr>
<td>R306Q</td>
<td>polycystin-mucolipin</td>
<td>polycystin-mucolipin interface</td>
</tr>
<tr>
<td>L314V</td>
<td>polycystin-mucolipin</td>
<td>β1 strand</td>
</tr>
<tr>
<td>R322W</td>
<td>polycystin-mucolipin</td>
<td>β1 strand</td>
</tr>
<tr>
<td>R325P</td>
<td>polycystin-mucolipin</td>
<td>β1 strand</td>
</tr>
<tr>
<td>R325Q</td>
<td>polycystin-mucolipin</td>
<td>β1 strand</td>
</tr>
<tr>
<td>C331S</td>
<td>polycystin-mucolipin</td>
<td>TLC1 hairpin</td>
</tr>
<tr>
<td>Y345C</td>
<td>polycystin-mucolipin</td>
<td>TLC1 hairpin</td>
</tr>
<tr>
<td>S349P</td>
<td>polycystin-mucolipin</td>
<td>TLC1/2 hairpin interface</td>
</tr>
<tr>
<td>A356P</td>
<td>polycystin-mucolipin</td>
<td>TLC2 hairpin</td>
</tr>
<tr>
<td>A384P</td>
<td>polycystin-mucolipin</td>
<td>TLC3 hairpin</td>
</tr>
<tr>
<td>G390S</td>
<td>polycystin-mucolipin</td>
<td>β2 strand</td>
</tr>
<tr>
<td>W414G</td>
<td>polycystin-mucolipin</td>
<td>α3 helix</td>
</tr>
<tr>
<td>G418V</td>
<td>polycystin-mucolipin</td>
<td>α3- β3 loop</td>
</tr>
<tr>
<td>R420G</td>
<td>polycystin-mucolipin</td>
<td>β3 strand</td>
</tr>
<tr>
<td>R440S</td>
<td>polycystin-mucolipin</td>
<td>β4 strand</td>
</tr>
<tr>
<td>T448K</td>
<td>polycystin-mucolipin</td>
<td>β4- β5 loop</td>
</tr>
<tr>
<td>D511V</td>
<td>VSD (voltage-sensing domain)</td>
<td>S3 helix</td>
</tr>
<tr>
<td>L517R</td>
<td>VSD (voltage-sensing domain)</td>
<td>S3 helix</td>
</tr>
<tr>
<td>N580K</td>
<td>PD (pore domain)</td>
<td>S4-S5 loop</td>
</tr>
<tr>
<td>A615T</td>
<td>PD (pore domain)</td>
<td>S5</td>
</tr>
<tr>
<td>F629S</td>
<td>PD (pore domain)</td>
<td>Pore helix 1</td>
</tr>
<tr>
<td>C632R</td>
<td>PD (pore domain)</td>
<td>Pore helix 1</td>
</tr>
<tr>
<td>R638C</td>
<td>PD (pore domain)</td>
<td>Pore helix 1</td>
</tr>
<tr>
<td>L656W</td>
<td>PD (pore domain)</td>
<td>S6</td>
</tr>
</tbody>
</table>

Table modified from Grieben et al. (5)
Figure 12. Locations of pathogenic ADPKD missense mutations in TRPP2. (A) ADPKD-causing missense mutations, indicated in red, are mapped onto the monomeric TRPP2 structure. (B) Orientation of tetrameric TRPP2 viewed from above the polycystin-mucolipin domain. (C) Sliced pore domain/VSD view of the tetramer. (B-D) ADPKD-causing missense mutations are illustrated as circles of varying color denoting the structural implications of these mutants, ranging from pink (severe) to yellow (mild). (D) Cartoon and molecular surface representation of the monomeric TRPP2 structure mapped with ADPKD-causing missense mutations, depicted as circles. Missense mutations displayed on the surface of the protein are indicated. Figure adapted from Shen et al. (A) (4) and Grieben et al. (B-D) (5).
Proposed interactions of several ADPKD mutations in the polycystin-mucolipin domain

**T448K and C331S**

Thr448 is located in the β4-β5 loop (Figures 12A, B, D) and interacts with the α1 helix of an adjoining polycystin-mucolipin domain through a hydrogen bond with Tyr248 on the β3-β4 loop (5). Hence, the T448K ADPKD mutation would presumably perturb channel assembly (4). Cys331 resides in TLC1 (Figures 12A, B, D) and forms a disulfide bond with Cys344, which appears to restrict the β-hairpin into a rigid conformation (5). Notably, in TRPP2MI, a disulfide bridge is formed between Cys331 and Cys344 which provides stability to TLC3 (6). Interestingly, this disulfide bridge is not observed in TRPP2SI (6) or the TRPP2CL structure reported by Grieben et al. (5), but is discerned in the TRPP2CL structure determined by Shen et al. (4). Therefore, the ADPKD mutation C331S would likely disrupt this critical inter-subunit interface within the tetrameric polycystin-mucolipin domain.

**Gly418 interacts with Tyr345, Arg 325, and Arg 420**

Gly418 is situated in the α3-β3 loop (Figure 12A) and interacts with several residues that can harbor an ADPKD mutation, including Tyr345 (5). Tyr345 resides in the TLC1 β-hairpin at the junction of TLC1-3, situated near the β3 strand (Figures 12A, D) (5). However, when Gly418 is mutated, the sidechain of Tyr345 would most likely be altered (5). Gly418 also interacts with two additional residues, Arg325 and Arg420, which are pathogenic upon mutation. Arg325 is located in β1 (Figures 12A, B), adjacent to the TLC2 and the α3-β3 loop, which forms a π-stacking interaction with Phe358 (5). Of note, two separate missense mutations in Arg325, R325P and R325Q, can lead to ADPKD (5). Arg420 is situated in the β3 strand between the polycystin-
mucolipin subunits (Figure 12A). Collectively, Gly418 and its respective interactions with Arg420 and Tyr345 presumably aid in the stabilization of the junction between TLC1-3 and the adjacent polycystin-mucolipin subunits (5). Thus, ADPKD-causing mutations G418V, R420G, and Y345C would likely disrupt the assembly of the polycystin-mucolipin domain (5).

**Arg306 and Leu314**

Arg306 is situated at the interface of the polycystin-mucolipin subunits and associates with TLC1 of the adjacent subunit as well as the S3-S4 extension (5). Of note, Arg306 is located next to one of the glycosylation sites, Asn305 (5). Therefore, disruption of this region by the pathogenic R306Q mutation may have an effect on the glycosylation state of Asn305 (5). Additionally, Leu314 is situated in β1 at a junction between the central beta-sheet and the extension between S3-S4 (5). Since both Arg306 and Leu314 are positioned at interfaces within the polycystin-mucolipin domain (Figure 12D), mutations in these residues would likely lead to the destabilization of this region (5).

**Examples of proposed interactions of ADPKD missense mutations in the VSD**

**D511V and L517R**

D511V is a well-characterized and common disease-causing missense mutation in ADPKD patients (13, 14) which is shown to lead to channel-inactivation (15). Notably, the D511V pathogenic mutation is situated on S3 and its side chain expands into the space amid the S1–S4 helices (Figure 12C) (5). Therefore, the side chain position of Asp511 is key to forming a hydrogen-bonding network incorporating Lys572 and Lys575 on S4, in addition to Tyr487, Glu491, and Asn508 (4, 5). In TRPP2SI, Asp511 interacts with Lys575, while it associates with Lys572 in TRPP2MI once conformational changes occur in S4 (6). Therefore, this interaction
between Asp511 in S3 and Lys572 or Lys575 in S4 may contribute to the stabilization of the open conformation within the lower gates (6). Hence, this D511V mutation is predicted to alter the VSD structure through conformational changes (5, 6).

Leu517 is situated midway up S3, exhibiting a sidechain that packs against S4 and projects into the lipid bilayer (Figure 12C) (5). Therefore, the pathogenic L517R mutation would introduce a destabilizing charged residue within the center of the membrane (5).

**Examples of proposed interactions of ADPKD missense mutations in the pore domain**

**R638C and C632R**

As noted earlier, Arg638 is located on pore helix 1 and its side chain protrudes into the upper region of the pore (Figure 12C) (4, 5). At the selectivity filter, Asp643 appears to form a cation-π interaction with Phe646 (4, 5). The ADPKD mutation R638C alters the charge and presumably eliminates this interaction, disrupting the association between the pore helices of two adjacent subunits (4). Similarly, Cys632 also resides in pore helix 1, containing a side chain that is situated between pore helix 1 and S5 (5). Like R638C, C632R leads to a change in charge which likely disrupts the pore architecture (5).
4.9 Conclusion

This chapter has provided a detailed summary of the published TRPP2 channel structures, which reveal several interesting features. Like other TRP channels, TRPP2 forms a tetrameric structure and its polycystin-mucolipin domain is very similar to that seen in the TRPML subfamily. Comparison of TRPP2 in its closed (TRPP2_{CL}) and open (TRPP2_{MI} and TRPP2_{SI}) states illustrated not only notable structural differences but also variations with regard to their respective ion-permeation pathways and glycosylation patterns. Since these structures lack resolved N- and C-terminal domains, it remains to be seen how these termini affect the overall structure of the full-length channel. Nonetheless, these structures have opened an entirely new avenue for structural investigation of TRPP2.
4.10 References


Epilogue

The studies that we conducted on TRPML1 and TRPP2 provide significant insights into the structural characterization of these channels. The aim of this section is to summarize our findings as well as discuss future investigations that can be conducted on these proteins.

