

Structural Study of Piezo Channel, a Novel  
Eukaryotes Mechanosensitive Channel.

Thesis by  
Aron Kamajaya

In Partial Fulfillment of the Requirements for  
the degree of  
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY  
Pasadena, California

2017  
(Defended July 12, 2016)

© 2017

Aron Kamajaya

All Rights Reserved

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank **Doug Rees** for his support and mentorship throughout my graduate school years. I am grateful that you gave me freedom to pursue an independent and unique project, which catalyzes my academic growth to become a scientist. Over the years, you have become my role model both as a scientist and a leader.

I would also express my gratitude for the members of my thesis committee: **Henry Lester, Bil Clemons, and David Chan**. Your insight, advice, and guidance are critical for my success in this journey. **Henry**, serving as TA in your class for four terms has ignited my interest in neuroscience. **Bil**, thank you for being such a cool and approachable person to talk to about my project outside the formal meeting. **David**, thank you for offering your expertise in cell biology. I love the close interaction between your lab and the Rees lab.

Next, I would like to thank the **members of the Rees group**, former and present, for creating such an environment that is conducive for scientific research. I also would like to give a shout out to my colleagues who have been instrumental to my success. **Phong Nguyen**, thanks for being an awesome partner in crime and fellow ‘ESL’ student. That *Costco Friday* and *monthly gathering* really helped ease out the pressure. Thanks to **Jonas Lee** and **Jens Kaiser** for being patient and teaching me crystallography. I also would like to thank **Troy Walton, Chris Gandhi, Janet Yang, Nadia Herrera**, and the ‘adopted Rees group member’ **Raymond Liu** for your help in proofreading my lengthy manuscripts and occasionally chatting about random cool science. I would like to express my appreciation to the two individuals whom I had the pleasure to work with: **Michelle Reid**

and **Pavle Nikolovski**. I expect to hear great things from you guys! **Allen Lee, Jeffrey Lai**, and **Phoebe Ray**, you have done a wonderful job to keep the lab running smoothly, especially in term of ordering reagents, ensuring the availability of common reagents, scheduling a timely service, resolving occasional conflict, and repair for the malfunction equipment, particularly AKTA and fermenter. I would also like to thank **Grant Jensen** for organizing the single particle EM sub-group, which helped me during the initial struggle to get started in my single particle EM journey. Thanks to **Alasdair McDowall** for aiding me with the single particle data collection and to **Haoqing Wang** for helping me in the EM data processing. Big thanks to **Jost Vielmetter** for allowing me to use the space at the Protein Expression Center (PEC) facility and providing me with a lot of useful discussion on different options to express my gigantic protein, and to **Angela Ho** for the assistance on developing the stable cell line.

It's been an interesting twist of fate and a thrilling journey for me to get here. Thus, I am going to mention several people who have been the heroes in my life. First, I would like to thank my wife, **Monica “the bunny” Lui**, who has been patiently listening to my rambling, so others don't have to. She is the reason why I seem to be cheerful and happy at all times in the workplace. Also, I want to acknowledge my best friend and current gym buddy, **Feng Chen**, who's always there with me even in my darkest hours. I want to thank **Stephen Smith**, my academic mentor, and **Kathleen Ryan**, my research advisor at UC Berkeley, for investing in my growth as a young scientist. I also want to thank the people in the Biological Technology program at Pasadena City College, **Wendie Johnston** and **Pamela Eversole-Cire**. You opened my eyes and showed me the joy of scientific research

and discoveries. I would like to thank **Kathleen Talaro** for believing in me when nobody did. You have been the miracle in my life, and none of these great things would have happened without you. Together with **my parents-in-law**, you have made the transition to a higher education possible. In addition, I would like to thank **my late uncle** and **my auntie** for bringing me to the United States and giving me a chance to succeed in America, the land of opportunity. I also thank my sister, **Dina Malounda**, for living the thrilling journey together with me as the generation zero in this foreign land. Last but not least, I want to thank **my parents** for bringing me into this world and teaching me that I can always achieve anything that I set my mind into. Thank you for investing in my education even when the finances got very tough. Thanks for all of your prayer, mom.

## ABSTRACT

Piezo is a unique family of eukaryotic mechanosensitive (MS) channel. With over 2500 amino acids per subunit, intact Piezo channel is one of the largest ion channels known to date. Two versions of Piezo can be found in vertebrates, namely PIEZO1 and PIEZO2. PIEZO1 appears to play roles in processes which control physiological homeostasis, whereas PIEZO2 assumes roles in mechanical somatosensation. A number of mutations mapped onto PIEZO1 or PIEZO2 are found in several hereditary human diseases, such as Dehydrated Hereditary Stomatocytosis, Gordon syndrome, and Distal Arthrogryposis. Although biochemical and functional studies provided many insightful findings, structural study of Piezo was very minimal. Herein, I described the structural investigation of Piezo channel. In the first study, we isolated a conserved soluble domain of Piezo (C-terminal loop 2, CTL2) from the *C. elegans* homolog, and provided the first molecular glimpse into this enigmatic MS channel. Subsequently, I described challenges that are associated with the expression and protein preparation of the full length Piezo channel. Recently, the full length mouse PIEZO1 structure solved by single particle cryo-EM revealed trimeric arrangement of the intact channel. CTL2 domain forms an extracellular cap which makes up the central core in this Piezo model. Lastly, we isolated a stable C-terminal fragment of Piezo. This fragment corresponds to the entire central core of Piezo channel and a few upstream transmembrane helices. This fragment can be localized to the plasma membrane. Further investigation is needed to look at the functionality of this fragment.

## PUBLISHED CONTENT AND CONTRIBUTIONS

Kamajaya, A., Kaiser, J.T., Lee, J., Reid, M., Rees, D.C. (2014) The structure of a conserved piezo channel domain reveals a topologically distinct  $\beta$  sandwich fold. *Structure* 22(10):1520-7. doi: 10.1016/j.str.2014.08.009.

A.K. designed the experiments, performed most of the experiments, analyzed the results, and wrote the manuscript. J.T.K. and J.L. provided mentorship in x-ray diffraction data processing. M.R. performed the early protein purification. D.C.R provided the mentorship, funding, and editorial service.

# TABLE OF CONTENTS

<b>Acknowledgements.....</b>	iii
<b>Abstract .....</b>	vi
<b>Published Content and Contributions.....</b>	vii
<b>Table of Contents.....</b>	viii
<b>Chapter 1: Primer on Mechanosensation and the Piezo Channel.....</b>	1
Mechanosensation .....	2
Bacterial Mechanosensitive (MS) Channels.....	4
Mechanosensation in Multicellular Eukaryotes .....	5
The Plant MSL channel Family .....	6
Diverse Families of eukaryotic MS Channels .....	7
Piezo Family of Metazoan MS Channel	
Initial Discovery .....	9
Physiological Role.....	10
Current Puzzles.....	13
The Scope of the Thesis .....	16
Chapter 1 Figure .....	17
References.....	18

**Chapter 2: The Structure of a Conserved Piezo Channel Domain****Reveals a Beta Sandwich Fold.....33**

Summary and Highlights.....34

Introduction.....35

**Results**

Construct Selection.....37

Architecture of the *C. elegans* CTL2 Domain.....38Structure of M31R *C. elegans* PIEZO CTL2 Domain Variant.....39

Discussion .....40

Experimental Procedures.....42

Acknowledgements and Author Contributions .....46

References.....47

Figures and Table .....52

<b>Chapter 3: The Pursuit Toward Obtaining a Full Length PIEZO Structure .....</b>	<b>67</b>
Historical Context and the Chapter Outline.....	68
Part 1: Exploring Various Protein Expression Systems	
Introduction.....	71
Bacterial Protein Expression Systems .....	72
Pichia Protein Expression System .....	73
The Journey Begins	
Mammalian Protein Expression System.....	75
Fluorescence Size Exclusion Chromatography Constructs ...	76
Exploring Other Ways of Expressing Full Length PIEZO .....	78
Part 2: Visualizing Full Length Mouse PIEZO1 Using Single Particle Electron Microscope (EM)	
Transition from X-Ray Crystallography to Single Particle EM .....	83
Initial Negative Stain EM Observation.....	84
Full Length Mouse PIEZO1 is a Trimer.....	86
Chapter Summary and Future Direction.....	88
Methods .....	89
Author Contributions.....	98
References.....	99
Figures.....	105

**Chapter 4: Isolation of a Stable C-terminal Piezo**

<b>Fragment .....</b>	126
Background and Motivation .....	127
C-terminal Piezo Constructs Expression and Detergent Extraction	
Screen .....	130
Isolation of a Stable Fragment of Mouse PIEZO1 .....	133
C-terminal Fragment in the Context of the Full Length Mouse PIEZO1	
Structure .....	134
Mouse PIEZO1 C-terminal Fragment can be Expressed and Localized to the	
Plasma Membrane .....	135
Future Directions .....	138
Methods.....	140
Author Contributions .....	145
Reference .....	146
Figures and Table.....	151

# **Chapter 1:**

## **Primer on Mechanosensation and the Piezo Channel**

## Mechanosensation

Life does not exist in a vacuum. The interactions between cells and their surroundings are crucial for cells' survival. To persist, cells must be able to detect changes in their environment and execute appropriate responses to adapt to these changes. Living organisms are continually experiencing dynamic mechanical forces, and the ability to detect these forces is essential for many biological processes [1]. Mechanosensation is one sensory modality which allows living organisms to detect mechanical force changes on their surroundings. The mechanistic explanation of how these forces are detected by biological mechanosensors remains incompletely understood. One way cells can convert mechanical cues into intracellular signals is through mechanosensitive (MS) channels. MS channels form permeation pathways upon activation by mechanical stimulation. Typically, ions flow through the permeation pathway, which in turn provides signals that lead to further downstream intracellular signaling cascades.

Challenges associated with studying MS channels, especially in multicellular eukaryotes, have been previously described [2]; this includes the scarcity of mechanosensory cells in multicellular eukaryotes, the minimal expression of MS channels even in the specialized mechanosensory cells, and the difficulty in developing an ideal assay to investigate the mechanosensory properties of the putative MS channels outside their native environment. Despite these challenges, several key advances have aided the quest to understand the underlying principle of mechanosensation. These include the development of various electrophysiological methods to assay putative MS channels,

bioinformatics tools to search for a putative MS channel in the available genomic data, and molecular cloning techniques which enable heterologous expression of a putative MS channel for further biochemical characterization.

Currently, there are two prominent models to explain the mechanisms by which MS channels mediate mechanosensation. The first model, and arguably the more prominent model, is commonly referred to as the “force from lipid” model [3-5]. In this model, perturbation of lipid bilayer is sufficient to promote the opening of MS channels [6]. Great examples for this model are the bacterial mechanosensitive channels. These channels retain their mechanosensing ability upon purification and reconstitution in a lipid bilayer [7]. The second model is the “tethered trapdoor” model [8, 9]. Here, MS channels are attached to either external or internal structures by a molecular tether. Perturbation in these structures causes the molecular tether to pull on the MS channels, which drives them to adopt an open channel state. This model was inspired by the geometry of the mechanotransduction apparatus in the hair cells of the inner ear of vertebrates [10, 11]. Here, MS channels are localized to the tip of the hair cells’ stereocilia where they are tethered by an extracellular tip links to the adjacent stereocilia. Although there is evidence to support both proposed models, there has been the suggestion that the “tethered trapdoor” model is a reiteration of the “force from lipid” model [4, 5].

## Bacterial Mechanosensitive Channels

The two best studied MS channels are the bacterial MS channels: the mechanosensitive channel of large conductance (MscL) and the mechanosensitive channel of small conductance (MscS). These MS channels, first described in 1987 [12] and subsequently cloned [13, 14], are thought to serve as safety valves which protect bacterial cells against osmotic downshock as well as intracellular pressure during cell growth [12, 15-17]. *Mycobacterium tuberculosis* MscL gene codes for 151 amino acids with molecular weight of 16 kDa, consisting of two transmembrane helices (TM1 and TM2), a short cytoplasmic N-terminal helix, and a C-terminal bundle [18]. In the closed state, TM1 segments from five MscL monomers form a constricted pore likely representing a non-conducting state [18, 19]. The single channel conductance of MscL is measured to be around 3 nS [19], corresponding to a channel pore diameter in the open state estimated to be at least 25 Å wide by Electron Paramagnetic Resonance (EPR) studies [20] as well as molecular sieving experiments using large organic ions [21]. Unexpectedly, the crystal structure of truncated *Staphylococcus aureus* MscL revealed a homotetrameric arrangement [19], and the variation in oligomeric state between the two MscL structures may reflect unintended consequences of truncation and detergent solubilization [22].

*Escherichia coli* MscS gene codes for a subunit of 286 amino acids with a molecular weight of almost 31 kDa. The MscS monomer consists of three N-terminal transmembrane segments (TM 1 - 3), a middle-β domain, and a cytoplasmic C-terminal domain [23]. Seven *E. coli* MscS monomers assemble together to form a channel with TM3 segments lining the permeation pathway. MscS exhibits a single-channel conductance of about 1 nS with an estimated open-state pore diameter of 11 Å [23, 24]. Despite the

availability of much biophysical and structural information on these channels, the gating mechanisms remain unclear. In addition to MscL and MscS, there are other bacterial MS channels, including five MscS homologs, such as MscK (potassium-dependent MS channel) and MscM (mini) [25]. The precise role of these MS channels, however, remains an active topic of investigation.