The TRPML1 I-II linker structures that we obtained were the first high-resolution crystal structures of a TRPML channel domain. Despite the changes in pH, the I-II linker structures at pH 4.5, 6.0, and 7.5 did not undergo significant conformational changes, indicating that these structures were in highly stable states. However, protein structures alone cannot demonstrate how proteins behave within the cell. Thus, functional studies enabled us to assess the dynamic qualities of these static TRPML1 I-II linker structures. Our findings demonstrate that the dual Ca\(^{2+}/pH\) regulatory mechanism of TRPML1 involves the complex interplay between the protonation of aspartates in the luminal pore and the blockade of Ca\(^{2+}\). Additionally, we characterized three MLIV-causing missense mutations, L106P, T232P, and C166F, which are present within the TRPML1 I-II linker. Our confocal imaging analyses as well as our structural and biochemical results comprehensively reveal that these missense mutations affect the structural integrity of the tetrameric I-II linker alone as well as within the greater context of the full-length TRPML1 channel.

Ultimately, our goal is to determine a high-resolution structure of full-length TRPML1 using cryo-EM. Our cryo-EM studies thus far have identified density for the extracellular domain, but not for the transmembrane segments. Therefore, further studies are needed in order to optimize our cryo-EM sample conditions. If indeed we are able to solve a full-length structure, it will be the first high-resolution cryo-EM structure of a channel in the TRPML subfamily.
An additional area of investigation that I focused on was to obtain a high-resolution cryo-EM structure of TRPP2. We have made great strides in optimizing TRPP2 protein purification and sample preparation procedures in order to pursue structural studies. As mentioned previously, many factors play a role in creating the proper conditions to generate protein particles of suitable quality for cryo-EM analysis. Every factor is a piece of the overall structural puzzle, and all of the pieces must align in order for the sample to yield a high-resolution structure. Using cryo-EM, we were able to obtain a very detailed view of human TRPP2 that exhibited the four-fold symmetrical axis, the central pore, as well as the S1-S4 helices within the tetramer. However, we had to overcome a significant preferred particle orientation dilemma. We presume that the preferred orientation of particles was due to a combination of factors including the presence of detergent, glycosylation of the protein, truncations, expression systems utilized, as well as grid preparation conditions. Although speculations can be made why our samples behaved differently compared to those of other groups, it is currently not possible to definitively conclude what exactly was impeding us from obtaining additional orientations. With the knowledge gained from the solved TRPP2 structures, we intend to try to resolve this dilemma in order to solve TRPP2 structures in different states.

It is important to keep in mind that the solved TRPP2 structures display only snapshots of the wild-type protein in a few states. Although the structural implications of ADPKD-causing missense mutations have been proposed based on the TRPP2 structures, additional studies are needed in order to determine whether these hypotheses are indeed correct. Thus, we intend to investigate various TRPP2 proteins harboring ADPKD-causing missense mutations using FSEC to ascertain if any of these proteins look promising for structural analysis. However, one limitation of structurally investigating these ADPKD-associated missense mutations is that the proteins may
misfold and aggregate due to inherent instability. Nonetheless, pathogenic missense mutations that do not extensively disrupt the structural integrity of TRPP2 can be further examined.

Moreover, we plan to characterize these ADPKD-causing missense mutations by determining their subcellular localization patterns using confocal fluorescence microscopy. Since TRPP2 normally localizes to the ER, it would be interesting to determine if pathogenic missense mutations result in mislocalization of the channels. Together with our structural findings, these studies would enable us to gain a better understanding of the mechanism of disease of ADPKD.

It is apparent that the wild-type TRPP2 solved structures are only part of the overall story of TRPP2 and the role that it plays in ADPKD. Taken together, although progress has been made in the TRPP2 field, many uncertainties and controversies exist, from the biophysical properties of TRPP2 to the role this protein plays in the pathogenesis of ADPKD. Therefore, further studies are necessary in order to better understand the function of this channel. Significantly, the structural determination of the PKD1/TRPP2 complex would provide valuable information that would significantly improve our knowledge of ADPKD. Due to continuous advancements being made in the cryo-EM field, this complex will undoubtedly be determined in the future. Hence, TRPP2 and PKD1 structures can be potential targets for the development of new therapeutic drugs for the treatment of ADPKD.
Appendix A:

Structural basis of dual Ca$^{2+}$/pH regulation of the endolysosomal TRPML1 channel
Structural basis of dual Ca\(^{2+}/\text{pH}\) regulation of the endolysosomal TRPML1 channel

Minghui Li\(^{1,6}\), Wei K Zhang\(^{2,3,6}\), Nicole M Benvin\(^{1,6}\), Xiaoyuan Zhou\(^4\), Deyuan Su\(^{2,5}\), Huan Li\(^{2,5}\), Shu Wang\(^{2,5}\), Ioannis E Michailidis\(^1\), Liang Tong\(^1\), Xueming Li\(^4\) & Jian Yang\(^{1,2}\)

The activities of organellar ion channels are often regulated by Ca\(^{2+}\) and H\(^+\), which are present in high concentrations in many organelles. Here we report a structural element critical for dual Ca\(^{2+}/\text{pH}\) regulation of TRPML1, a Ca\(^{2+}\)-release channel crucial for endolysosomal function. TRPML1 mutations cause mucolipidosis type IV (MLIV), a severe lysosomal storage disorder characterized by neurodegeneration, mental retardation and blindness. We obtained crystal structures of the 213-residue luminal domain of human TRPML1 containing three missense MLIV-causing mutations. This domain forms a tetramer with a highly electronegative central pore formed by a novel luminal pore loop. Cysteine cross-linking and cryo-EM analyses confirmed that this architecture is fundamental to TRPML1’s regulation, assembly and pathogenesis.

Many types of ion channels are present in the membranes of intracellular organelles, such as the endoplasmic reticulum, Golgi apparatus, mitochondria, endosomes and lysosomes\(^{1-6}\). These channels help to establish and maintain ionic-concentration gradients across organelar membranes, and/or serve as release channels for key signaling ions, including Ca\(^{2+}\) and H\(^+\). Many organelles have high concentrations of Ca\(^{2+}\) and H\(^+\) in their lumens\(^{1,3-5,7}\). These ions in turn regulate the activities of numerous organellar ion channels\(^{8-12}\). Elucidating the molecular mechanisms underlying this universal regulation is instrumental in understanding the physiological and pathophysiological functions of organellar ion channels.

TRPML1 is a member of the transient receptor potential mucolipin (TRPML) channel subfamily, which was first identified as a genetic determinant of mucolipidosis type IV (MLIV)\(^{13-15}\), one of ~50 lysosomal storage disorders. Young children with MLIV often show cognitive, linguistic and motor deficits and are sometimes blind\(^{6,16-18}\). TRPML1 is located primarily in lysosomes, where its main role is to conduct Ca\(^{2+}\) from the lysosome lumen to the cytoplasm\(^{5,6,19-23}\). This Ca\(^{2+}\)-release step is important in many lysosome-dependent cellular events, including exocytosis\(^{5,6,19,23}\), membrane trafficking\(^{5,6,25-27}\) and autophagy\(^{5,6,24,25,28-30}\). TRPML1 activity is regulated by luminal and extracellular Ca\(^{2+}\) and H\(^+\) (refs. 6,19–23,31,32). In the presence of divalent ions, TRPML1 currents are greatly potentiated by H\(^+\), via unknown mechanisms\(^{6,19-23,31,32}\). The dual regulation by Ca\(^{2+}\) and H\(^+\) has clear physiological relevance: TRPML1 exists primarily in lysosomes, and during biogenesis and lysosomal exocytosis, it may also be present in other organelles, including late endosomes and the plasma membrane\(^{23-25,33}\). These distinct subcellular compartments have different pH levels and Ca\(^{2+}\) concentrations\(^{1,5,6,23}\). The extracellular side of the plasma membrane has 1.8–2 mM Ca\(^{2+}\) and a near-neutral pH of 7.2–7.4, whereas late endosomes have 0.5 mM Ca\(^{2+}\) and an acidic pH of 5.5–6.0. Lysosomes have 0.5–0.6 mM Ca\(^{2+}\) and an even more acidic pH of 4.5–5.0, a condition crucial for the activity of their native hydrolases\(^{23,34}\). Thus, depending on its subcellular location, TRPML1 may be regulated to different extents by Ca\(^{2+}\) and pH and may consequently exhibit different activities. The molecular mechanism of this dual regulation is unknown.

The crucial role of TRPML1 in cellular processes is further demonstrated by the existence of >20 mutations in the TRPML1 gene that have been linked to MLIV\(^{6,18,35}\). Among these mutations, some (such as frameshift or nonsense mutations) completely abolish TRPML1 expression, but many (>12) are single–amino acid missense mutations\(^{6}\). How these missense mutations alter the structure and function of TRPML1 is largely unclear.