### **Mechanosensation in Multicellular Eukaryotes**

The ability to sense mechanical cues is notably important for multicellular eukaryotes. Many essential cell-to-cell communications in multicellular eukaryotes rely on mechanical cues, such as gravitropism in plants [26], proper formation of neural networks [27], biosynthesis of extracellular matrix (ECM) by chondrocytes in response to dynamic loading [28, 29], and many others. Dysregulation of mechanotransduction is linked to the abnormal cell proliferation that leads to the development of cancer [30]. Furthermore, sensing mechanical cues allows multicellular organisms to interact with their ecosystem, and perhaps to improve the likelihood of their survival. For example, many animals rely on their ability to detect sound waves as a way to avoid predators. Also, mechanical somatosensation and proprioception are crucial for the proper execution of locomotive action as well as the detection of pain and pleasure. Moreover, actively climbing plants rely heavily on their ability to identify a solid surface through which they can climb to gain access to sunlight [31]. In addition, some plants and animals rely on the mechanical sensing to forage for food [32, 33].

## Role of Bacterial MS Channel Homologs in Plants: The MSL Channel Family

In plants, a number of putative mechanosensitive (MS) channel families have been identified. Interestingly, ten homologs of the bacterial MscS channel have been found in the *Arabidopsis thaliana* genome [34]. They are named the MSL (MscS-Like) family, with the individual genes identified as MSL 1 – 10. The functional properties of these channels have been studied by the Haswell group. Knocking out the MSL2 gene causes aberrant leaf morphology [35]. Both MSL2 and MSL3 are localized to the plastid envelope, and *msl2* *msl3* double mutants exhibit an abnormal plastid size and shape [36]. It is remarkable that MSL3 is able to rescue MJF456, an *E. coli* strain which lacks MS channels [36]. Furthermore, MSL9 and MSL10 are found on the plasma membrane of root cells, and are required for mechanosensitivity [37].

With an exception of the highly conserved MscS domain, members of MSL family are topologically diverse. MSL 1 – 3 are predicted to have five transmembrane segments with a large C-terminal domain, whereas MSL 4 – 10 are predicted to have six transmembrane helices with both N-terminals and C-terminal ends facing the cytoplasmic compartment [37-39]. The single channel conductances of plant MSL9 and MSL10 are about one order of magnitude less than that of MscS. MSL9 and MSL10 have single channel conductance of 45 pS and 137 pS, respectively, in root protoplast [37], whereas, heterologously expressed MSL10 in Xenopus oocytes exhibit slightly lower conductance (100 pS) and a moderate preference for anions [40]. Currently, there is no structural information on any member of the MSL MS channel family. The gating mechanism and the oligomeric state of these channels have yet to be established.

## Diverse Families of Ion Channel Mediating Mechanosensation in Eukaryotes

Beside plant MSLs, other eukaryotic channel families have been shown to play roles in mechanosensation, including the Degenerin/ Epithelial sodium (DEG/ENaC) channels, two-pore domain potassium ( $K_2P$ ) channels, Transient Receptor Potential (TRP) channels, Trans-Membrane Channel-like (TMC) protein, and Piezo channels.

The Degenerin (*deg*) genes were first identified in *C. elegans* mutants that showed constitutively active sodium currents, leading to the degeneration of mechanosensitive neurons [41]. The DEG/ENaC channel family can be activated by multiple stimuli, including mechanical force [42]. Members of this family are proposed to mediate gentle touch, as well as perform other physiological roles, in *C. elegans* and *Drosophila* [42, 43]. The structure of an evolutionarily related channel, Acid Sensing Ion Channel (ASIC), has been solved by x-ray crystallography [44]. This structure established that ASIC1 assembles as trimer where each monomer possesses two transmembrane helices connected by extensive extracellular domain.

Two-pore domains potassium ( $K_2P$ ) channels were originally identified in yeast [45]. Among the 15 members of  $K_2P$  channel family, only three of them have been shown to have mechanosensing capability: TREK-1, TREK-2, and TRAAK [2]. Interestingly, transmission of mechanical cues seems to be directly mediated by the surrounding membrane [46], indicative of the “force from lipid” model. Intact  $K_2P$  channels are formed by two subunits, where each subunit possesses two ‘pore domains’, in contrast to the typical organization of potassium channel as tetramers where each subunit has one pore domain. Each pore domain consists of a pore helix that serves as a selectivity filter between two transmembrane helices [47]; parts of the first transmembrane helix makes up the

extracellular helical cap [46]. The channel gating mechanism of K<sub>2</sub>P family has been proposed based on several crystal structures of TREK-2 and TRAAK [48-51].

The Transient Receptor Potential (TRP) channel family consists of ion channels that respond to a wide variety of stimulants, including chemical, temperature, and mechanical stimuli. Aside from the yeast TRPY, members of this family are classified into seven subgroups: TRPC, TRPV, TRPA, TRPM, TRPP, TRPN, and TRPML [52]. In general, TRP channels possess six transmembrane segments with varying accessory domains at their cytosolic N- and C-terminal ends [53]. A number of TRP channels are linked to mechanosensation [54]. Through previous biochemical and mutagenesis studies, TRP channels topologies are thought to resemble voltage-gated potassium channels [53, 55]. In fact, the six transmembrane segments of Kv1.2 fit nicely in the low resolution (19 Å) TRPV1 cryo-EM map [56]. Atomic structures of TRP channels have recently been solved by single particle cryo-electron microscopy (cryo-EM) [57-59]. Several members of TRP channel family possess Ankyrin Repeat Domains (ARD) at varying lengths; among them, TRPN and TRPA1 have unusually long ARDs. This domain has been proposed to serve as the gating spring which enables these channels to respond to mechanical cues. In fact, ARD was previously demonstrated to have nanospring behavior using atomic force microscopy [60]. Attaching ARD from NOMPC, a member of TRPN, into voltage-gated channels causes these channels to be mechanosensitive [61]. Previously, TRPA1 was suggested to be the primary mechanotransducer of the auditory system [62]. However, there is no sign of deafness in TRPA1 knock-out mice [63, 64].

The Trans-Membrane Channel-like (TMC) family has also been proposed to be the primary mechanotransducer for hair cells. Members of the TMC family, TMC1 and TMC2,

are expressed in mice cochlea and vestibular organs during the early postnatal periods [65, 66]. TMC1 becomes a strong candidate for the mechanosensor of the auditory system because a mutation in this gene has been found in patients that have trouble hearing and in the *Beethoven* mice, which is a mouse model for deafness [65, 67]. There is currently limited structural information on this channel family. TMC1 is predicted to have six transmembrane segments with both N- and C- termini facing the cytosol [68]. Although the predicted topology for TMC1 suggests similarities to the topology of TRP channels, the identity of the putative pore domain is still unclear [4].

### **Piezo Family of Metazoan Mechanosensitive Channel – Initial Discovery**

Despite the number of ion channel families implicated in mechanical sensory transduction, the role of these ion channels as the primary transducer of mechanical cues has not typically been established [2]. In 2010, the Patapoutian group performed an RNAi knockdown screen in a cell line with a consistent mechanically-activated (MA) current, mouse Neuroblastoma cells (Neuro2A), in a quest to find a novel family of eukaryotic MS channels [69]. From this screen, they found that knocking down the *Fam38A* (Family with sequence similarity 38) gene resulted in significant decrease of MA current; *Fam38A* was later renamed as *Piezo*. Vertebrates possess two different versions of Piezo: *Fam38A* (*Piezo1*) and *Fam38B* (*Piezo2*). They found that both *Fam38A* and *Fam38B* are expressed in putative mechanosensory organs, and the protein product of these genes can be localized to the plasma membrane [69]. MA currents can be detected on a cell line that is not naturally mechanosensitive (human embryonic kidney – HEK293T), upon heterologous

expression of PIEZO1 or PIEZO2 [69]. Therefore, Piezo satisfies the criteria for a transducer of mechanical signals [2, 70, 71].

Subsequent investigation revealed that Piezo proteins are non-selective cation channel that can be efficiently blocked by 30 – 50  $\mu\text{M}$  of ruthenium red and gadolinium ion [69, 72] that have been previously identified as general inhibitors of MS channels. Mouse PIEZO1 has a subunit molecular weight of ~250 kDa, while the molecular weight of intact mouse PIEZO1 was estimated as 1.2 MDa, based on native PAGE as detected by western blot using custom polyclonal antibodies raised against a synthetic peptide that corresponds to the conserved region of mouse PIEZO1 [72]. Together with results from photobleaching experiments as well as crosslinking studies, PIEZO1 was proposed to form a tetrameric complex [72]. Mouse PIEZO1 exhibits a single channel conductance of 30 – 50 pS [72] which exceeds the typical conductance of shaker family potassium channel, 18 – 30 pS [73-76], or typical eukaryotic ion channels in general (1 – 30 pS) [77]; it is, nevertheless, about one to two order of magnitude less compared to the bacterial MscS and MscL. Upon reconstitution into lipid bilayers, mouse PIEZO1 exhibited discrete channel opening that could be blocked by ruthenium red. However, the channel conductance of reconstituted mouse PIEZO1 is significantly higher (110 – 120 pS), and does not exhibit mechanosensing activity [72].

### **Piezo Family of Metazoan Mechanosensitive Channel – Physiological Role**

Piezo is found in almost all eukaryotes, including vertebrates, invertebrates, plants, and pathogenic protozoa [69, 78, 79]. The Piezo knock out in *Drosophila* strain is viable

and shows no defects involving either locomotion coordination or the bristle mechanoreceptor; the mutant, however, does exhibit defects in detecting noxious mechanical stimulus at the larval stage [80]. Since the *Drosophila* mutant lacking Piezo does not exhibit any defects in locomotion, Piezo-dependent detection of noxious mechanical stimuli is thought to function in parallel with PPK26 (pickpocket), a member of DEG/ENac channel family that was previously shown to be essential for the detection of noxious mechanical stimuli [80-83].

In general, PIEZO1 appears to play roles in processes which control physiological homeostasis of multicellular organisms. Before Piezo was established as a MS channel, PIEZO1 (previously known as Fam38a and Mib) was shown to be expressed in astrocytes which are positioned at the proximity of senile plaques in the Alzheimer's disease mice model; indeed, PIEZO1 expression is significantly increased in rat astrocytes that have been treated with  $\beta$ -amyloid [84]. Also, Piezo is proposed to be involved in sensing the nano-roughness of a cell's surrounding terrain in rat hippocampal primary neuron culture [85]. Later, PIEZO1 was shown to direct the differentiation of neuronal-glial progenitor cells by controlling cellular localization of Yap and Taz [86], transcriptional co-activators that were previously shown to change their cellular localization in response to the stiffness of the surrounding substrate [87, 88]. Furthermore, PIEZO1 is shown to activate integrin [89], and the loss of PIEZO1 increases 2D and 3D cell migration in matrigel matrix [90], a complex media formulation which mimics the extracellular matrix (ECM) of an organism [91, 92]. In fact, expression of PIEZO1 in multiple small cell lung cancer (SCLC) cell

lines is notably reduced compared to the normal lung epithelial cell line [90].

Human *Piezol1* mRNA in the periodontal ligament cells (hPDLC) is upregulated within 0.5 hr in response to the application of static compressive force, and the high mRNA level is maintained for the next 12 hours [93]. hPDLC cells were previously shown to induce differentiation of the progenitor into osteoclasts or osteoblasts in response to mechanical signals [94, 95]. Since GsMTx4, a general chemical inhibitor for MS channels [96], was shown to inhibit osteoclastogenesis, MS channels such as PIEZO1 are thought to direct the cellular differentiation events [93]. Also, live cell extrusion of epithelial cells in response to overcrowding requires PIEZO1; knocking down PIEZO1 in zebrafish prevents live cell extrusion, which leads to formation of epithelial cell aggregates [97]. In addition, PIEZO1 was shown to be essential for the proper vascular development in mice [98, 99] as well as for the maintenance of red blood cell volume in zebrafish [100].

In contrast to vertebrate's PIEZO1, PIEZO2 assumes roles in mechanical somatosensation, which is also referred to as the discriminative touch. First, PIEZO2 is a putative downstream target of Epac [101], a guanine nucleotide exchange factor that changes pain threshold in a cAMP-dependent fashion [102, 103]. Activation of Epac sensitizes PIEZO2, which contributes to mechanical allodynia [101]. It is also shown that activation of the bradykinin receptor (BDKRB2) elevates PIEZO2-dependent current as well as slows the inactivation of mammalian PIEZO2 [104]; bradykinin is a potent inflammatory peptide which sensitizes peripheral mechanonociceptor causing acute mechanical hyperalgesia [105-107]. Furthermore, PIEZO2 has been demonstrated to be the primary mechanosensor for light touch in mouse Merkel cells [108-110], general low-threshold mechanoreceptors (LTMR) [111, 112], and zebrafish's Rohon-Beard cells [113]. In tactile-foraging

waterfowls, such as ducks and mallards, PIEZO2 is expressed in trigeminal ganglia (TG), and it is thought to be the mechanosensor which enables them to forage for food without using their vision [33]. In addition, both PIEZO1 and PIEZO2 are shown to be the transduction channels for high-strain stimuli on the articular cartilage [114].

In addition, Piezo mutations in human have been associated with multiple clinical disorders [115]. Multiple point mutations found in patients suffering from Dehydrated Hereditary Stomatocytosis (DHS) are mapped onto human PIEZO1 [116-118]. DHS is an autosomal dominant genetic disorder which is characterized by the inability of red blood cells to regulate their volume [119]. Patients suffering from this disorder exhibit clinical hemolytic anemia phenotype; a similar phenotype has also been described in zebrafish [100]. Furthermore, mutation of the human PIEZO2 at a conserved isoleucine residue (I802F) as well as a conserved glutamate residue located at the putative C-terminal cytoplasmic domain (E2727del) are associated with the Gordon syndrome (GS) and Distal Arthrogryposis (DA) type 5 [120, 121]. Both of these disorders are rare autosomal dominant genetic disorders that are characterized by multiple congenital contractures of the hands and feet [122-125].

### **Piezo Family of Metazoan Mechanosensitive Channel – Current Puzzles**

Despite the rapid progress that has been made to uncover Piezo's role in cellular physiology, we've only scratched the surface on the quest to understand the mechanism by which this enigmatic MS channel works. Currently, channel electrophysiology characteristics of Piezo have been extensively explored [69, 72, 126-128]. From biochemical and electrophysiological studies, Piezo has been proposed to assemble as an

oligomer, potentially a tetramer, with no requirement for direct interaction with other proteins to function [72]. It is unclear whether the Piezo oligomer is assembled to form one permeation pathway or multiple independent pores. Considering that the single channel conductance of Drosophila Piezo is 9-fold less (3.3 pS) than the mouse PIEZOs [72], it is interesting to see whether fly Piezo also assembles in similar oligomeric assembly. Human PIEZO1 is suggested to have a pore diameter of at least 8 Å since it is able to permeate organic cations TMA and TEA [128]. Furthermore, efforts to determine the channel topology are limited to bioinformatics predictions with minimal experimental results. Due to differences in the topology prediction algorithms for eukaryotic membrane proteins, different topology prediction algorithms predict that Piezo consists of between 30 to 40 transmembrane segments (**Fig 1**). Generally, the Piezo protein sequence is conserved toward the C-terminal portion of the protein, especially in the regions which are predicted to be the transmembrane helices [129]; this may highlight the crucial role of Piezo's C-terminal region in the general function of the channel. In line with this idea, the first two clinically-relevant point mutations of human PIEZO1, M2225R and R2456H, are mapped onto predicted soluble domains at the conserved C-terminal region of the protein [116]. These mutant PIEZOs exhibit delayed inactivation [127, 130]; in fact, all clinically-relevant human PIEZO1 mutants exhibit similar phenotype [117]. In the absence of any structural information, however, it is difficult to propose any mechanistic models to explain how these mutations alter the gating kinetics of the channel.

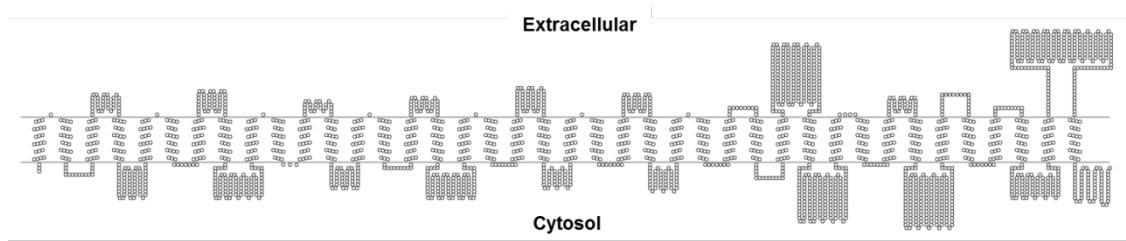
There are several proteins that are known to modulate Piezo channel's activity. Stomatin-like protein 3 (STOML3) protein is shown to lower the opening threshold for both PIEZO1 and PIEZO2 under the substrate deflection assay [131]. STOML3 is the

homolog of stomatin-domain family, MEC-2, which is required for mechanotransduction in *C. elegans* [132-134]. When overexpressed in HEK293 cells, STOML3 was shown to physically interact with both PIEZO1 and PIEZO2 [131]. The mechanism by which STOML3 binds to Piezo and sensitizes the channel awaits further investigation. Furthermore, wildtype Polycystin2 (PC2), a member of Transient Receptor Potential P (TRPP) subfamily, is shown to inhibit PIEZO1-dependent stretch-activated current in the proximal convoluted tubule cells; a mutant form of Polycystin2 (PC2-740X) is a stronger inhibitor of PIEZO1 compared to the wildtype Polycystin2 [135]. The two proteins were shown to physically interact via the N-terminal domain of Polycystin2 [135]. Polycystin2 is responsible for sensing the deflection of primary cilia of the kidney tubular cells, and mutation of this channel causes polycystic kidney disease (PKD) [136-139]. The nature of the interaction, the interplay between these two channels, and the role of PIEZO1 in kidney pathology are still a mystery.

In addition, it is unclear why Piezo maintains its enormous size across the Piezo homologs. One may suggest that Piezo utilizes its large size to survey the lipid bilayer environment more efficiently since a significant portion of this channel is predicted to be transmembrane helices. Alternatively, it has been suggested that some parts of the Piezo channel play non-mechanosensory roles, such as protein network interactions and others [115]. This begs the question whether there are minimal Piezo variants which still maintain mechanosensing capabilities. Last but not least, the Patapoutian group has recently demonstrated that mouse PIEZO1 can also be modulated and mildly-activated by a chemical agonist, Yoda1; this raises the possibility for an endogenous chemical agonist for this channel [140].

## The Scope of the Thesis

There are many interesting aspects of mechanosensitive (MS) channels that are yet to be explored. In this thesis, my focus will exclusively be on the structural characterization of the eukaryotic Piezo MS channel. When I started this project, there was no structural information on Piezo. My primary objective is to use a structural biology approach to investigate fundamental questions concerning the structure and mechanism of Piezo channels. The Patapoutian lab generously provided me with five Piezo constructs: mouse PIEZO1, mouse PIEZO2, human PIEZO1, *Drosophila* PIEZO, and *C. elegans* PIEZO. However, the sizes as well as the eukaryotic nature of these channels provide formidable challenges just to get the amount of protein suitable for structural characterization using x-ray crystallography. Therefore, the first objective of my study was to isolate smaller Piezo constructs that would be more amenable for x-ray crystallography. In Chapter 2, I describe the isolation of a large-and-conserved soluble domain of Piezo and the structural determination of this domain. In the following chapter (Chapter 3), I will describe challenges that we encountered in our effort to optimize protein expression and purification of full length Piezo, as well as our journey to arrive at a low resolution single particle electron microscope (EM) structure of this protein. Recently, the medium resolution Piezo structure, solved by single particle cryo-EM, revealed that mouse PIEZO1 assembles as a trimer where its C-terminal region forms the central core of the channel [141]; CTL2 domain was shown to make up the extracellular cap of the channel. In the last chapter of this thesis (Chapter 4), I will describe our most recent work on the isolation and characterization of a stable C-terminal fragment of mouse PIEZO1; this fragment corresponds to the central core of Piezo channel.



**Figure 1. Topology diagram of mouse PIEZO1.** Due to differences in the current membrane topology prediction algorithms for mammalian membrane protein, Piezo is variably predicted to have between 30 to 40 transmembrane segments. Here, mouse Piezo1 is predicted to have 40 putative membrane segments using the Topcons algorithm [142, 143].

## References

1. Peyronnet, R., et al., *Mechanosensitive channels: feeling tension in a world under pressure*. Front Plant Sci, 2014. **5**: p. 558.
2. Arnadottir, J. and M. Chalfie, *Eukaryotic mechanosensitive channels*. Annu Rev Biophys, 2010. **39**: p. 111-37.
3. Kung, C., B. Martinac, and S. Sukharev, *Mechanosensitive channels in microbes*. Annu Rev Microbiol, 2010. **64**: p. 313-29.
4. Ranade, S.S., R. Syeda, and A. Patapoutian, *Mechanically Activated Ion Channels*. Neuron, 2015. **87**(6): p. 1162-79.
5. Kung, C., *A possible unifying principle for mechanosensation*. Nature, 2005. **436**(7051): p. 647-54.
6. Haswell, E.S., R. Phillips, and D.C. Rees, *Mechanosensitive channels: what can they do and how do they do it?* Structure, 2011. **19**(10): p. 1356-69.
7. Sukharev, S., *Purification of the small mechanosensitive channel of Escherichia coli (MscS): the subunit structure, conduction, and gating characteristics in liposomes*. Biophys J, 2002. **83**(1): p. 290-8.
8. Gillespie, P.G. and R.G. Walker, *Molecular basis of mechanosensory transduction*. Nature, 2001. **413**(6852): p. 194-202.
9. Chalfie, M., *Neurosensory mechanotransduction*. Nat Rev Mol Cell Biol, 2009. **10**(1): p. 44-52.
10. Hudspeth, A.J. and D.P. Corey, *Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli*. Proc Natl Acad Sci U S A, 1977. **74**(6): p. 2407-11.

11. Pickles, J.O., S.D. Comis, and M.P. Osborne, *Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction.* Hear Res, 1984. **15**(2): p. 103-12.
12. Martinac, B., et al., *Pressure-sensitive ion channel in Escherichia coli.* Proc Natl Acad Sci U S A, 1987. **84**(8): p. 2297-301.
13. Sukharev, S.I., et al., *A large-conductance mechanosensitive channel in E. coli encoded by mscL alone.* Nature, 1994. **368**(6468): p. 265-8.
14. Levina, N., et al., *Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity.* EMBO J, 1999. **18**(7): p. 1730-7.
15. Booth, I.R. and P. Louis, *Managing hypoosmotic stress: aquaporins and mechanosensitive channels in Escherichia coli.* Curr Opin Microbiol, 1999. **2**(2): p. 166-9.
16. Nakamaru, Y., et al., *Mechanosensitive channel functions to alleviate the cell lysis of marine bacterium, Vibrio alginolyticus, by osmotic downshock.* FEBS Lett, 1999. **444**(2-3): p. 170-2.
17. Martinac, B., *Mechanosensitive ion channels: molecules of mechanotransduction.* J Cell Sci, 2004. **117**(Pt 12): p. 2449-60.
18. Chang, G., et al., *Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel.* Science, 1998. **282**(5397): p. 2220-6.
19. Liu, Z., C.S. Gandhi, and D.C. Rees, *Structure of a tetrameric MscL in an expanded intermediate state.* Nature, 2009. **461**(7260): p. 120-4.

20. Perozo, E., et al., *Open channel structure of MscL and the gating mechanism of mechanosensitive channels*. Nature, 2002. **418**(6901): p. 942-8.
21. Cruickshank, C.C., et al., *Estimation of the pore size of the large-conductance mechanosensitive ion channel of Escherichia coli*. Biophys J, 1997. **73**(4): p. 1925-31.
22. Dorwart, M.R., et al., *S. aureus MscL is a pentamer in vivo but of variable stoichiometries in vitro: implications for detergent-solubilized membrane proteins*. PLoS Biol, 2010. **8**(12): p. e1000555.
23. Bass, R.B., et al., *Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel*. Science, 2002. **298**(5598): p. 1582-7.
24. Lai, J.Y., et al., *Open and shut: crystal structures of the dodecylmaltoside solubilized mechanosensitive channel of small conductance from Escherichia coli and Helicobacter pylori at 4.4 Å and 4.1 Å resolutions*. Protein Sci, 2013. **22**(4): p. 502-9.
25. Naismith, J.H. and I.R. Booth, *Bacterial mechanosensitive channels--MscS: evolution's solution to creating sensitivity in function*. Annu Rev Biophys, 2012. **41**: p. 157-77.
26. Bastien, R., et al., *Unifying model of shoot gravitropism reveals proprioception as a central feature of posture control in plants*. Proc Natl Acad Sci U S A, 2013. **110**(2): p. 755-60.
27. Anava, S., et al., *The regulative role of neurite mechanical tension in network development*. Biophys J, 2009. **96**(4): p. 1661-70.

28. Sah, R.L., et al., *Biosynthetic response of cartilage explants to dynamic compression*. J Orthop Res, 1989. **7**(5): p. 619-36.
29. O'Conor, C.J., et al., *TRPV4-mediated mechanotransduction regulates the metabolic response of chondrocytes to dynamic loading*. Proc Natl Acad Sci U S A, 2014. **111**(4): p. 1316-21.
30. Yoon, A.R., et al., *COX-2 dependent regulation of mechanotransduction in human breast cancer cells*. Cancer Biol Ther, 2015. **16**(3): p. 430-7.
31. Isnard, S. and W.K. Silk, *Moving with climbing plants from Charles Darwin's time into the 21st century*. Am J Bot, 2009. **96**(7): p. 1205-21.
32. Darwin, C., *Insectivorous Plants*1875, London: John Murray.
33. Schneider, E.R., et al., *Neuronal mechanism for acute mechanosensitivity in tactile-foraging waterfowl*. Proc Natl Acad Sci U S A, 2014. **111**(41): p. 14941-6.
34. Kloda, A. and B. Martinac, *Common evolutionary origins of mechanosensitive ion channels in Archaea, Bacteria and cell-walled Eukarya*. Archaea, 2002. **1**(1): p. 35-44.
35. Jensen, G.S. and E.S. Haswell, *Functional analysis of conserved motifs in the mechanosensitive channel homolog MscS-Like2 from Arabidopsis thaliana*. PLoS One, 2012. **7**(6): p. e40336.
36. Haswell, E.S. and E.M. Meyerowitz, *MscS-like proteins control plastid size and shape in Arabidopsis thaliana*. Curr Biol, 2006. **16**(1): p. 1-11.
37. Haswell, E.S., et al., *Two MscS homologs provide mechanosensitive channel activities in the Arabidopsis root*. Curr Biol, 2008. **18**(10): p. 730-4.