We used a structural approach to investigate the molecular mechanisms of dual Ca\(^{2+}/\text{pH}\) regulation of TRPML1 and MLIV pathogenesis, focusing on the linker between the first two transmembrane segments (S1 and S2) of TRPML1, which we refer to as the I–II linker. Like other TRP-channel subunits, TRPML1 contains six transmembrane segments (S1–S6; Supplementary Fig. 1a). The I–II linker accounts for

\(^{1}\)Department of Biological Sciences, Columbia University, New York, New York, USA. \(^{2}\)Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences/Key Laboratory of Bioactive Peptides of Yunnan Province; and Ion Channel Research and Drug Development Center, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China. \(^{3}\)Department of Pharmacology, School of Pharmaceutical Sciences, South-Central University for Nationalities, Wuhan, China. \(^{4}\)Beijing Advanced Innovation Center for Structural Biology, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China. \(^{5}\)Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China. \(^{6}\)These authors contributed equally to this work. Correspondence should be addressed to J.Y. (jy160@columbia.edu).

Received 28 September 2016; accepted 12 December 2016; published online 23 January 2017; doi:10.1038/nsmb.3362
more than one-third of the channel's length (Supplementary Fig. 1b). In lysosomes and endosomes, this linker faces the lumen (and is therefore also referred to as the ‘luminal linker’); on the plasma membrane, it faces the extracellular solution (Supplementary Fig. 1a). Moreover, it is the site of three single–amino acid missense mutations that cause MLIV5,17,18,36,37 (Supplementary Fig. 1b). We solved the crystal structures of the luminal linker at three different pH conditions (corresponding to the pH in lysosomes, endosomes and the extracellular milieu) and elucidated its role in the dual regulation of TRPML1 by luminal Ca$^{2+}$ and pH, TRPML1-channel assembly and MLIV pathogenesis.

**RESULTS**

**TRPML1-channel activity is regulated by luminal Ca$^{2+}$ and pH**

To obtain human TRPML1 currents, we followed a well-established method by expressing a constitutively active mutant channel designated TRPML1$^{VP}$ (tagged with EGFP on its N terminus) in HEK 293T cells20,38. TRPML1$^{VP}$ contains the V432P mutation, which mimics a spontaneous gain-of-function mutation (A419P) in mouse TRPML3 that causes the varaint-waddler (Va) phenotype39–41. The V432P mutation enables TRPML1 to traffic to the plasma membrane and to become spontaneously active20,38. In whole-cell patch-clamp recordings, TRPML1$^{VP}$ produced a strong inward-rectifying current in a nominal divalent-ion-free (NDF) extracellular/luminal solution at pH 7.4 (Fig. 1a,b). Addition of Ca$^{2+}$ to the extracellular/luminal solution inhibited the current in a dose-dependent manner. A dose–response curve yielded an apparent half-maximal inhibitory concentration (IC$_{50}$) of 0.27 mM, with a Hill coefficient of 1 (Fig. 1c), thus suggesting a one-to-one blocking mechanism.

Ca$^{2+}$ inhibition of TRPML1 was strongly dependent on extracellular/luminal pH, and became much weaker when the pH of the extracellular/luminal solution was lowered from 7.4 to 4.6, the pH of the lysosome lumen. The dose–response curve was shifted to the right at pH 4.6 (Fig. 1c), thereby increasing the apparent IC$_{50}$ by 14-fold, to 3.8 mM. Intriguingly, the shape of the dose–response curve was different, and the Hill coefficient was changed from 1 to 0.5 (Fig. 1c), thus suggesting negative cooperativity between extracellular/luminal Ca$^{2+}$ and H$^{+}$. Owing to attenuated Ca$^{2+}$ inhibition at pH 4.6, TRPML1$^{VP}$ currents were augmented when the extracellular/luminal pH was lowered from 7.4 to 4.6 (Fig. 1d).

It should be mentioned that the dual regulation by luminal Ca$^{2+}$ and pH described above was observed in a gain-of-function mutant rather than the wild-type (WT) channel. Because the V432P mutation enables the channel to open constitutively, any effects of luminal Ca$^{2+}$ and H$^{+}$ on gating might not be revealed.

**Crystal structure of the luminal linker**

Because the TRPML1 luminal linker faces the endolysosomal lumen and constitutes a large part (>1/3) of the channel, we investigated its role in the Ca$^{2+}$/pH regulation of TRPML1. We first determined the X-ray crystal structure of the luminal linker (G84–S296) of human TRPML1 at a 2.3-Å resolution at pH 6.0 (Fig. 2 and Table 1). The structure shows that the linker forms a tetramer with four-fold rotational symmetry (Fig. 2a). The structure is ~90 Å wide (outer diameter) and ~30 Å tall. At the center of the tetramer exists a highly electronegative pore, termed the luminal pore (Fig. 2b). Each protomer has a structural fold consisting of two long α-helices (α1 and α2, with 22 and 17 residues, respectively), two short α-helices (α3 and α4), eight β-strands and three loops with >15 residues (between α1 and α2, α3 and β1, and β4 and β5, respectively) (Fig. 2c). The four α-helices are packed tightly against a five-stranded β-sheet formed by β1, β4, β5, β6 and β7. One side of this β-sheet faces the luminal pore, and the other side faces the outside of the tetramer (Fig. 2a,c). Two intrasubunit disulfide bonds are present, one between C166 in β1 and C192 in β4, and another between C253 in β6 and C284 in the short loop between β7 and β8. The structure of E199–K219 (which constitutes part of the loop between β4 and β5) was unresolved, presumably because of its flexibility.

The TRPML1 I–II linker structure has a number of unique and notable features. (i) The luminal pore is lined by a novel 16–amino acid, reentrant luminal pore loop (L106–T121) that connects to helix α2. (ii) Each protomer contains a large disulfide loop (between α1 and α2), a short loop (between α3 and β1), and a short loop (between β4 and β5). (iii) The first 22 amino acids of the luminal linker (G84–F105) form an α–helix (α1) that may be of strategic importance for allosteric modulation: it connects directly to S1 at the N terminus and to the luminal pore loop at the C terminus (Fig. 2a). When the TRPML1 I–II linker is placed on top of the transmembrane segments of the TRPV1 structure42, α1 aligns reasonably well with S1 (Supplementary Fig. 2). Indeed, α1 appears to form a continuous α-helix with S1, thereby anchoring the I–II linker to the rest of the channel. We speculate that in the full-length channel, sequential movements of cytoplasmic and transmembrane regions may propagate to the luminal pore through α1.

Figure 1 Dual regulation of TRPML1 by Ca$^{2+}$ and pH. (a) Families of TRPML1$^{VP}$ currents at pH 7.4 and the indicated concentrations of extracellular Ca$^{2+}$. (b) Current–voltage (I–V) relationship of the currents in a at the indicated Ca$^{2+}$ concentrations. NDF, nominal divalent-cation free. (c) Dose–response relationships of Ca$^{2+}$ inhibition of TRPML1$^{VP}$ at pH 7.4 and 4.6, at a potential of −80 mV. Error bars, s.e.m. Numbers of recordings (n) are shown in parentheses. Solid curves are fits to the Hill equation. (d) Time course of TRPML1$^{VP}$ currents at the indicated pH, with 1 mM Ca$^{2+}$ and either Na$^{+}$ or NMDG$^{+}$ as the charge carrier. Source data for experiments in a, c and d are available in Supplementary Data Set 2.
Another α-helix (α2, R122–A138), whose N terminus is connected to the luminal pore loop, forms a flat perimeter on the top of the luminal pore (Fig. 2a).

Confirming the luminal-linker structure in full-length TRPML1

The structure described above was obtained from an isolated channel domain. Is this tetrameric structure preserved in the full-length channel? We used two approaches to address this question. First, the crystal structure shows that each luminal linker interacts with two neighboring linkers (Fig. 2a); therefore, we performed biochemical experiments to examine cross-linking among them. In this experiment, we used Caenorhabditis elegans instead of human TRPML1 because, for unknown reasons, the human TRPML1 protein remained tetrameric even in reducing SDS–PAGE (data not shown). On the basis of the intersubunit interface of the human TRPML1 I–II linker, we engineered three complementary pairs of cysteines in full-length TRPML1:

\[
\begin{align*}
D115 &\quad D114 \\
A113 &\quad A112 \\
D111 &\quad S110 \\
Y109 &\quad L106 \\
L107 &\quad T116 \\
F117 &\quad A118 \\
A119 &\quad Y120 \\
T121 &\quad T122
\end{align*}
\]

Figure 2 Crystal structure of the TRPML1 I–II linker. (a,b) Ribbon representation (a) and electrostatic-potential surface representation (b) of the I–II linker structure at pH 6.0. Top panels, top-down views from the extracellular/luminal side of the membrane. Bottom panels, side views parallel to the membrane. (c) Stereo top-down view of a protomer. (d) Stereo side view of the luminal pore loop. The front subunit was removed for clarity.