38. Hamilton, E.S., A.M. Schlegel, and E.S. Haswell, *United in diversity: mechanosensitive ion channels in plants*. Annu Rev Plant Biol, 2015. **66**: p. 113-37.
39. Haswell, E.S., *MscS-Like Proteins in Plants*. Curr. Top. Membr., 2007. **58**: p. 329-59.
40. Maksaev, G. and E.S. Haswell, *MscS-Like10 is a stretch-activated ion channel from Arabidopsis thaliana with a preference for anions*. Proc Natl Acad Sci U S A, 2012. **109**(46): p. 19015-20.
41. Chalfie, M. and J. Sulston, *Developmental genetics of the mechanosensory neurons of Caenorhabditis elegans*. Dev Biol, 1981. **82**(2): p. 358-70.
42. Kellenberger, S. and L. Schild, *Epithelial sodium channel/degenerin family of ion channels: a variety of functions for a shared structure*. Physiol Rev, 2002. **82**(3): p. 735-67.
43. Adams, C.M., et al., *Ripped pocket and pickpocket, novel Drosophila DEG/ENaC subunits expressed in early development and in mechanosensory neurons*. J Cell Biol, 1998. **140**(1): p. 143-52.
44. Jasti, J., et al., *Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH*. Nature, 2007. **449**(7160): p. 316-23.
45. Ketchum, K.A., et al., *A new family of outwardly rectifying potassium channel proteins with two pore domains in tandem*. Nature, 1995. **376**(6542): p. 690-5.
46. Brohawn, S.G., *How ion channels sense mechanical force: insights from mechanosensitive K2P channels TRAAK, TREK1, and TREK2*. Ann N Y Acad Sci, 2015. **1352**: p. 20-32.

47. Enyedi, P. and G. Czirjak, *Molecular background of leak K<sup>+</sup> currents: two-pore domain potassium channels*. Physiol Rev, 2010. **90**(2): p. 559-605.
48. Brohawn, S.G., E.B. Campbell, and R. MacKinnon, *Physical mechanism for gating and mechanosensitivity of the human TRAAK K<sup>+</sup> channel*. Nature, 2014. **516**(7529): p. 126-30.
49. Brohawn, S.G., E.B. Campbell, and R. MacKinnon, *Domain-swapped chain connectivity and gated membrane access in a Fab-mediated crystal of the human TRAAK K<sup>+</sup> channel*. Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2129-34.
50. Brohawn, S.G., J. del Marmol, and R. MacKinnon, *Crystal structure of the human K2P TRAAK, a lipid- and mechano-sensitive K<sup>+</sup> ion channel*. Science, 2012. **335**(6067): p. 436-41.
51. Dong, Y.Y., et al., *K2P channel gating mechanisms revealed by structures of TREK-2 and a complex with Prozac*. Science, 2015. **347**(6227): p. 1256-9.
52. Clapham, D.E., et al., *International Union of Pharmacology. XLIX. Nomenclature and structure-function relationships of transient receptor potential channels*. Pharmacol Rev, 2005. **57**(4): p. 427-50.
53. Montell, C., *The TRP superfamily of cation channels*. Sci STKE, 2005. **2005**(272): p. re3.
54. Eijkelkamp, N., K. Quick, and J.N. Wood, *Transient receptor potential channels and mechanosensation*. Annu Rev Neurosci, 2013. **36**: p. 519-46.
55. Clapham, D.E., *TRP channels as cellular sensors*. Nature, 2003. **426**(6966): p. 517-24.

56. Moiseenkova-Bell, V.Y., et al., *Structure of TRPV1 channel revealed by electron cryomicroscopy*. Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7451-5.
57. Cao, E., et al., *TRPV1 structures in distinct conformations reveal activation mechanisms*. Nature, 2013. **504**(7478): p. 113-8.
58. Liao, M., et al., *Structure of the TRPV1 ion channel determined by electron cryomicroscopy*. Nature, 2013. **504**(7478): p. 107-12.
59. Paulsen, C.E., et al., *Structure of the TRPA1 ion channel suggests regulatory mechanisms*. Nature, 2015. **525**(7570): p. 552.
60. Lee, G., et al., *Nanospring behaviour of ankyrin repeats*. Nature, 2006. **440**(7081): p. 246-9.
61. Zhang, W., et al., *Ankyrin Repeats Convey Force to Gate the NOMPC Mechanotransduction Channel*. Cell, 2015. **162**(6): p. 1391-403.
62. Corey, D.P., et al., *TRPA1 is a candidate for the mechanosensitive transduction channel of vertebrate hair cells*. Nature, 2004. **432**(7018): p. 723-30.
63. Kwan, K.Y., et al., *TRPA1 contributes to cold, mechanical, and chemical nociception but is not essential for hair-cell transduction*. Neuron, 2006. **50**(2): p. 277-89.
64. Bautista, D.M., et al., *TRPA1 mediates the inflammatory actions of environmental irritants and proalgesic agents*. Cell, 2006. **124**(6): p. 1269-82.
65. Kurima, K., et al., *Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function*. Nat Genet, 2002. **30**(3): p. 277-84.

66. Kawashima, Y., et al., *Mechanotransduction in mouse inner ear hair cells requires transmembrane channel-like genes*. J Clin Invest, 2011. **121**(12): p. 4796-809.
67. Vreugde, S., et al., *Beethoven, a mouse model for dominant, progressive hearing loss DFNA36*. Nat Genet, 2002. **30**(3): p. 257-8.
68. Labay, V., et al., *Topology of transmembrane channel-like gene 1 protein*. Biochemistry, 2010. **49**(39): p. 8592-8.
69. Coste, B., et al., *Piezol1 and Piezo2 are essential components of distinct mechanically activated cation channels*. Science, 2010. **330**(6000): p. 55-60.
70. Ernstrom, G.G. and M. Chalfie, *Genetics of sensory mechanotransduction*. Annu Rev Genet, 2002. **36**: p. 411-53.
71. Delmas, P. and B. Coste, *Mechano-gated ion channels in sensory systems*. Cell, 2013. **155**(2): p. 278-84.
72. Coste, B., et al., *Piezo proteins are pore-forming subunits of mechanically activated channels*. Nature, 2012. **483**(7388): p. 176-81.
73. Heginbotham, L. and R. MacKinnon, *Conduction properties of the cloned Shaker K<sup>+</sup> channel*. Biophys J, 1993. **65**(5): p. 2089-96.
74. Conforti, L., et al., *O<sub>2</sub>-sensitive K<sup>+</sup> channels: role of the Kv1.2 -subunit in mediating the hypoxic response*. J Physiol, 2000. **524 Pt 3**: p. 783-93.
75. Conforti, L. and D.E. Millhorn, *Selective inhibition of a slow-inactivating voltage-dependent K<sup>+</sup> channel in rat PC12 cells by hypoxia*. J Physiol, 1997. **502 ( Pt 2)**: p. 293-305.

76. Grissmer, S., et al., *Pharmacological characterization of five cloned voltage-gated K<sup>+</sup> channels, types Kv1.1, 1.2, 1.3, 1.5, and 3.1, stably expressed in mammalian cell lines*. Mol Pharmacol, 1994. **45**(6): p. 1227-34.
77. Hille, B., *Ion Channels of Excitable Membranes*. Third edition ed2001, Sunderland, CT: Sinauer Associates.
78. Prole, D.L. and C.W. Taylor, *Identification and analysis of putative homologues of mechanosensitive channels in pathogenic protozoa*. PLoS One, 2013. **8**(6): p. e66068.
79. Hedrich, R., *Ion channels in plants*. Physiol Rev, 2012. **92**(4): p. 1777-811.
80. Kim, S.E., et al., *The role of Drosophila Piezo in mechanical nociception*. Nature, 2012. **483**(7388): p. 209-12.
81. Gorczyca, D.A., et al., *Identification of Ppk26, a DEG/ENaC Channel Functioning with Ppk1 in a Mutually Dependent Manner to Guide Locomotion Behavior in Drosophila*. Cell Rep, 2014. **9**(4): p. 1446-58.
82. Guo, Y., et al., *The role of PPK26 in Drosophila larval mechanical nociception*. Cell Rep, 2014. **9**(4): p. 1183-90.
83. Zhong, L., R.Y. Hwang, and W.D. Tracey, *Pickpocket is a DEG/ENaC protein required for mechanical nociception in Drosophila larvae*. Curr Biol, 2010. **20**(5): p. 429-34.
84. Satoh, K., et al., *A novel membrane protein, encoded by the gene covering KIAA0233, is transcriptionally induced in senile plaque-associated astrocytes*. Brain Res, 2006. **1108**(1): p. 19-27.

85. Blumenthal, N.R., et al., *Stochastic nanoroughness modulates neuron-astrocyte interactions and function via mechanosensing cation channels*. Proc Natl Acad Sci U S A, 2014. **111**(45): p. 16124-9.
86. Pathak, M.M., et al., *Stretch-activated ion channel Piezo1 directs lineage choice in human neural stem cells*. Proc Natl Acad Sci U S A, 2014. **111**(45): p. 16148-53.
87. Dupont, S., et al., *Role of YAP/TAZ in mechanotransduction*. Nature, 2011. **474**(7350): p. 179-83.
88. Zhao, B., et al., *The Hippo-YAP pathway in organ size control and tumorigenesis: an updated version*. Genes Dev, 2010. **24**(9): p. 862-74.
89. McHugh, B.J., et al., *Integrin activation by Fam38A uses a novel mechanism of R-Ras targeting to the endoplasmic reticulum*. J Cell Sci, 2010. **123**(Pt 1): p. 51-61.
90. McHugh, B.J., et al., *Loss of the integrin-activating transmembrane protein Fam38A (Piezo1) promotes a switch to a reduced integrin-dependent mode of cell migration*. PLoS One, 2012. **7**(7): p. e40346.
91. Hughes, C.S., L.M. Postovit, and G.A. Lajoie, *Matrigel: a complex protein mixture required for optimal growth of cell culture*. Proteomics, 2010. **10**(9): p. 1886-90.
92. Benton, G., et al., *Advancing science and technology via 3D culture on basement membrane matrix*. J Cell Physiol, 2009. **221**(1): p. 18-25.
93. Jin, Y., et al., *Functional role of mechanosensitive ion channel Piezo1 in human periodontal ligament cells*. Angle Orthod, 2015. **85**(1): p. 87-94.
94. Kanzaki, H., et al., *Periodontal ligament cells under mechanical stress induce osteoclastogenesis by receptor activator of nuclear factor kappaB ligand up-*

- regulation via prostaglandin E2 synthesis.* J Bone Miner Res, 2002. **17**(2): p. 210-20.
95. Kim, S.J., et al., *Compressive stress induced the up-regulation of M-CSF, RANKL, TNF-alpha expression and the down-regulation of OPG expression in PDL cells via the integrin-FAK pathway.* Arch Oral Biol, 2013. **58**(6): p. 707-16.
96. Bowman, C.L., et al., *Mechanosensitive ion channels and the peptide inhibitor GsMTx-4: history, properties, mechanisms and pharmacology.* Toxicon, 2007. **49**(2): p. 249-70.
97. Eisenhoffer, G.T., et al., *Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia.* Nature, 2012. **484**(7395): p. 546-9.
98. Li, J., et al., *Piezol integration of vascular architecture with physiological force.* Nature, 2014. **515**(7526): p. 279-82.
99. Ranade, S.S., et al., *Piezol, a mechanically activated ion channel, is required for vascular development in mice.* Proc Natl Acad Sci U S A, 2014. **111**(28): p. 10347-52.
100. Faucherre, A., et al., *Piezol plays a role in erythrocyte volume homeostasis.* Haematologica, 2014. **99**(1): p. 70-5.
101. Eijkelkamp, N., et al., *A role for Piezo2 in EPAC1-dependent mechanical allodynia.* Nat Commun, 2013. **4**: p. 1682.
102. Hucho, T. and J.D. Levine, *Signaling pathways in sensitization: toward a nociceptor cell biology.* Neuron, 2007. **55**(3): p. 365-76.
103. Bos, J.L., *Epac proteins: multi-purpose cAMP targets.* Trends Biochem Sci, 2006. **31**(12): p. 680-6.

104. Dubin, A.E., et al., *Inflammatory signals enhance piezo2-mediated mechanosensitive currents*. Cell Rep, 2012. **2**(3): p. 511-7.
105. Steranka, L.R., et al., *Bradykinin as a pain mediator: receptors are localized to sensory neurons, and antagonists have analgesic actions*. Proc Natl Acad Sci U S A, 1988. **85**(9): p. 3245-9.
106. Leeb-Lundberg, L.M., et al., *International union of pharmacology. XLV. Classification of the kinin receptor family: from molecular mechanisms to pathophysiological consequences*. Pharmacol Rev, 2005. **57**(1): p. 27-77.
107. Mizumura, K., et al., *Excitation and sensitization of nociceptors by bradykinin: what do we know?* Exp Brain Res, 2009. **196**(1): p. 53-65.
108. Ikeda, R., et al., *Merkel cells transduce and encode tactile stimuli to drive Abeta-afferent impulses*. Cell, 2014. **157**(3): p. 664-75.
109. Ikeda, R. and J.G. Gu, *Piez02 channel conductance and localization domains in Merkel cells of rat whisker hair follicles*. Neurosci Lett, 2014. **583**: p. 210-5.
110. Woo, S.H., et al., *Piez02 is required for Merkel-cell mechanotransduction*. Nature, 2014. **509**(7502): p. 622-6.
111. Ranade, S.S., et al., *Piez02 is the major transducer of mechanical forces for touch sensation in mice*. Nature, 2014. **516**(7529): p. 121-5.
112. Schrenk-Siemens, K., et al., *PIEZO2 is required for mechanotransduction in human stem cell-derived touch receptors*. Nat Neurosci, 2015. **18**(1): p. 10-6.
113. Faucherre, A., et al., *piezo2b regulates vertebrate light touch response*. J Neurosci, 2013. **33**(43): p. 17089-94.

114. Lee, W., et al., *Synergy between Piezol1 and Piezo2 channels confers high-strain mechanosensitivity to articular cartilage*. Proc Natl Acad Sci U S A, 2014. **111**(47): p. E5114-22.
115. Volkers, L., Y. Mechtioukhi, and B. Coste, *Piezo channels: from structure to function*. Pflugers Arch, 2015. **467**(1): p. 95-9.
116. Zarychanski, R., et al., *Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis*. Blood, 2012. **120**(9): p. 1908-15.
117. Albuission, J., et al., *Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels*. Nat Commun, 2013. **4**: p. 1884.
118. Andolfo, I., et al., *Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1*. Blood, 2013. **121**(19): p. 3925-35, S1-12.
119. Bruce, L.J., *Hereditary stomatocytosis and cation-leaky red cells--recent developments*. Blood Cells Mol Dis, 2009. **42**(3): p. 216-22.
120. McMillin, M.J., et al., *Mutations in PIEZO2 cause Gordon syndrome, Marden-Walker syndrome, and distal arthrogryposis type 5*. Am J Hum Genet, 2014. **94**(5): p. 734-44.
121. Coste, B., et al., *Gain-of-function mutations in the mechanically activated ion channel PIEZO2 cause a subtype of Distal Arthrogryposis*. Proc Natl Acad Sci U S A, 2013. **110**(12): p. 4667-72.
122. Gordon, H., D. Davies, and M. Berman, *Camptodactyly, cleft palate, and club foot. A syndrome showing the autosomal-dominant pattern of inheritance*. J Med Genet, 1969. **6**(3): p. 266-74.

123. Ioan, D.M., et al., *Distal arthrogryposis with autosomal dominant inheritance and reduced penetrance in females: the Gordon syndrome*. Clin Genet, 1993. **43**(6): p. 300-2.
124. Say, B., et al., *The Gordon syndrome*. J Med Genet, 1980. **17**(5): p. 405.
125. Bamshad, M., A.E. Van Heest, and D. Pleasure, *Arthrogryposis: a review and update*. J Bone Joint Surg Am, 2009. **91 Suppl 4**: p. 40-6.
126. Gottlieb, P.A. and F. Sachs, *Piezo1: properties of a cation selective mechanical channel*. Channels (Austin), 2012. **6**(4): p. 214-9.
127. Bae, C., P.A. Gottlieb, and F. Sachs, *Human PIEZO1: removing inactivation*. Biophys J, 2013. **105**(4): p. 880-6.
128. Gnanasambandam, R., et al., *Ionic Selectivity and Permeation Properties of Human PIEZO1 Channels*. PLoS One, 2015. **10**(5): p. e0125503.
129. Kamajaya, A., et al., *The structure of a conserved piezo channel domain reveals a topologically distinct beta sandwich fold*. Structure, 2014. **22**(10): p. 1520-7.
130. Bae, C., et al., *Xerocytosis is caused by mutations that alter the kinetics of the mechanosensitive channel PIEZO1*. Proc Natl Acad Sci U S A, 2013. **110**(12): p. E1162-8.
131. Poole, K., et al., *Tuning Piezo ion channels to detect molecular-scale movements relevant for fine touch*. Nat Commun, 2014. **5**: p. 3520.
132. Huang, M., et al., *A stomatin-like protein necessary for mechanosensation in C. elegans*. Nature, 1995. **378**(6554): p. 292-5.

133. O'Hagan, R., M. Chalfie, and M.B. Goodman, *The MEC-4 DEG/ENaC channel of Caenorhabditis elegans touch receptor neurons transduces mechanical signals.* Nat Neurosci, 2005. **8**(1): p. 43-50.
134. Wetzel, C., et al., *A stomatin-domain protein essential for touch sensation in the mouse.* Nature, 2007. **445**(7124): p. 206-9.
135. Peyronnet, R., et al., *Piezol-dependent stretch-activated channels are inhibited by Polycystin-2 in renal tubular epithelial cells.* EMBO Rep, 2013. **14**(12): p. 1143-8.
136. Delmas, P., *Polycystins: from mechanosensation to gene regulation.* Cell, 2004. **118**(2): p. 145-8.
137. Nauli, S.M. and J. Zhou, *Polycystins and mechanosensation in renal and nodal cilia.* Bioessays, 2004. **26**(8): p. 844-56.
138. Patel, A. and E. Honore, *Polycystins and renovascular mechanosensory transduction.* Nat Rev Nephrol, 2010. **6**(9): p. 530-8.
139. Weinbaum, S., et al., *Mechanotransduction in the renal tubule.* Am J Physiol Renal Physiol, 2010. **299**(6): p. F1220-36.
140. Syeda, R., et al., *Chemical activation of the mechanotransduction channel Piezol.* Elife, 2015. **4**.
141. Ge, J., et al., *Architecture of the mammalian mechanosensitive Piezol channel.* Nature, 2015.
142. Bernsel, A., et al., *TOPCONS: consensus prediction of membrane protein topology.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W465-8.
143. Tsirigos, K.D., et al., *The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides.* Nucleic Acids Res, 2015.

# **Chapter 2:**

## **The Structure of a Conserved Piezo Channel Domain Reveals a Novel Beta Sandwich Fold**

Published as Kamajaya, A., Kaiser, J.T., Lee, J., Reid, M., Rees, D.C. (2014) The structure of a conserved piezo channel domain reveals a topologically distinct  $\beta$  sandwich fold. Structure 22(10):1520-7. doi: 10.1016/j.str.2014.08.009.

*Chapter 2 : The Structure of a Conserved Piezo Channel Domain Reveals a Novel Beta Sandwich Fold*

## Summary

Piezo has recently been identified as a family of eukaryotic mechanosensitive channels composed of subunits containing over 2000 amino acids, without recognizable sequence similarity to other types of channels. Here, we present the crystal structure of a large, conserved extracellular domain located just before the last predicted transmembrane helix of *C. elegans* PIEZO, which adopts a novel beta sandwich fold. The structure was also determined of a point mutation located in this domain at the equivalent position to the mutation in human PIEZO1 found in the Dehydrated Hereditary Stomatocytosis (DHS) patients (M2225R). While the point mutation does not change the overall domain structure, it does alter the surface electrostatic potential, which may perturb the ability of this domain to interact with a yet-to-be identified ligand or protein. The lack of structural similarity between this domain and any previously characterized fold, including those of eukaryotic and bacterial channels, highlights the distinctive nature of the Piezo family of eukaryotic mechanosensitive channels.

## Highlights

- A conserved loop of the Piezo channel is identified that forms a stable domain.
- The structure of this domain reveals a novel beta sandwich fold.
- The novel fold highlights the distinctive nature of the Piezo channel family.
- A disease related mutation in this domain has minimal structural consequences.

## **Introduction**

The ability to translate environmental cues (chemical, electrical, and mechanical) into intracellular signals is crucial to the functioning of all cells. Although membrane protein receptors for many of these environmental signals have been identified, our understanding of the receptors for mechanical signal transduction remains incomplete. The best characterized models of mechanotransduction are bacterial mechanosensitive (MS) channels that protect cells from osmotic downshock [1]. In multicellular eukaryotes, mechanosensitive channels are proposed to be involved in diverse physiological and developmental processes, such as somatosensory and auditory detection, change in blood osmolarity, muscle stretch, and others [2]. The identity of the MS channels involved in these physiological processes remains contentious, however.

A notable exception to the unknown molecular identities of eukaryotic mechanosensors is provided by the Piezo family of mechanosensitive channels [3]. Piezo was first identified through an siRNA knock-down screen and was shown to be necessary and sufficient for mechanically-activated currents [4]. Piezos have been found in both multicellular and unicellular eukaryotes with the exception of yeast [4, 5]. A striking characteristic of Piezo is the significant size, with all characterized homologs (human, mice, and *Drosophila*) containing over 2000 amino acids; moreover, mouse PIEZO1 has been demonstrated to form a tetrameric complex corresponding to a total molecular weight for the channel of ~1 MDa [3]. No sequence similarities have been recognized between Piezo and other types of channels, nor have internal sequence repeats within the Piezo subunit been detected. PIEZO1 can be inhibited by a general blocker of mechanosensitive channels, GsMTx4 [6], and is shown to be sensitive to ruthenium red [3, 4]. Piezo

channels are involved in several physiological processes, such as maintaining normal cell density in zebrafish and mammalian epithelial tissue [7], mediating mechanical nociception in *Drosophila* larvae [8], and light touch in Rohon-Beard neurons in zebrafish [9]. In humans, point mutations in PIEZO1 have been identified in patients suffering from Dehydrated Hereditary Stomatocytosis (DHS), an autosomal dominant hemolytic anemia disease [10-12]. In addition, inactivation of PIEZO1 in lung epithelial cells causes integrin-independent amoeboid cell migration [13].

Understanding the gating mechanism of Piezo will enrich our current knowledge of how mechanical cues are converted into intracellular signals. A challenge in developing a gating model for Piezo arises from the absence of any structural information on the channel. Both the large size and the eukaryotic nature of Piezo create formidable technical challenges to the direct crystallographic analysis of the intact channel. The substantial size suggests, however, that the channel likely consists of smaller domains that may be more amenable to study. Here, we present the crystal structure at 2.5 Å resolution of a Piezo channel soluble domain designated CTL2 (C-Terminal Loop 2), located just before the last transmembrane helix and the C-terminal tail. This loop is predicted to be the largest soluble domain that is conserved across Piezo homologs. Moreover, there is accumulating evidence that the C-terminal portion of Piezo, including the CTL2, is involved in the gating kinetics of the channel. One of the human PIEZO1 point mutants found in DHS patients (M2225R) maps onto this loop and is shown to exhibit a delayed activation as well as slower inactivation [14]. A combination of a naturally-occurring mutation (M2225R) and an artificial mutation (R2456K in the C-terminal loop following the last transmembrane helix) in a human PIEZO1 double mutant construct eliminates the inactivation of the channel

[15]. We also present the structure of *C. elegans* PIEZO CTL2 domain carrying a point mutation (M31R) at the equivalent position to the mutation found in DHS patients (M2225R).

## **Results**

### **Construct Selection**

As a starting point for structural analysis of the Piezo channel, we targeted large, well-conserved extramembrane loops. To gain a broad overview of the threading of Piezo protein through the membrane, we used multiple topology prediction web servers: TMHMM2 [16], Phobius [17], TOPCONS [18], TMPred [19], and SMART [20]. Although similar trends were evident, different algorithms predicted different numbers of transmembrane helices (between 30 and 40) for all characterized Piezos and there were differences in the predicted orientation of some of the extramembrane loops (Fig. 1). For our structural study, we selected a Piezo loop that satisfied the following criteria:

1. it is consistently predicted by all the topology-prediction algorithms used,
2. it is larger than 100 amino acid residues,
3. it is highly conserved across the Piezo family.

Under these constraints, we identified a predicted soluble domain corresponding to mouse PIEZO1 residues 2210 – 2453. These residues form an extramembrane loop preceding the last predicted transmembrane helix that we designate as the C-terminal loop 2 (CTL2) domain. 3 of the 4 topology prediction web servers predict that CTL2 is oriented in the extracellular environment (Fig. 1). An alignment of Piezo homologs using MAFFT (Fig S1; [21]) reveals that CTL2 contains over 200 amino acid residues with a calculated

molecular weight between 24 – 32 kDa, making CTL2 the largest conserved loop among the predicted loops of the Piezo channel. In addition, the Phyre2 webserver [22] predicts that CTL2 has well-defined secondary structure (Fig S2). Therefore, we selected CTL2 for crystallographic analysis.

#### Architecture of the *C. elegans* CTL2 Domain

Following a screen of various Piezo orthologues (see Materials and Methods), the *C. elegans* CTL2 domain was found to be amenable for crystallographic analysis. The structure of wild type *C. elegans* CTL2 establishes that this domain is organized around a core of three beta sheets flanked by two pairs of helices with overall dimensions of ~60 Å x 40 Å x 30 Å (Fig 2). The secondary structure observed in the crystal structure generally matches that predicted by the Phyre2 webserver (Fig S2). The three beta sheets are arranged such that sheet 1 (strands 1, 2, 5, 6, with strands numbered from the N-terminus) is nearly parallel to sheet 2 (strands 3, 4, 11, 8), and perpendicular to sheet 3 (strands 7, 9, 10). Interestingly, analysis of residue conservation done with Profunc [23, 24] and Consurf [25] reveals that CTL2 possess conserved patch on one side but not the other (Fig 3). Analysis done by Profunc is more automated, such that residue conservation is calculated using 50 of the closest *C. elegans* CTL2 homologs as obtained by BLAST search against Uniprot database. Consequently, Profunc shows a higher residue conservation at this patch compared to calculation done by Consurf.