Table 1 Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>TRPML1 I–II linker, pH 6.0 (PDB 5TJA)</th>
<th>TRPML1 I–II linker, pH 6.0, K2Pt(CN)4 derivative (PDB 5TJB)</th>
<th>TRPML1 I–II linker, pH 4.5 (PDB 5TJC)</th>
<th>TRPML1 I–II linker, pH 7.5 (PDB 5TJD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td>I4 22</td>
<td>I422</td>
<td>P421 2</td>
<td>F432</td>
</tr>
<tr>
<td>Space group</td>
<td>I4 22</td>
<td>I422</td>
<td>P421 2</td>
<td>F432</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>125.3, 125.3, 76.7</td>
<td>125.3, 125.3, 76.5</td>
<td>94.4, 94.4, 50.7</td>
<td>182.8, 182.8, 182.8</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.5418</td>
<td>1.5418</td>
<td>1.5418</td>
<td>1.0750</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.0–2.3 (2.38–2.30)</td>
<td>20.0–2.6 (2.69–2.60)</td>
<td>20.0–2.4 (2.49–2.40)</td>
<td>20.0–2.4 (2.49–2.40)</td>
</tr>
<tr>
<td>Rwp (%)</td>
<td>9.9 (37.7)</td>
<td>13.8 (64.4)</td>
<td>6.8 (43.4)</td>
<td>5.9 (34.0)</td>
</tr>
<tr>
<td>Rwork (%)</td>
<td>20.0 (4.6)</td>
<td>14.6 (2.8)</td>
<td>19.0 (2.8)</td>
<td>22.3 (3.7)</td>
</tr>
<tr>
<td>CR1/2</td>
<td>0.982 (0.944)</td>
<td>0.961 (0.866)</td>
<td>0.968 (0.872)</td>
<td>0.969 (0.882)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.0 (99.6)</td>
<td>99.9 (99.8)</td>
<td>99.3 (98.3)</td>
<td>98.0 (98.6)</td>
</tr>
<tr>
<td>Redundancy (%)</td>
<td>7.2 (6.1)</td>
<td>6.9 (6.4)</td>
<td>4.7 (3.8)</td>
<td>4.2 (4.1)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.0–2.3</td>
<td>20.0–2.4</td>
<td>20.0–2.4</td>
<td>20.0–2.4</td>
</tr>
<tr>
<td>No. reflections</td>
<td>13,298 (1,337)</td>
<td>8,777 (927)</td>
<td>9,088 (1,033)</td>
<td>9,088 (1,033)</td>
</tr>
<tr>
<td>Rwork / Rfree</td>
<td>0.228 / 0.253</td>
<td>0.212 / 0.252</td>
<td>0.216 / 0.253</td>
<td>0.216 / 0.253</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>1,518</td>
<td>1,268</td>
<td>1,426</td>
<td>1,426</td>
</tr>
<tr>
<td>Water</td>
<td>111</td>
<td>104</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td>B factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>33.0</td>
<td>42.1</td>
<td>50.3</td>
<td>47.6</td>
</tr>
<tr>
<td>Water</td>
<td>37.2</td>
<td>45.5</td>
<td>47.6</td>
<td>47.6</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.010</td>
<td>0.007</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.7</td>
<td>1.5</td>
<td>1.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Each data set was collected from a single crystal. Values in parentheses are for the highest-resolution shell.
C. elegans TRPML1 that were predicted to form intersubunit disulfide bonds; these included D104C in linker A and T293C in linker B; H111C in linker A and Y136C in linker B; and D158C in linker A and D137C in linker B (Fig. 3a). We also engineered six noncomplementary pairs of cysteines in two neighboring linkers that were predicted not to form intersubunit disulfide bonds: D104C and Y136C; D104C and D137C; H111C and T293C; H111C and D137C; D158C and T293C; and D158C and Y136C. The results were as predicted: under nonreducing conditions (i.e., without DTT), purified full-length TRPML1 proteins containing complementary cysteine mutations formed oligomeric bands in SDS–PAGE at the expense of the monomeric bands (Fig. 3b, top, lanes 2, 6 and 10; uncropped images shown in Supplementary Data Set 1). However, those proteins containing noncomplementary cysteine mutations remained monomeric (Fig. 3b, top, lanes 1, 3, 4, 5, 7 and 8). (The D158C Y136C protein in lane 9 appeared to have been degraded or precipitated and did not yield a band.) Under reducing conditions (i.e., with DTT), all proteins existed as monomers (Fig. 3b, bottom). The existence of the oligomer bands only in the complementary cysteine pairs and only under nonreducing conditions was consistent with disulfide-bond formation between the engineered cysteines, even though the molecular weight of the oligomer bands was larger than that of a predicted tetramer, possibly because of tetramer aggregation.

Second, we carried out single-particle cryo-EM analysis of the purified full-length WT C. elegans TRPML1 protein (Supplementary Fig. 3). The cryo-EM density of the extracellular domain (ED) was clear and robust, but that of the transmembrane domain (TMD) was unclear and heterogeneous (Supplementary Fig. 3cd). The ED and TMD were distinguishable and recognizable on the basis of their dimensions and the gross similarity of the TRPML1 and TRPV1 TMDs. We therefore used a masking procedure that allowed us to segregate the ED from the TMD during subsequent cryo-EM analysis (Supplementary Fig. 3c). A cryo-EM density map of the ED was obtained at an overall resolution of 5.28 Å with C₄ symmetry imposed (Fig. 3c and Supplementary Fig. 3c). Even at this limited resolution, the ED, ~79% of which is made up of the luminal linker, is visible as a tetramer with a large pore in the center (Fig. 3c), thus confirming two of the most salient hallmarks of the crystal structure of the luminal linker. Given the limited resolution, we did not build an atomic model or identify secondary structures on this density map; however, the crystal structure of the luminal linker could be fitted with a correlation coefficient of 0.75, thus indicating a reasonably good fit (Fig. 3d). The extra density not covered by the luminal-linker crystal structure in the composite may come from the E199–K219 loop (which is unresolved in the crystal structure), other extracellular loops and/or glycosylation.

Together, the above biochemical and cryo-EM studies suggest that the TRPML1 luminal linker has the same structure in the full-length channel as in the isolated domain: it forms a tetramer and contains a large central pore. This conclusion was further supported by structure-guided mutagenesis studies described directly below.

The luminal-linker tetramer is crucial in TRPML1 assembly

The crystal structure of the luminal linker shows that it forms a tight tetramer through extensive intersubunit interactions. These interactions include hydrogen-bonding, formation of ion pairs, hydrophobic interactions and formation of an antiparallel, domain-swapped, three-stranded β-sheet between two neighboring subunits (Fig. 4a and Supplementary Fig. 1b). The interface between two adjacent linkers has a buried surface area of 1,430 Å². To examine whether the tetrameric assembly of the luminal linker is important in the tetrameric assembly of full-length TRPML1, we mutated L144 and R146, two residues that engage in intersubunit interactions (Fig. 4a), to lysine and serine, respectively, in both the isolated luminal linker and full-length human TRPML1. We then examined the oligomeric states of WT and mutant linker domains or full-length proteins by native gel electrophoresis. In a nondenaturing gel, the isolated luminal linker containing the L144K R146S mutation (Fig. 4b, top, lane 2) migrated at a much lower molecular weight than did the WT linker (Fig. 4b, top, lane 1), thus suggesting their being a monomer and a tetramer, respectively, and indicating that the double mutation disrupted the tetrameric assembly of the isolated linker.
In a blue native gel, full-length WT TRPML1 showed a distinct band, which presumably corresponded to tetramers (Fig. 4b, bottom, lane 1). This band was absent for mutant full-length TRPML1, which migrated as aggregated oligomers (Fig. 4b, bottom, lane 2), thus suggesting that the L144K R146S mutation disrupted the tetrameric assembly of full-length TRPML1.

In agreement with the above biochemical observations, a mutant TRPML1VP channel containing the L144K R146S mutation (termed TRPML1VP-LR) produced low whole-cell currents when it was expressed in HEK 293T cells, in an NDF solution at either pH 7.4 or 4.6 (Fig. 4c). We then examined the subcellular localization of TRPML1-LR, containing the L144K R146S mutation, by using confocal fluorescence microscopy. The WT and mutant channels tagged with EGFP were expressed in live HeLa cells, which were labeled with the lysosomal marker LysoTracker to identify lysosomes. WT TRPML1 had a punctate distribution and localized to lysosomes, as previously observed43–46; TRPML1-LR, however, had a diffuse distribution throughout the cell and did not localize to lysosomes (Fig. 4d). Altogether, these results indicate that proper formation of the luminal-linker tetramer is critical for the proper assembly and proper subcellular organelle targeting of TRPML1.

**MLIV mutations in the luminal linker disrupt TRPML1 assembly and localization**

Three MLIV-causing single–amino acid missense mutations (L106P, C166F and T232P) are present in the TRPML1 luminal linker (Fig. 5a). These mutations produce milder MLIV symptoms than do truncation or frame-shift mutations18,47,48, which generally result in a complete loss of the TRPML1 protein. The pathogenic mechanisms of these missense mutations at the channel level are unknown. The crystal structure of the luminal linker provides a structural context to investigate how these mutations affect TRPML1 structure, assembly, localization and activity.

In the structure, L106 is at the junction of the anchoring α-helix (e1) and the luminal pore loop, C166 forms a disulfide bond with C192, and T232 is located on β5 of a core β-sheet (Fig. 5a and Supplementary Figs. 1b and 4a). The luminal-linker protein containing any one of the three pathogenic mutations was expressed but with low yields. CD measurements of the purified proteins revealed marked changes in the secondary structures of all three mutant linkers (Fig. 5b). In a blue native gel, the full-length WT protein migrated as a distinct putative tetramer band (Fig. 5c, lane 1), whereas the full-length mutant proteins migrated as aggregated oligomers (Fig. 5c, lanes 2, 3 and 4). Size-exclusion chromatography further confirmed increased aggregation of the mutant proteins (Fig. 5d). These results indicated that all three mutations disrupt the structure of the luminal pore and the tetrameric assembly of the full-length channel. These detrimental effects were expected, because two of the mutations (L106P and T232P) introduce a proline, and one of the mutations (C166F) disrupts an intrasubunit disulfide bond and replaces a small amino acid with a larger hydrophobic one.