To identify whether any protein of known structure shares a similar fold with CTL2, we used SSM [26] as well as the DALI server [27] to perform a structure-based alignment. None of the identified proteins exhibited close similarities to CTL2 (Max z-

score: 2.8). Most of these candidate proteins possess a beta sandwich fold composed of two beta sheets; this is different than our Piezo soluble domain structure that consists of three beta sheets. However, if we consider sheet 1 and sheet 2 of CTL2 as the beta sandwich core with the rest of the protein as connectors, we obtain a two beta sheet sandwich core with a 2D connectivity diagram illustrated in (Fig S3a). A comparison between the Piezo loop structure and the identified proteins did not reveal any significant connectivity similarity (Fig S3b). Subsequently, we explored the structural classification database CATH [28] for two-sheet beta sandwich containing proteins (CATH ID: 2.60). Examination of the strand order in the beta sheets of these proteins failed to identify any close matches between CTL2 and previously solved structures. Therefore, we conclude that Piezo C-terminal soluble domain exhibits a novel beta sandwich fold.

#### Structure of M31R *C. elegans* PIEZO CTL2 Domain Variant

The single substitution in human PIEZO1 of methionine 2225 to arginine has been linked to the disease Dehydration Hereditary Stomatocytosis (DHS) [10]. From the multiple sequence alignment of Piezo homologs, we determined that M31 of our *C.elegans* PIEZO CTL2 construct, located in strand 2 of sheet 1, is equivalent to residue M2225 in human PIEZO1 (Fig. S4). This residue is located at the conserved patch of CTL2 domain (Fig. 3). The M31R *C.elegans* mutant CTL2 protein crystallized in a distinct crystal form from the WT domain and the structure was solved by molecular replacement. Overall, the structure of the mutant CTL2 is similar to the WT structure with some minor rearrangements (RMSD C $\alpha$  = 1.3 Å overall, with 0.7 Å for residues in alpha-helices and beta-strands (Fig. 4a)). In *C.elegans* PIEZO CTL2, M31 is spatially adjacent to a

neighboring arginine residue, R94 (Fig. 4b). The most significant change in the M31R mutant CTL2 structure is the shift of the R94 sidechain ~3 Å away from R31, presumably due to electrostatic repulsion (Fig. 4b,c, Fig S5). However, R94 in *C. elegans* PIEZO loop is not conserved in the multiple sequence alignment (Fig. S4). R94 aligns to a glutamate residue that is primarily conserved in the vertebrate Piezo1 and Piezo2. Even though the *C. elegans* PIEZO CTL2 does not have a glutamate at this position, there are three glutamates in proximity to M31: E30, E32, and E98. This results in a net acidic local environment surrounding M31 in the *C. elegans* PIEZO.

## **Discussion**

As a starting point for the structural analysis of Piezo, we have identified a predicted extracellular domain designated as CTL2 (C-terminal Loop 2) as a promising target for three reasons. (1) There is a moderate degree of sequence conservation in this region through all Piezo homologs, which may imply its significance in the general protein function. (2) We were looking for a target loop with at least 100 amino acids because larger domains are more likely to have a stable core structure. With ~250 residues, CTL2 is the largest conserved extramembrane loop found among Piezo homologues. Moreover, the boundaries at which this loop starts and ends as defined by different topology prediction methods are similar, simplifying the design of our constructs. (3) A point mutation in the CTL2 domain of human PIEZO1 is found in DHS patients. Removal of this loop, together with the last transmembrane segment and the C-terminal tail, severely impaired the inactivation kinetics of the channel [14], leading to an increased ion influx that may contribute to the disease phenotype [10-12].

The model of CTL2 establishes that the region forms a compact domain containing a beta sandwich with three sheets. The CTL2 fold is apparently unrelated to any other family of proteins, including the  $\beta$  sheet domains present in other eukaryotic channels, specifically the extracellular domains of the two eukaryotic ion channels of known structure [29, 30], ASIC (CATH ID: 2.60.470.10) and P2X4 (CATH ID: 2.60.490.10). There is no evidence from either structure of the wild type or M31R mutant that this domain forms a tetramer that might help stabilize the oligomeric state of the intact channel. Mutating the conserved methionine residue into arginine (M31R), corresponding to a human DHS mutation (M2225R), does not cause a significant conformational change in *C. elegans* PIEZO CTL2 domain, suggesting that the effect of this mutation may be to more subtly perturb conformational energetics through changes in electrostatics, or perhaps to interfere with the ability to interact with other domains by changing the interface region.

In conclusion, we have described the first structural characterization of a domain of a eukaryotic mechanosensitive channel, the CTL2 domain of *C. elegans* PIEZO. We have shown that this extramembrane region forms a stable domain that adopts a previously uncharacterized beta sandwich fold. The functional role of this domain in the context of the full length channel remains unclear. In view of the monomeric nature of the CTL2 domain, it is unlikely to mediate tetramerization of the full length channel. As this domain is generally conserved across Piezo homologs, it may have a universal role such as mediating proper intracellular trafficking, tethering the intact channel to an unknown extracellular protein, or binding of a ligand; the search for binding partners is currently under

investigation. The lack of structural similarity between this domain and any known protein fold, including those of eukaryotic and bacterial channels, reinforces the lack of recognizable sequence homology between Piezo and other channels, and highlights the distinctive nature of this enigmatic family of eukaryotic mechanosensitive channels.

### **Experimental Procedures**

#### Protein cloning and expression

The CTL2 regions from the mouse PIEZO1, mouse PIEZO2, human PIEZO1, *Drosophila* PIEZO and *C. elegans* PIEZO orthologues were cloned into a pET21 vector between the NdeI and XhoI restriction sites, generating a construct with a C-terminal 6xHis tag. The plasmid DNA was transformed into BL-21 gold *E. coli* cells (Stratagene). Cells were grown in Terrific Broth medium until they reached an OD<sub>600</sub> of 1.0, when 1 mM IPTG was added. Protein induction was carried on overnight at 20 °C. The CTL2 of *C. elegans* PIEZO (GenBank ID: 392900032) was found to be best suited for structural analysis; the protein and nucleotide sequences of the coding region for this construct are:

MSLLNQIGTISMPEKVTLRISIEGYPPLYEMEAQGSNHDNAELGMIKPDQLASLNQ  
ALTDSYTTTRDTNSILRSRMSVSYLKGTYEDILIVRFRPESEIYWPISQDSRNAMID  
KLSRNTSVNFEVSLEFTRPYDPNENAALKHSKSWLVPISLDMTIRAKIQSALRGDP  
GHPILIPQSIPAFIQVPNQGELTLPTSIGNTIINDGNPRINTTGMEKSDEARAWFDSLTL  
LNLEQGKSQNEKMWIATSEHPGDQNAKLWIKTANTTYSGRPYLQVVGFIDRAFPS  
LEHHHHHH

**Protein residue numbers are assigned according to their position in this CTL2 construct.**

ATGTCGCTACTCAAATCGGAACATCAGTATGCCAGAAAAAGTTACGT  
TGAGGATTTCGATTGAAGGATATCCACCACTTACGAGATGGAGGCACAAGG  
AAGTAATCACGATAACGCTGAACTTGGAATGATAAAACCTGATCAATTGGCAT  
CTTGAAATCAAGCATTGACGGATAGTTACACGACACGAGACACTAATTCTATT  
TTAAGATCAAGAATGTCTGTATCATATCTGAAAGGATACACATATGAGGATAT  
TCTGATTGTGCGATTCCGTCCAGAACATCTGAAATTATTGGCCGATTCACAGGA  
TTCGAGAAATGCGATGATCGATAAATTGAGCCGAAACACGTCAGTCAATTG  
AAGTATCTCTGGAATTCACTCGGCCATATGATCCAATGAAAACGCTGCTTA  
AACACATTGAAATCGTGGCTTCCAATCTCCCTGGATATGACGATTGAGC  
AAAAATTCAaAGTGCCTCCGAGGAGATCCAGGTCACTCGATTCTCATTCCACA  
GTCAATTCCCTGCGTTATTCAAGTTCAAATCAAGGAGAATTGACCCTCCAAC  
ATCAATCGAAATACTATTCAATGACGGAAATCCACGGATTAATACGACCG  
GGATGGAAAAATCAGACGAAGCTCGGGCTGGTCGACTCGTTAACCTTGAAT  
CTCGAACAGGAAAATCGCAAAACGAGAAAATGTGGATCGCCACATCCGAAC  
ATCCAGGAGACCAAAATGCGAAACTCTGGATAAAGACTGCAAATACGACGTA  
TTCGGGAAGGCCGTACCTCAAGTTGTCGGATTATTGATAGACCATTCCCATC  
ACTCGAGCACCACCACCACTGA

### Protein purification

Cells were homogenized in lysis buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, lysozyme, DNase, PMSF) at a ratio of 1 g cells / 10 mL of buffer. Cells were lysed using a microfluidizer (Microfluidics, Model M-110L), and the lysate was centrifuged to remove insoluble material. The protein was purified in three steps. First, lysate was run through a NiNTA column, washed with 40 mM imidazole for 15 column volumes, and eluted with 250 mM imidazole. Next, the eluate from the NiNTA column was run through a Superdex-200 16/60 column (GE Healthcare), and the corresponding peak fraction was collected. Finally, the resulting protein fraction was run through an anion exchange column, HiTrap Q Sepharose FF (GE Healthcare), to obtain the final protein preparation. The identities of the purified wild type (WT) and M31R variant proteins were confirmed by mass spectrometry.

### Crystallization & model refinement

WT CTL2 was crystallized by the sitting drop vapor diffusion method against a reservoir containing 0.1 M phosphate-citrate buffer pH 4.0, 26% PEG1000, and 0.2 M Li<sub>2</sub>SO<sub>4</sub> at 20 °C in a 96-well plate, at a ratio 0.2 µL protein to 0.2 µL reservoir drop, using intelli sitting drop crystallization plates (ArtRobbins). The protein concentration was 10-12 mg/mL. WT CTL2 crystals appeared after 18 hrs and grew to maturation after 3 days to approximate dimensions of 50 µm x 50 µm x 150 µm. M31R CTL2 was crystallized against a reservoir containing 0.1 M phosphate-citrate buffer pH 4.8, 30% PEG3350, and 0.2 M Li<sub>2</sub>SO<sub>4</sub> at 20 °C using similar procedures. These crystals grew more slowly than the WT crystals.

A heavy atom derivative of the WT CTL2 crystals was prepared by soaking in 10 mM potassium tetrani troplatinate (II) for three days, followed by back-soaking in the mother liquor for 2 hrs. Protein crystals were flash frozen in the presence of Fomblin Y 16/6 oil (Sigma), and were sent to Stanford Synchrotron Radiation Lightsource (SSRL). Diffraction data were collected at SSRL beamline 12-2 equipped with a PILATUS 6M PAD detector. Diffraction patterns were indexed and integrated using iMOSFLM [31] from the CCP4 software package [32] and XDS [33]. The integrated data was scaled using SCALA [34]. Phases for the WT CTL2 were obtained experimentally by Single-wavelength Anomalous Dispersion (SAD) phasing based on the platinum derivative. Five platinum sites were located, and electron density maps were generated using Autosol [35] from the PHENIX suite [36]. Phases for the M31R CTL2 structure were generated with PhaserMR [37] by molecular replacement using the WT structure. Model building was done in Coot [38], and the structure refinements were done using Refmac5 [39]. The final molecular structures were displayed using PyMOL [40].

**Acknowledgments:**

We thank A. Patapoutian for providing the *C. elegans* PIEZO construct, S. Hess from the Proteome Exploration Laboratory at the Beckman Institute at the California Institute of Technology for mass spectrometry analysis of the protein crystals, and P. Nikolovski from the Molecular Observatory at the Beckman Institute at the California Institute of Technology. This research was supported by National Institute of Health (NIH GM84211). We acknowledge the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program and the Beckman Institute at Caltech for their generous support of the Molecular Observatory at Caltech. Operations at SSRL are supported by the US DOE and NIH.

**Authors Contributions**

A.K. designed the experiments, performed most of the experiments, analyzed the results, and wrote the manuscript. J.T.K. and J.L. provided mentorship in x-ray diffraction data processing. M.R. performed the early protein purification. D.C.R provided the mentorship, funding, and editorial service.

**References:**

1. Levina, N., et al., *Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity.* EMBO J, 1999. **18**(7): p. 1730-7.
2. Chalfie, M., *Neurosensory mechanotransduction.* Nat Rev Mol Cell Biol, 2009. **10**(1): p. 44-52.
3. Coste, B., et al., *Piezo proteins are pore-forming subunits of mechanically activated channels.* Nature, 2012. **483**(7388): p. 176-81.
4. Coste, B., et al., *Piez01 and Piez02 are essential components of distinct mechanically activated cation channels.* Science, 2010. **330**(6000): p. 55-60.
5. Prole, D.L. and C.W. Taylor, *Identification and analysis of putative homologues of mechanosensitive channels in pathogenic protozoa.* PLoS One, 2013. **8**(6): p. e66068.
6. Bae, C., F. Sachs, and P.A. Gottlieb, *The mechanosensitive ion channel Piez01 is inhibited by the peptide GsMTx4.* Biochemistry, 2011. **50**(29): p. 6295-300.
7. Eisenhoffer, G.T., et al., *Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia.* Nature, 2012. **484**(7395): p. 546-9.
8. Kim, S.E., et al., *The role of Drosophila Piezo in mechanical nociception.* Nature, 2012. **483**(7388): p. 209-12.
9. Faucherre, A., et al., *piezo2b Regulates Vertebrate Light Touch Response.* Journal of Neuroscience, 2013. **33**(43): p. 17089-94.