In agreement with the biochemical characterizations, confocal imaging showed that the mutant channels expressed in HeLa cells were diffusely distributed in the cells and did not localize to lysosomes, where the WT channels were primarily found (Fig. 5e and Supplementary Fig. 4b). As expected, because of abnormal assembly, the mutant channels produced little or no current in HEK 293T cells when the mutations were introduced in TRPML1VP (Fig. 5f).

**The luminal pore is critical for Ca^2+ and pH regulation of TRPML1**

The existence of a luminal pore is a distinct feature of TRPML1. The manual docking model (Supplementary Fig. 2) suggests that the luminal pore is a part of the ion-conduction pathway. To test this hypothesis, we simultaneously mutated three amino acids with relatively small side chains (S110, G112 and A113) in the luminal pore loop to cysteine in TRPML1VP, and examined the effect of MTSET
(a sulfhydryl-specific modifying reagent with a head group 5.8 Å wide) on ion conduction in the mutant channel (termed TRPML1VP-3C). The whole-cell current of TRPML1VP did not change significantly after extracellular application of 5 mM MTSET, but the current of TRPML1VP-3C was quickly and irreversibly inhibited (Supplementary Fig. 5b), thus supporting the hypothesis above. In agreement with the structural revelation of a wide luminal pore, the current reduction after MTSET modification was only partial (Supplementary Fig. 5b). Further supporting the notion of a wide luminal pore and the specificity of thiold modification, the currents of TRPML1VP channels carrying a S110C G112C or a G112C A113C double mutation were not inhibited by 5 mM extracellular MTSET (data not shown).

The presence of 12 aspartate residues in proximity within the luminal pore (Fig. 2a,d) strongly suggests that this pore region plays a role in Ca2+ and pH regulation. We tested this possibility by simultaneously mutating all 12 aspartate residues (D111, D114 and D115 in each subunit) to glutamine in TRPML1VP, thus generating a mutant channel termed TRPML1VP-3DQ. These mutations greatly decreased Ca2+ inhibition at pH 7.4, increasing the apparent IC50 from 3.8 mM to >10 mM (Fig. 6a). Decreasing the pH from 7.4 to 4.6 attenuated Ca2+ inhibition of TRPML1VP (Fig. 1c). The aspartate-glutamine mutations further decreased Ca2+ inhibition at pH 4.6, thereby increasing the apparent IC50 from 3.8 mM to >10 mM (Fig. 6b). These results indicated that the luminal-pore aspartate residues are important for Ca2+ and pH regulation of TRPML1.

The high electronegativity of the luminal pore is also expected to enhance Ca2+ conductance by attracting and concentrating Ca2+. However, the effect on Ca2+ conductance is likely to be complex and pH dependent. On the one hand, the luminal-pore aspartates attract and concentrate Ca2+ and hence increase Ca2+ conductance; on the other hand, they bind Ca2+, thereby blocking the luminal pore and decreasing Ca2+ conductance. Thus, the end result on Ca2+ conductance is a balance between these two opposite effects, depending on the protonation state of the luminal-pore aspartates.

To examine the importance of the luminal pore on Ca2+ conduction, we used Ca2+ imaging to measure Ca2+ influx into HEK 293T cells expressing TRPML1VP after switching from an extracellular solution without Ca2+ to one with 3 mM Ca2+. We chose this concentration to achieve a balance between a sufficiently strong Ca2+ signal and physiological relevance. Intracellular Ca2+ increased gradually over time after the concentration was switched from 0 to 3 mM Ca2+ (Fig. 6c). The slow rise time was not due to a slow solution exchange, which was complete in 100–300 ms, but it might have been due to a slow accumulation of Ca2+ inside the cells. Nonetheless, the rise time was significantly faster at pH 4.6 than at pH 7.4, thus presumably reflecting a faster Ca2+ conductance at pH 4.6. This outcome was probably mostly due to a weaker inhibition of the channel by Ca2+, even though the Ca2+-concentrating effect was also attenuated at this acidic pH.

We next examined the effect of the aspartate-glutamine mutations on Ca2+ conductance. These mutations are expected to produce two opposite effects: (i) less Ca2+ inhibition and hence increased Ca2+ conductance and (ii) less of a Ca2+-concentrating effect and hence decreased Ca2+ conductance. The initial rise time of the Ca2+ increase was similar at pH 4.6 and 7.4 for TRPML1VP-3DQ, in agreement with the expected effect of the aspartate-glutamine substitutions (Fig. 6d). At pH 7.4, the time course of Ca2+ increase of TRPML1VP-3DQ was similar to that of TRPML1VP (Fig. 6e), which was strongly inhibited by Ca2+ at that pH. However, at pH 4.6, the aspartate-glutamine mutations significantly slowed the time course of Ca2+ increase (Fig. 6f). This result can probably be attributed to a markedly decreased Ca2+-concentrating effect in the mutant channel. Altogether, these results suggest that the luminal-pore aspartate residues play a role in Ca2+ conductance.
Conduction through TRPML1. However, the TRPML1VP-3DQ channel currents clearly still exhibited strong inward rectification, thus indicating that this property is not controlled by the luminal-pore aspartate residues (Supplementary Fig. 6).

There are two possibilities for how acidic pH attenuates Ca\(^{2+}\) inhibition of TRPML1. First, low pH changes the structure of the luminal pore, either globally or locally. Second, low pH simply protonates the luminal-pore aspartates without causing substantial structural changes.

To distinguish between these possibilities, we obtained crystal structures of the TRPML1 luminal linker at pH 4.5, mimicking its native physiological state, and at pH 7.5, mimicking the pH at the plasma membrane. Both structures were determined at a 2.4-Å resolution (Fig. 7a and Table 1). Because residues E199–K219 were unresolved in the pH 6.0 structure, we removed R200–E213 in the constructs used.

The luminal-pore structure is not substantially altered by pH

There are two possibilities for how acidic pH attenuates Ca\(^{2+}\) inhibition of TRPML1. First, low pH changes the structure of the luminal pore, either globally or locally. Second, low pH simply protonates the luminal-pore aspartates without causing substantial structural changes.

To distinguish between these possibilities, we obtained crystal structures of the TRPML1 luminal linker at pH 4.5, mimicking its native physiological state, and at pH 7.5, mimicking the pH at the plasma membrane. Both structures were determined at a 2.4-Å resolution (Fig. 7a and Table 1). Because residues E199–K219 were unresolved in the pH 6.0 structure, we removed R200–E213 in the constructs used.
in the pH 4.5 and pH 7.5 structures. The different structures show some differences in the periphery; for example, at pH 4.5, the first nine amino acids of α1 became unstructured, and several β-strands (β2, β3 and β8') could not be resolved. These differences might have been caused by different crystal packing in different pH conditions.

Remarkably, the structures are almost identical in most regions, including the luminal pore loop (Fig. 7a,b). Even the side chain orientation of the luminal-pore aspartates is virtually the same (Fig. 7b), although the crystals at different pH are in completely different crystal forms (Table 1). These structures indicate that the I–II linker exists in a highly stable state. Thus, the attenuated Ca\(^{2+}\) inhibition at the acidic pH (Fig. 1) is probably due to protonation of the luminal-pore aspartates rather than to large conformational changes in the luminal pore.

**DISCUSSION**

The long luminal/extracellular linker between the first two transmembrane segments is a unique but shared feature of the TRPML and TRPP subfamilies of TRP channels. Our crystal structure of the TRPML1 I–II linker is, to our knowledge, the first high-resolution structure of a TRPML channel. Because the I–II linkers of TRPML2 and TRPML3 share high amino acid sequence homology with TRPML1 (Supplementary Fig. 1b), they probably have the same or similar structures. Recently, a 3.0-Å-resolution cryo-EM structure of a human TRPP2 construct containing the ED and TMD has been obtained\(^{49}\). Comparisons of the structures of the TRPML1 and TRPP2 I–II linkers show that despite divergent primary sequences (Supplementary Fig. 7a), the two linkers have a very similar overall structural fold (Supplementary Fig. 7b) and form tightly bound tetramers (Supplementary Fig. 7c,d). We suggest that, given its unique presence in both TRPP and TRPML channels, the I–II linker domain be named the ‘polycystin–mucolipin domain’ instead of the ‘polycystin domain’, as has been proposed recently\(^{49}\). However, structural comparisons also reveal that the highly electronegative luminal pore loop is a unique hallmark of TRPML1 (Supplementary Fig. 7e). In TRPML1, this loop extends downward toward the ion-selectivity filter, whereas in TRPP2 it bends upward toward the luminal/extracellular entryway.

Our studies show that the novel luminal pore of TRPML1 has an important physiological function. Because of TRPML1's crucial role in numerous cellular processes, especially lysosome-dependent events, the subcellular localization and activity of TRPML1 must be tightly regulated. Although TRPML1 localizes mainly to lysosomes, it can be inserted into the plasma membrane under certain circumstances, such as during lysosomal exocytosis\(^{23–25,33}\). Cells appear to use multiple regulatory mechanisms to ensure proper levels of TRPML1 activity in different subcellular compartments—high activity in lysosomes and low activity in the plasma membrane. One mechanism is subcellular-compartment–dependent regulation by phosphoinositides. Thus, TRPML1 is activated by phosphatidylinositol 3,5-bisphosphate (PI(3,5)P\(_2\)) (ref. 22), which is enriched on the lysosome membrane\(^{23,50}\). However, TRPML1 is inhibited by PI(4,5)P\(_2\) (ref. 33), which is abundant on the plasma membrane. The Ca\(^{2+}\)/PH dual regulation is another subcellular compartmentalized regulatory mechanism. The acidic environment of the lysosome ensures high TRPML1 conductance in lysosomes by diminishing TRPML1 inhibition by luminal Ca\(^{2+}\). In contrast, the neutral extracellular pH ensures low TRPML1 conductance on the cell surface by enhancing TRPML1 inhibition by extracellular Ca\(^{2+}\).

Notably, the dual regulation by luminal Ca\(^{2+}\) and pH described here was observed in a gain-of-function mutant rather than the WT channel. Because the V432P mutation enables the channel to open constitutively, any allosteric effect of luminal Ca\(^{2+}\) and H\(^{+}\) on channel gating might not have been revealed. Our studies suggest, however, that the dual Ca\(^{2+}\)/pH regulation is due to Ca\(^{2+}\) block of the luminal pore and modulation of this block by protonation. This dual regulation is largely conferred by the luminal-pore aspartate residues, which either form a low-affinity Ca\(^{2+}\)-binding site or attract extracellular/luminal Ca\(^{2+}\) through electrostatic interactions, thus resulting in Ca\(^{2+}\) binding at a site in the transmembrane pore. From our structural and functional studies, we propose a molecular model for the dual regulation of TRPML1 by extracellular/luminal Ca\(^{2+}\) and pH. At pH 7.4, virtually all luminal-pore aspartates carry negative charges (Fig. 7c, top left). These negative charges attract and bind extracellular/luminal Ca\(^{2+}\), thereby blocking Ca\(^{2+}\) as well as monovalent-cation conduction (Fig. 7c, top right). Lowering the pH to 4.6, which is close to the pK\(_{a}\) of the aspartate side chain, results in protonation of the aspartates, thereby decreasing the net negative charges and attenuating the Ca\(^{2+}\) block (Fig. 7c, bottom right). There is residual Ca\(^{2+}\) block at pH 4.6, because some of the aspartates are unprotonated and can attract and bind Ca\(^{2+}\). The negative cooperativity observed at pH 4.6 between H\(^{+}\) and Ca\(^{2+}\) (Fig. 1c) is in good agreement with the protonation hypothesis. Mutating the aspartates to glutamines removes all the negative charges in the luminal pore and markedly further decreases Ca\(^{2+}\) block (Fig. 7c, bottom left).

A docking model (Supplementary Fig. 2) suggests that the TRPML1 I–II linker constitutes virtually all of the exposed surface of TRPML1 in the endolysosomal lumen and thus may contain all the interaction sites between TRPML1 and the lipids and proteins in the lumen. Notably, the structure of E199–K219 in the TRPML1 I–II linker was unresolved (Fig. 2 and Supplementary Fig. 1b), probably because of its flexibility. This 21-amino acid loop may protrude into the endolysosomal lumen, beyond an otherwise rigid structure, and is thus a candidate interaction site with endolysosomal lumen molecules. Finally, the architecture and strong electronegativity of the luminal pore suggest that it may be a suitable site for structure-based design of TRPML1-specific blockers.

**METHODS**

Methods, including statements of data availability and any associated access codes and references, are available in the online version of the paper.

**ACKNOWLEDGMENTS**

We thank Y. Chen and M. Chalfie (Columbia University) for providing the *C. elegans* cDNA library; R. Prywes (Columbia University) for providing HEla cells; T. Patel and S. Banta for help with the CD experiment; and the staff at X29 of the National Synchrotron Light Source, Brookhaven National Laboratory, for synchrotron support. This work was supported by grants to J.Y. from the National Key Basic Research Program of China (2011CB910301), the National Institutes of Health (R01GM085234 and R01NS053494), the National Natural Science Foundation of China (31370821), the Top Talents Program of Yunnan Province (2011HA012) and the High-level Overseas Talents of Yunnan Province; and grants to X.L. from the China Youth 1000-Talent Program of the State Council of China, the Beijing Advanced Innovation Center for Structural Biology, the Tsinghua-Peking Joint Center for Life Sciences and the National Natural Science Foundation of China (31570730).

**AUTHOR CONTRIBUTIONS**

M.L. and J.Y. conceived and initiated the project. M.L. obtained the first crystal structure at pH 6.0 and contributed to most of the other experiments. W.K.Z. performed most of the electrophysiology experiments. N.M.B. obtained the crystal structures at pH 4.5 and 7.5 and performed the imaging experiments.
COMPEITING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

ONLINE METHODS

Mutagenesis and cloning. The human TRPML1 clone (GenBank BC005149) was purchased from OpenBiosystems. The C. elegans TRPML1 (CUP-5) was cloned from a cDNA library. Site-directed mutagenesis was carried out with overlapping PCR or a QuikChange Site-Directed Mutagenesis Kit (TransGen Biotech) to produce all the mutations. The DNA sequences of all constructs were verified by DNA sequencing. Depending on the experiment, WT and mutant constructs were cloned into different expression vectors, as detailed in the respective sections.

Expression and purification of the TRPML1 I–II linker protein. A DNA fragment encoding the I–II linker (residues 84–296) of human TRPML1 was obtained by PCR and cloned into a modified pET26b(+) vector. A maltose-binding protein (MBP) tag was added to the N terminus of the I–II linker, and a DNA fragment encoding the recognition site of thrombin was inserted between the MBP tag and I–II linker. The resulting construct thus encoded a fusion protein consisting of an N-terminal MBP tag followed by the thrombin-recognition site, the I–II linker, and a C-terminal hexa-histidine tag.

Rosetta-gami 2 (DE3) cells were used to express the I–II linker to allow proper disulfide-bond formation. Transformed cells were selected and grown at 37°C in 10 l of LB medium containing 50 μg/ml kanamycin and 34 μg/ml chloramphenicol in an incubator shaker at 250 r.p.m. When the optical density at 600 nm reached 1.0, the culture was cooled to 22°C, and 0.1 mM isopropyl β-D-thiogalactopyranoside (IPTG) was added. The bacterial cells were incubated for another 12 h in the shaker at 22°C before they were harvested by centrifugation at 3,000 g for 15 min. The cells were resuspended in a solution containing 50 mM sodium phosphate, pH 8.0, 300 mM NaCl and 2.5% (w/v) glycerol (solution A). Then 5 mM imidazole, 0.5 mg/ml lysozyme, 25 μg/ml DNase and 2 mM PMSF were added, and the cells were disrupted by sonication. Insoluble cell debris was removed by centrifugation at 17,000 g for 30 min. The supernatant was incubated with Ni-NTA His-Bind resin (Novagen) with gentle agitation at 4°C for 1 h. The beads were spun down at 800 g for 1 min and transferred to a gravity-flow chromatography column. After the resin was washed with ten volumes of solution A containing 5 mM imidazole, the bound protein was eluted with 500 mM imidazole in solution A. The eluted protein was incubated with amylose resin (NEB) at 4°C for 2.5 h. The resin was collected by centrifugation at 800 g for 1 min, transferred to a column, and washed with solution A. The MBP-tagged protein was eluted with 20 mM maltose in solution A. Thrombin (Sigma–Aldrich) was added at 4 U per milligram protein to cleave the MBP tag. After incubation at 16°C overnight, the protein was purified and concentrated on a 200 column (GE Healthcare) to remove the MBP. The gel-filtration solution contained 10 mM HEPES and 150 mM NaCl, pH 7.5. The peak fractions corresponding to the tetrameric I–II linker protein were collected and concentrated to 4 mg/ml for crystallization.

Crystallization, data collection, structure determination and refinement. Crystallization of the TRPML1 I–II linker was carried out by using the hanging-drop vapor-diffusion method at 16°C (for pH 6.0) or 20°C (for pH 4.5 and 7.5). For the pH 6.0 crystal, the protein solution was mixed with reservoir solution at a 1:1 ratio. The reservoir solution contained 1.38 M sodium phosphate monobasic and 0.42 M potassium phosphate dibasic, pH 6.0, and 5% pentaerythritol ethoxylate (3/4 EO/0H) (Hampton Research). The best crystals grew from drops formed with 4 mg/ml protein. For the pH 4.5 crystal, the protein concentration was 4 mg/ml, and the reservoir solution contained 200 mM magnesium sulfate, 5.3% PEG 3350 and 100 mM acetate, pH 4.5. Crystals at pH 7.5 were obtained by macroseeding in a reservoir solution of 100 mM magnesium sulfate, 4% PEG 3350 and 100 mM HEPES, pH 7.5, with a protein concentration of 4 mg/ml. Heavy atom–derivative crystals were prepared by soaking crystals in a solution containing 1.53 M sodium phosphate monobasic and 0.47 M potassium phosphate dibasic, pH 6.0 and 1 mM K2Pt(CN)4 for 24 h.