10. Zarychanski, R., et al., *Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis*. Blood, 2012. **120**(9): p. 1908-15.
11. Albuison, J., et al., *Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels*. Nat Commun, 2013. **4**: p. 1884.
12. Andolfo, I., et al., *Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1*. Blood, 2013. **121**(19): p. 3925-35, S1-12.
13. McHugh, B.J., et al., *Loss of the integrin-activating transmembrane protein Fam38A (Piezo1) promotes a switch to a reduced integrin-dependent mode of cell migration*. PLoS One, 2012. **7**(7): p. e40346.
14. Bae, C., et al., *Xerocytosis is caused by mutations that alter the kinetics of the mechanosensitive channel PIEZO1*. Proc Natl Acad Sci U S A, 2013. **110**(12): p. E1162-8.
15. Bae, C., P.A. Gottlieb, and F. Sachs, *Human PIEZO1: removing inactivation*. Biophys J, 2013. **105**(4): p. 880-6.
16. Sonnhammer, E.L., G. von Heijne, and A. Krogh, *A hidden Markov model for predicting transmembrane helices in protein sequences*. Proc Int Conf Intell Syst Mol Biol, 1998. **6**: p. 175-82.
17. Kall, L., A. Krogh, and E.L. Sonnhammer, *Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W429-32.
18. Bernsel, A., et al., *TOPCONS: consensus prediction of membrane protein topology*. Nucleic Acids Res, 2009. **37**(Web Server issue): p. W465-8.

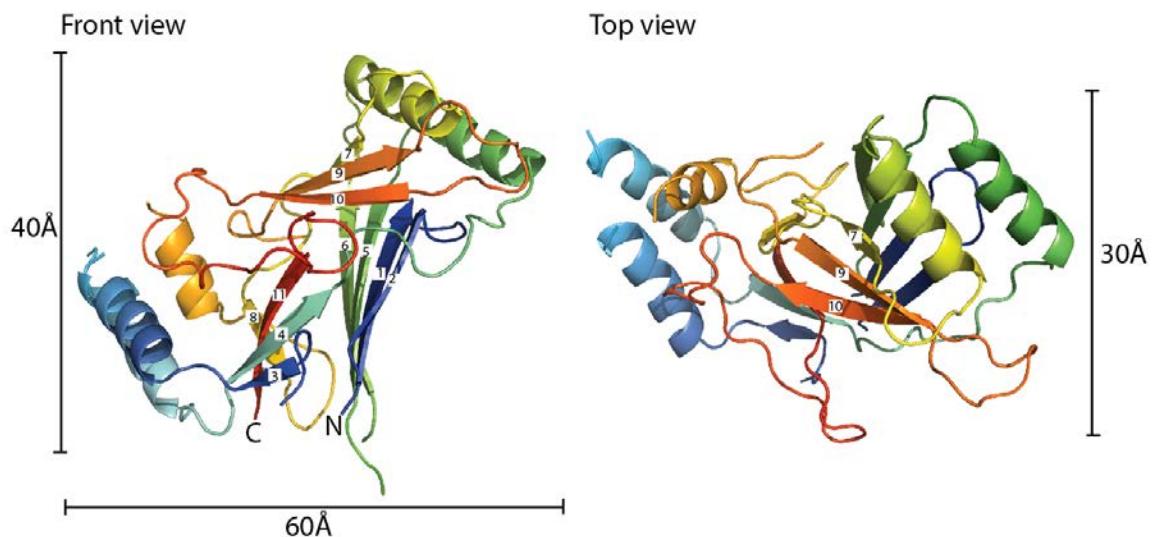
19. Hofmann, K. and W. Stoffel, *TMBASE - A database of membrane spanning protein segments*. Biol. Chem. Hoppe-Seyler, 1993. **374**: p. 166.
20. Schultz, J., et al., *SMART, a simple modular architecture research tool: identification of signaling domains*. Proc Natl Acad Sci U S A, 1998. **95**(11): p. 5857-64.
21. Katoh, K., et al., *MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform*. Nucleic Acids Res, 2002. **30**(14): p. 3059-66.
22. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server*. Nat Protoc, 2009. **4**(3): p. 363-71.
23. Laskowski, R.A., J.D. Watson, and J.M. Thornton, *Protein function prediction using local 3D templates*. J Mol Biol, 2005. **351**(3): p. 614-26.
24. Laskowski, R.A., J.D. Watson, and J.M. Thornton, *ProFunc: a server for predicting protein function from 3D structure*. Nucleic Acids Res, 2005. **33**(Web Server issue): p. W89-93.
25. Ashkenazy, H., et al., *ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W529-33.
26. Krissinel, E. and K. Henrick, *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2256-68.
27. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.

28. Sillitoe, I., et al., *New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures*. Nucleic Acids Res, 2013. **41**(Database issue): p. D490-8.
29. Jasti, J., et al., *Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH*. Nature, 2007. **449**(7160): p. 316-23.
30. Kawate, T., et al., *Crystal structure of the ATP-gated P2X(4) ion channel in the closed state*. Nature, 2009. **460**(7255): p. 592-8.
31. Battye, T.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.
32. Potterton, E., et al., *A graphical user interface to the CCP4 program suite*. Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 7): p. 1131-7.
33. Kabsch, W., *Integration, scaling, space-group assignment and post-refinement*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 133-44.
34. Evans, P., *Scaling and assessment of data quality*. Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 1): p. 72-82.
35. Terwilliger, T.C., et al., *Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard*. Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 6): p. 582-601.
36. Zwart, P.H., et al., *Automated structure solution with the PHENIX suite*. Methods Mol Biol, 2008. **426**: p. 419-35.
37. McCoy, A.J., *Solving structures of protein complexes by molecular replacement with Phaser*. Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 1): p. 32-41.

38. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
39. Murshudov, G.N., et al., *REFMAC5 for the refinement of macromolecular crystal structures.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 355-67.
40. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.3r1.* 2010.
41. Johns, S.J., *TOPO2, Transmembrane protein display software.*

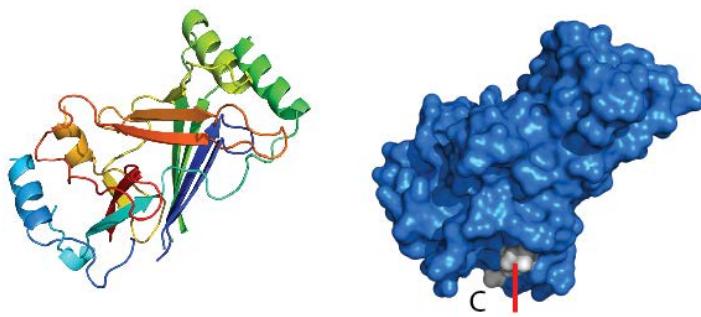


**Figure 1. Membrane topology of mouse PIEZO1 as predicted by various membrane topology prediction algorithms.** Depending on the membrane topology prediction algorithm, mouse PIEZO1 is predicted to have 30 to 40 transmembrane segments. The numbers of amino acids (aa) for the larger predicted loops are indicated. The loop of interest, CTL2 (red box), is predicted to be an extracellular loop by 3 of the 4 topology prediction algorithms. Since mouse PIEZO1 is the most extensively characterized Piezo channel, it is used to illustrate the topology predictions; Diagrams are drawn using TOPO2 [41] <http://www.sacs.ucsf.edu/TOPO2/>.

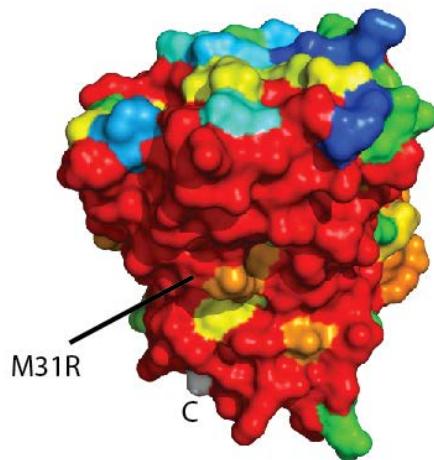


**Figure 2. Structure of the wild type *C. elegans* CTL2 domain.** The structure of the wild type Piezo CTL2 is colored in a rainbow scheme, progressing from the N-terminus (blue) to C-terminus (red). Front view (left): The loop is oriented so that the connections to the putative transmembrane helices are positioned toward the bottom. The β-strands are numbered sequentially from the N-terminus. Top view (right): the front view, rotated 90° around the horizontal axis.

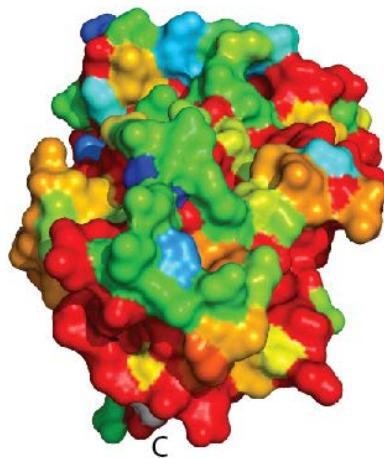
A



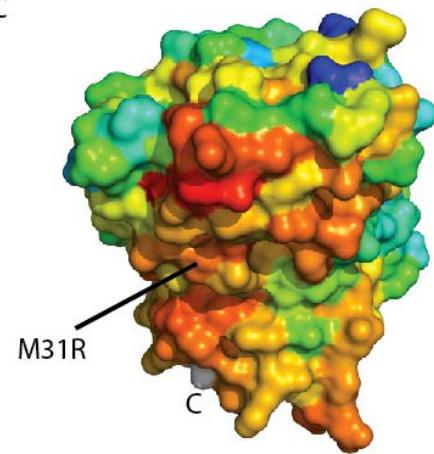
B

 $\sim +90^\circ$  $\sim -90^\circ$ 

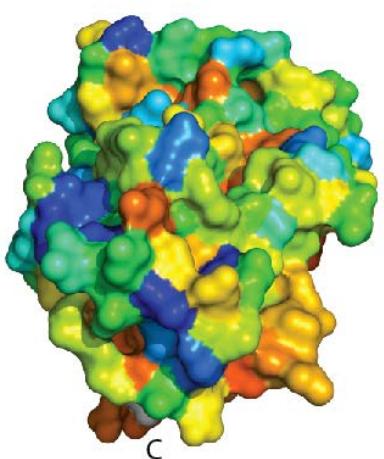
ProFunc



C



ConSurf



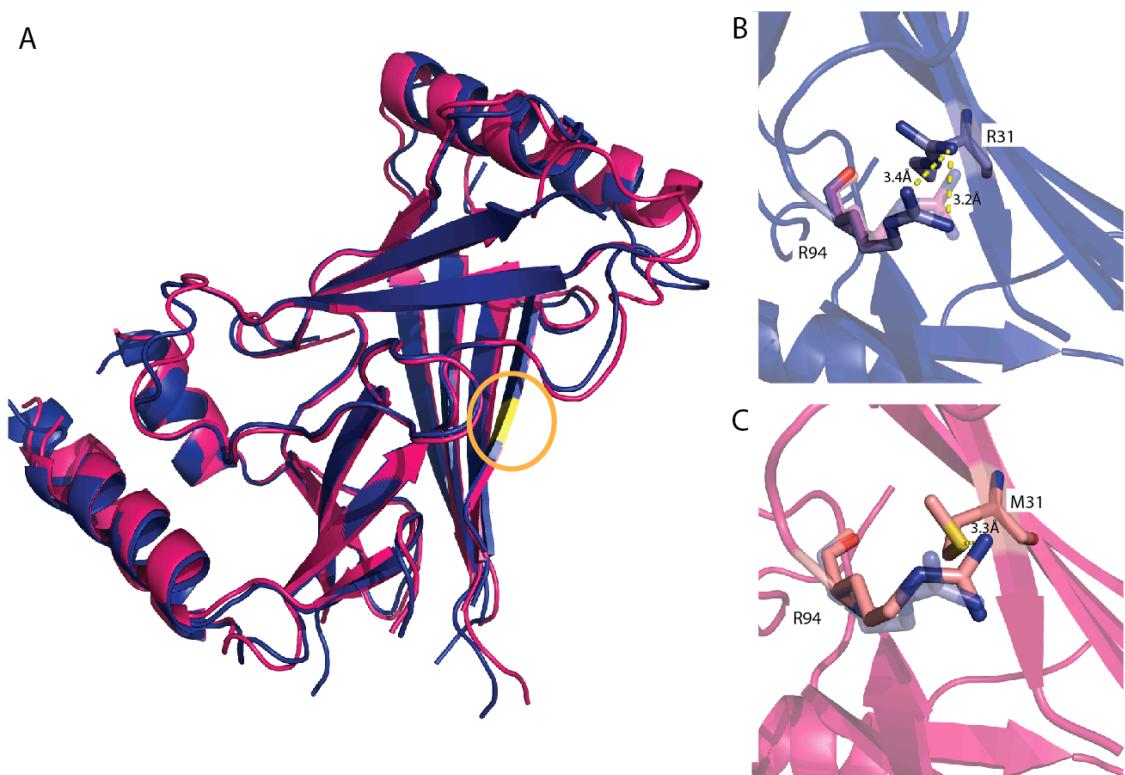
Low conservation



High conservation

**Figure 3. Residue conservation analysis of CTL2**

(A) Surface representation of the WT CTL2 domain (marine blue) is oriented in similar fashion to the one displayed in Figure 2 (inside the black square). N and C terminus are colored in light gray. (B) The top panel displays surface conservation representation from the closest Piezo homologs as calculated by ProFunc using the pdb coordinate of the WT CTL2 domain structure, whereas the bottom panel display surface conservation from representative Piezo homologs taken from broad branches of eukaryotic kingdom as calculated by ConSurf. Left panel shows the surface of CTL2 upon 90 degree rotation along the y-axis, showing the right-side view of CTL2. Right panel shows the surface of CTL2 upon -90 degree rotation, showing the left-side view of CTL2. Surface conservation is colored in a rainbow gradient scheme in such a way that the least conserved region is colored in blue and the most conserved region is colored in red. M31R is located at the conserved patch of CTL2.