Single crystals were flash-cooled in liquid nitrogen with Paratone-N (Hampton Research) for pH 4.5 and 6.0 crystals) or 30% glycerol as a cryoprotectant (for pH 4.5 and 6.0 crystals) or 30% glycerol as a cryoprotectant (for pH 4.5 and 6.0 crystals) or 30% glycerol as a cryoprotectant (for pH 4.5 and 6.0 crystals). X-ray diffraction data for native and derivative crystals (at pH 4.5 and 6.0) were collected at 100 K on a RAXIS-IV detector with Cu Kα radiation (λ = 1.5418 Å) from a Rigaku RuiH3R X-ray generator. X-ray diffraction data at pH 7.5 were collected at the National Synchrotron Light Source (NSLS) beamline X29 at Brookhaven National Laboratory.

The diffraction images were processed and scaled with the HKL package. The structure of the I–II linker at pH 4.5 was solved with the single isomorphous replacement with anomalous scattering (SIRAS) method with native and platinum-derivative data sets. Two platinum sites were found with Shake and Bake, and were input into SOLVE to calculate the initial phases. Density modification was done with RESOLVE to improve phase accuracy. The quality of the density map from RESOLVE was sufficiently good to locate most residues. The presence of two intrasubunit disulfide bonds aided in the tracing. The initial model was built with COOT. Refinement with CNS and manual rebuilding were done iteratively. Structures of the I–II linker at pH 4.5 and 7.5 were solved by molecular replacement with PHASER, with the structure at pH 6 used as the search model. Cycles of refinement with CNS and manual model correction with COOT were carried out. Crystallographic statistics are summarized in Table 1.

Expression and purification of the TRPML1 I–II linker for circular dichroism. TRPML1 I–II linkers were expressed through viral infection in Hi5 insect cells (Expression Systems, not tested for mycoplasma). The linkers were cloned in the pFastBac1 vector, which was modified by the insertion of a honeybee melittin secretion signal peptide coding region followed by BamHI/XhoI restriction sites and a hexa-histidine-tag coding region. DNA fragments encoding the WT or mutant I–II linkers were inserted into the BamHI/XhoI sites of the modified vector.

Recombinant baculovirus was generated with the Bac-to-Bac method (Invitrogen). The virus was amplified in Sf9 cells (Invitrogen, not tested for mycoplasma) and was used to infect Hi5 cells in ESF920 medium (Expression Systems). The expressed protein containing a hexa-histidine tag on its C terminus was secreted into the extracellular medium, where the honeybee melittin secretion signal peptide was cleaved. Forty-eight hours after transfection, cells were removed by centrifugation, and the medium containing the desired protein was collected. The medium was concentrated, dialyzed against solution A containing 5 mM imidazole and incubated with TALON metal-affinity resin (Clontech) at 4°C for 1 h. The resin was collected and washed with solution A containing 10 mM imidazole. The histidine-tagged protein was eluted with 300 mM imidazole in solution A. The protein was concentrated and changed to a solution containing 10 mM sodium phosphate, 150 mM NaCl, pH 7.0, by dialysis for CD analysis.

Circular dichroism spectroscopy. CD was performed with a J-815 CD spectrometer (Jasco). All proteins were adjusted to a concentration of 1 mg/ml and loaded into a quartz cuvette with a path length of 0.01 cm for measurement. A buffer containing no protein was used for subtraction of the baseline signal. Measurements were performed from 185 to 250 nm with a 0.1-nm interval, 1-nm bandwidth and scanning speed of 50 nm/min. For each protein, three samples were measured, and three accumulations were measured for each sample and averaged. The MRE was calculated from the raw CD signal, θ, with the equation [θ]MRE = θ/(10 × Cr × l), where Cr is the protein concentration (M × residue number), and l is the cuvette path length (cm).

Cross-linking. The DNA fragment encoding C. elegans full-length TRPML1 was cloned into a modified pFastBac1 vector. An MBP tag was added before the N terminus of TRPML1, and a hexa-histidine tag was added before the N terminus of MBP. A linker region containing a TEV protease–recognition sequence (underlined; NNNNNNENLYFQGGGGS) was inserted between the MBP tag and TRPML1. Cysteine substitutions were subsequently introduced into this background construct.

The baculovirus of TRPML1 was generated in Sf9 cells with the standard Bac-to-Bac method and was used to infect Hi5 insect cells. Forty-eight hours after infection, cells were harvested by centrifugation at 4°C and suspended in a buffer containing 50 mM HEPES-NaOH, pH 7.4, 500 mM NaCl and 5% glycerol, in the presence of a protease-inhibitor cocktail (Pierce). After cell disruption by sonication, the cell debris was removed by centrifugation at 4,000 r.p.m for 10 min at 4°C, and membranes were pelleted by ultracentrifugation at 45,000 r.p.m at 4°C for 1 h. The membranes were suspended in buffer A containing a protease-inhibitor cocktail and homogenized with a glass Dounce homogenizer. TRPML1 proteins were extracted with 1% lauryl maltose neopentyl glycol (LMNG, Anatrace) for 1 h at 4°C. The solubilized membranes were clarified by ultracentrifugation for 30 min and incubated with amylose resin (NEB) for 2 h at 4°C with gentle agitation.
The resin was collected by low-speed centrifugation at 2,000 r.p.m., transferred into a gravity column, and washed with buffer A containing 0.5 mM LMG and 0.1 mg/ml soybean lipids (Avanti Polar Lipids). This wash buffer with 20 mM maltose was used to elute the MBP-tagged TRPML1 proteins. TEV protease was added to the eluted proteins and incubated overnight at 4 °C to cleave the MBP tag. The proteins were mixed with 3× SDS loading buffer with or without DTT and analyzed with 10% SDS–PAGE.

**Native gel electrophoresis and fluorescence-detection size-exclusion chromatography.** The WT and L144K R146S mutant I–II linker proteins were obtained as described above in the CD experiment. The proteins were loaded on a 6% native acrylamide gel. After electrophoresis, the proteins were transferred onto a PVDF membrane for western blot analysis. HisTag monoclonal antibody (EMD Millipore, cat. no. 70796) was used as the primary antibody. Alexa Fluor 680 goat anti-mouse IgG (Invitrogen, cat. no. A-21058) was used as the secondary antibody. Images were scanned and analyzed with an Odyssey Infrared Imaging System (LI-COR).

For native gel electrophoresis and fluorescence-detection size-exclusion chromatography (FSEC), full-length *C. elegans* TRPML1 was cloned in the pEGFP-C1 vector (Clontech), which was modified by replacing the multiple cloning site with a TEV protease-recognition-site coding region and SbfI/Ascl restriction sites. The DNA fragment was added into the SbfI/Ascl restriction site. The expressed protein thus had an EGFP tag on its N terminus and a TEV protease-recognition site in between the EGFP tag and the channel. HEK 293T (ATCC) cells were transfected with the TRPML1 plasmid with Lipofectamine 2000 (Invitrogen). Forty-eight hours after transfection, cells were washed and resuspended in solution A containing Halt protease-inhibitor cocktail (Thermo). DDM was added to a final concentration of 2%, and lysates were incubated at 4 °C for 1 h. The solution was clarified by centrifugation at 16,000 r.p.m. for 10 min. This cell lysate was used for blue native PAGE and FSEC57.

Blue native PAGE was performed with a NativePAGE Bis-Tris Gel System (Invitrogen). The solubilized cell lysate in DDM was mixed with NativePAGE sample buffer and Coomassie blue G-250. Electrophoresis was carried out on 3–12% Bis-Tris gel at 4 °C. The proteins were transferred to PVDF membranes for western blotting. Anti-GFP monoclonal antibody (Santa Cruz, cat. no. 99966) and Alexa Fluor 680 goat anti-mouse IgG (Invitrogen, cat. no. A-21058) were used for detecting the GFP-tagged TRPML1 protein. Images were scanned and analyzed with an Odyssey Infrared Imaging System (LI-COR). All primary antibodies have been validated for their species and applications, per information provided by the manufacturers.

FSEC was performed with a spectrophotometric detector RF-10AXL (Shimadzu) for fluorescence detection and a Superose 6 10/300 column (GE Healthcare) for size-exclusion chromatography.

**Electrophysiology.** All electrophysiology source data are presented in Supplementary Data Set 2.

**Constructs.** Human TRPML1 tagged with GFP on the N terminus and containing the V432P point mutation (TRPML1VP) was cloned in the pEGFP-C1 vector. TRPML1VP-3DQ, TRPML1VP-3C and TRPML1VP-LR were also cloned in this vector. TRPML1VP was also cloned in pCDNA3.1, in which TRPML1VP-L106P, TRPML1VP-C166F and TRPML1VP-T232P were cloned.

**Cell culture and transfection.** HEK 293T cells were grown in DMEM (30-2002, GIBCO) supplemented with 0.5% penicillin/streptomycin (Sigma, P-0781) and 10% FBS (Standard quality, PAA laboratory, 95025-534) with standard procedures. Transfections were performed with Lipofectamine In vitro DNA Transfection Reagent Ver 2 (Invitrogen). Seventeen hours after transfection, the cells were loaded with 100 nM of the lysosomal marker LysoTracker Red DND-99 (Invitrogen) for 1 h in a Ca2+-free solution. The fluorescence ratios of F340/F380 were measured with a fluorescence microscopic system. The Ca2+-free solution contained 150 mM NaCl, 1 mM MgCl2 and 10 mM HEPES, pH 7.4 with NaOH. In the Ca2+-containing solution, 3 mM CaCl2 was added to the solution above.

**Confocal imaging.** HEK cells (gift of R. Prywes, Columbia University) were maintained in DMEM ( Gibco) containing 10% FBS (Atlanta Biologicals) and 100 µg/ml of penicillin/streptomycin (Invitrogen). Cells were cultured and plated onto poly-ℓ-lysine hydrobromide (Sigma)-coated coverslips 18 to 24 h before transfection and grown on the coverslips in DMEM plus 10% FBS supplemented with 100 µg/ml of penicillin/streptomycin at 37 °C until cells became ~80% confluent. Cells were transfected with WT or mutant full-length TRPML1, all tagged with EGFP on the N terminus and cloned in the pcDNA3.1 vector, with Lipofectamine 2000 (Invitrogen). Forty-eight hours after transfection, cells were loaded with Fura-2-AM (10 µM) and Pluronic F-127 (0.02%) (Molecular Probes) at 37 °C for 1 h in a Ca2+-free solution. The fluorescence ratios of F340/F380 were measured with a fluorescence microscopic system. The Ca2+-free solution contained 150 mM NaCl, 1 mM MgCl2 and 10 mM HEPES, pH 7.4 with NaOH. In the Ca2+-containing solution, 3 mM CaCl2 was added to the solution above.

**Cryo-EM sample preparation, data acquisition and image processing.** *C. elegans* TRPML1 was expressed in Hi5 insect cells and purified with amylase resin with the same procedures as described above in the cross-linking experiment. Eluted TRPML1 protein from amylase resin was mixed with amphipol A8-35 (Anatrace) at 1:6 (w/w), and TEV protease was added at 1:20 (w/w) to cleave the MBP tag. The mixture was incubated with gentle agitation at 4 °C overnight, and detergent was removed with Bio-Beads SM-2 (Bio-Rad) at 4 °C. After Bio-Beads removal, the TRPML1 protein was concentrated and further purified on a Superose 6 10/300 column in a buffer composed of 20 mM HEPES–NaOH, 150 mM NaCl, pH 7.4. The peak corresponding to the tetrameric TRPML1 protein was collected for cryo-EM analysis.

A drop of 4 μl of the amphipol-solubilized protein at 0.6 mg/ml concentration was loaded onto a glow-discharged Quantifoil R1.2/1.3 holey carbon grid, incubated for 3 s, blotted and then plunged into liquid ethane cooled with liquid nitrogen by using a Vitrobot device (FEI Company) at 100% humidity and 8 °C. The grids were imaged with a Titan Krios microscope operated at 300 kV and equipped with a K2 Summit electron-counting camera (Gatan Company). UCSFImage4 (ref. 58) was used for data collection with nominal magnification of 22,500×, corresponding to an image pixel size of 0.66 Å under super-resolution counting mode. Images were recorded with a defocus range from 2.1 to 3.1 μm. The dose rate was set to 8.2 counts/physical pixel/s on the camera.
plane. The total exposure time was 8 s, thus resulting in a total accumulated dose of ~50 e−/Å². Each micrograph was fractionated into 32 frames, each with an exposure time of 0.25 s.

All micrographs were first 2 × 2 binned, thus generating a pixel size of 1.32 Å. Motion correction was performed with MotionCorr59. The aligned frames were integrated and used for further processing. Defocus for all images was determined with CTFFIND3 (ref. 60). Particle-picking and all 2D and 3D classification and refinement steps were performed with RELION61. Approximately 200,000 autopicked particles were screened with several rounds of 2D classifications to remove most of the ‘junk’ and bad particles. The particles were 4 × 4 binned to a pixel size of 2.64 Å for further image analysis. A star-shaped density map with C₄ symmetry was made and used as an initial reference for 3D classification. 71,052 good particles were selected by 3D classification and subjected to first-round 3D refinement with C₄ symmetry imposed, thus yielding a reconstruction at a resolution of 8.12 Å. The transmembrane region showed much poorer quality than the soluble region, in agreement with the asymmetric features shown in 2D class averages. To improve the resolution of the soluble region, we applied a soft-edged mask to remove the possible transmembrane region. Further 3D refinement of the soluble region generated a reconstruction at a resolution of 5.28 Å. The resolutions were estimated on the basis of the gold-standard FSC = 0.143 criterion. ResMap62 was used to calculate the local resolution map. The crystal structure of the human I–II linker at pH 6.0 was fit into this cryo-EM density map with Chimera63.

Data availability. 3D cryo-EM density maps of full-length C. elegans TRPML1 or the I–II linker, without and with low-pass filtering and amplitude modification, have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-6669 and EMD-6670, respectively. The coordinates of the atomic model of the human TRPML1 I–II linker at different pH values have been deposited in the Protein Data Bank under accession numbers PDB 5TJB (pH 4.5), PDB 5TJA (pH 6.0) and PDB 5TJC (pH 7.5). Data for electrophysiology experiments in Figures 1, 3, 4, 5 and 6 and for Supplementary Figs. 5 and 6 are provided as Supplementary Data Set 2. Other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Amino acid sequence alignment of TRPML subunits.

(a) Subcellular localization and transmembrane topology of TRPML1. (b) Amino acid sequence alignment of TRPML subunits. Green and yellow mark identical and similar amino acids, respectively. Putative S1-S6 segments are indicated in gray. Secondary structures of TRPML1 I-II linker are colored in the same scheme as in Fig. 2c. Bold black residues boxed in red are involved in intersubunit interactions. Red triangles mark the luminal pore-loop aspartates. Stars mark the amino acids whose point mutation causes MLIV.
Supplementary Figure 2

A docking model of the TRPML1 I–II linker.

The TRPML1 I-II linker structure is visually and manually docked onto the transmembrane domain (TMD) of the TRPV1 structure (PDB code: 3J5P). (a,c) and (b,d) show the ribbon and surface representations, respectively, of the docked structures. Upper panels, top down views from the extracellular/luminal side of the membrane. Lower panels, side views parallel to the membrane. The $\alpha_1$ helix of the TRPML1 I-II linker and S1 of TRPV1 are highlighted in red and gold, respectively, in (a,c).
Supplementary Figure 3

Single-particle cryo-EM analysis of TRPML1. (a) A representative micrograph. Typical particles are marked with yellow boxes. (b) Fourier power spectrum of the micrograph shown in a with the Thon ring extending to 3 Å. (c) Outline of image processing. After multiple rounds of 2D and 3D classification to remove unwanted particles, 71,052 particles were selected from 201,010 autopicked particles and subjected to refinement with C4 symmetry imposed, yielding a reconstruction at a resolution of 8.12 Å. Due to the possible flexibility of the transmembrane domain (TMD), a soft mask surrounding the more rigid extracellular domain (shown in green) was applied for further refinement, resulting in a reconstruction at a resolution of 5.28 Å. The angular distribution of the final structure is also shown. (d) Enlarged views of representative 2D classes. Orange and green arrows point to TMD regions with different densities, a phenomenon suggestive of flexibility in the TMD. (e) FSC
curves of the final TRPML1 reconstruction with (blue) or without (black) mask. The pixel size used was 2.64 Å; thus, the blue curve crossing the Nyquist frequency yields a 5.28 Å resolution.
**Supplementary Figure 4**

MLIV-causing mutations cause TRPML1 mislocalization.

(a) Location of three MLIV-causing missense mutations in the I-II linker structure, marked in red. (b) Confocal images of live HeLa cells expressing the indicated GFP-tagged channels. Red indicates LysoTracker-labeled lysosomes.
Supplementary Figure 5

Cysteine modification in the luminal pore reduces ion conduction. (a, b) Time course of whole-cell currents of the indicated channels in response to 5 mM extracellular MTSET. NDF: nominal divalent cation free. (c) Current-voltage curves taken at the time points indicated in b. (d) Normalized and averaged current amplitude of TRPML1<sup>WT</sup>-3C at -80 mV of at pH 7.4. Number of recordings is indicated inside the bar. Error bars represent SEM. * p<0.05 with Student’s t-test.
Supplementary Figure 6

The luminal-pore aspartate mutations do not affect inward rectification.
(a) Families of TRPML1<sup>W73</sup>-3DQ currents at the indicated pH 7.4 and extracellular Ca<sup>2+</sup> concentrations. (b) Current-voltage relationship of the currents in a.
Supplementary Figure 7
Comparison of the structures of the I–II linkers of TRPML1 and TRPP2.
(a) Amino acid sequence alignment of the TRPML1 and TRPP2 I–II linkers. Secondary structures are indicated. Bold black residues are identical residues in both sequences. Bold red residues are the luminal pore-loop aspartates. (b) Superposition of the I–II linker protomer structures of TRPML1 (pH 6.0) and TRPP2 (PDB code: 5T4D), aligned by the β strands. (c, d) Superposition of the TRPML1 and TRPP2 I–II linker tetramer structures, aligned by the β strands and viewed from above (c) or parallel (d) to the membrane. e, Same
view as in (d) but showing only two diagonally opposed subunits, highlighting the different orientations of the luminal pore-loop of TRPML1 and its counterpart in TRPP2.
Appendix B:

Vita and Publications
Vita and Publications

2011  Bachelor of Arts, Cornell University
2013  Master of Arts, Columbia University
2014  Master of Philosophy, Columbia University
2017  Doctor of Philosophy, Columbia University