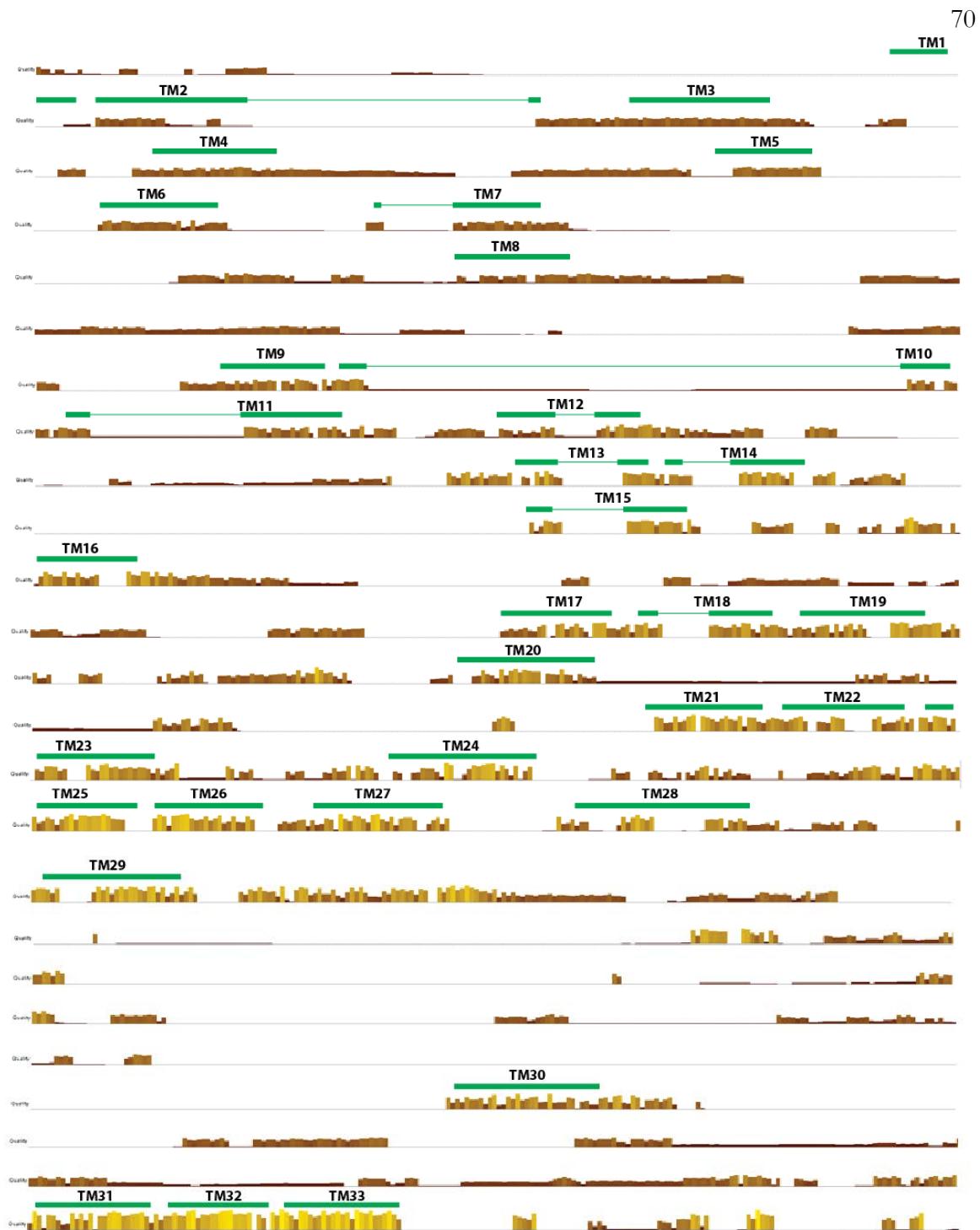


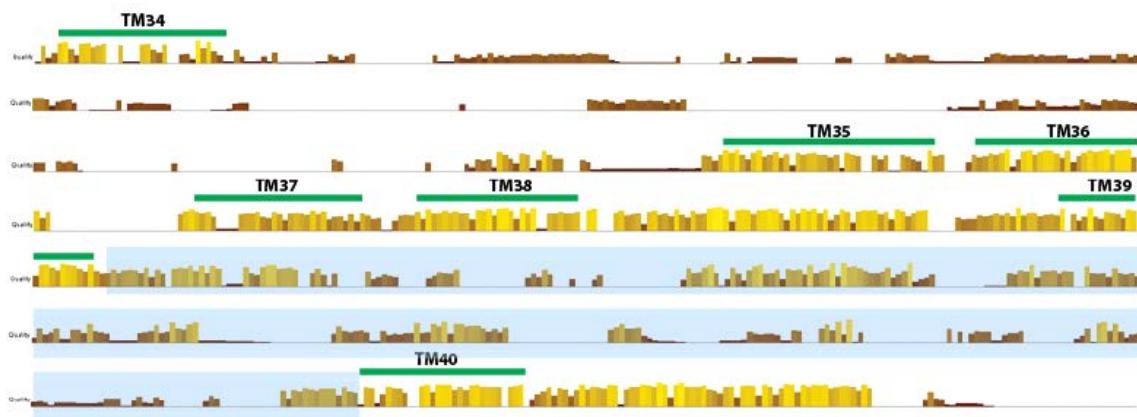
**Figure 4. Superposition of wild type and M31R substituted CTL2 domains.** (A) Superposition of the wild type (pink) and M31R substituted (blue) CTL2 structures. The position of residue 31 is highlighted in yellow (circled). (B, C). The M31R mutant structure is displayed in blue and the WT structure is displayed in pink; this substitution is accompanied by a modest re-arrangement of the R94 side chain.

**Table 1. Data collection and refinement statistics of the wild type and mutant**

Data set	Wildtype CTL2	CTL2 M31R
Protein Data Bank number	4PKE	4PKX
Data collection statistics		
Space group	H 3 2	P 1 2 <sub>1</sub> 1
a, b, c (Å)	72.59, 72.59, 241.6	45.18, 37.25, 73.56
α, β, γ (°)	90, 90, 120	90, 91.58, 90
Wavelength (Å)	1.0716	1.0332
Resolution (Å) <sup>a</sup>	33.09 – 2.45 (2.57 – 2.45)	38.97 – 2.54 (2.68 – 2.54)
No. of unique reflection <sup>a</sup>	9377 (1211)	7746 (887)
Mean (I/sigma I) <sup>a</sup>	17.7 (3.0)	28.4 (11.2)
Completeness <sup>a</sup>	99.9% (100%)	93.6% (75.4%)
Average multiplicity <sup>a</sup>	10.1 (10.6)	3.1 (2.8)
Anom. Completeness <sup>a</sup>	99.9% (99.9%)	
DelAnom. correlation between half-sets	0.915 (0.061)	
<sup>a</sup>		
R-merge <sup>a</sup>	0.093 (0.810)	0.028 (0.085)
Matthew Coefficient	1.951	1.92
Refinement statistics		
Resolution	31.17 – 2.50	33.23 – 2.54
Average B factor (Å <sup>2</sup> )	83.9	38.6
R <sub>work</sub> /R <sub>free</sub> (%)	22.81 / 23.84	21.79 / 25.91
Rmsd bonds (Å)	0.006	0.004
Rmsd angles (°)	1.23	0.97
Ramachandran plot (%)		
Favored	97.5	98.2
Outliers	0	0
Rotamer outliers (%)	5.9	2.9
Molprobity clashscore	13.8	12.8
Molprobity overall score	2.34	1.97

<sup>a</sup> Values in parentheses represents the highest resolution shell



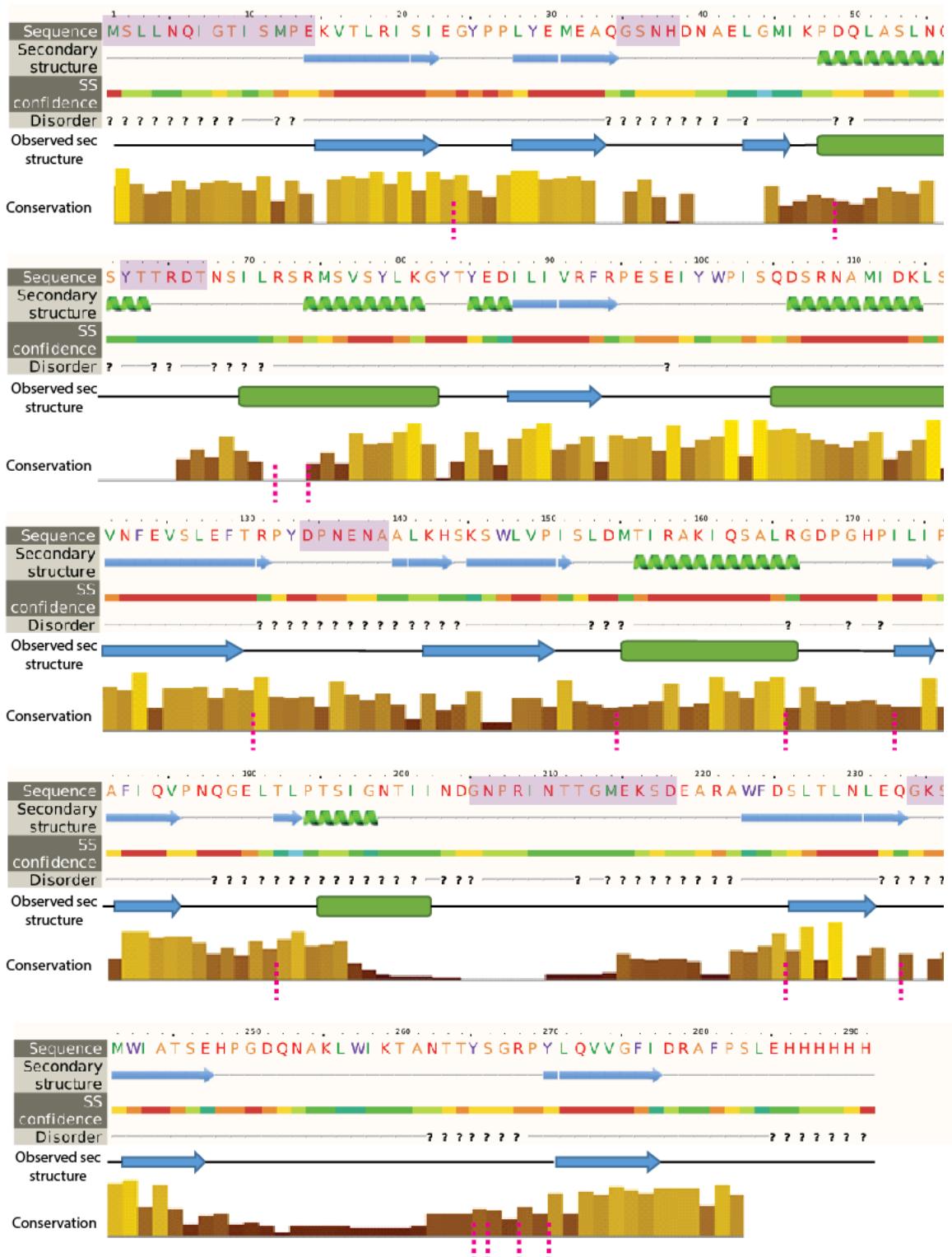


**Figure S1. Multiple sequence alignment of Piezo homologs.** Multiple sequence alignment was performed using representative Piezo homologs (listed below) from vertebrate Piezo1, vertebrate Piezo2, invertebrate Piezo, plant Piezo, and unicellular eukaryote Piezo. Protein sequence conservation is represented by the shading and height of the bar corresponding to each residue, with the tall yellow bar, short brown bar, and no bar representing high levels, low levels, and poor sequence conservation, respectively. The green bars highlight the transmembrane (TM) segment for mouse PIEZO1 as predicted by Topcons [18]. CTL2 (shaded in light blue) is the largest conserved Piezo soluble domain. The Piezo sequence conservation is relatively higher in the C-terminal region. The diagram was generated using the MAFFT multiple sequence alignment program [21], available at the EMBL-EBI web page.

List of Piezo homologs used:

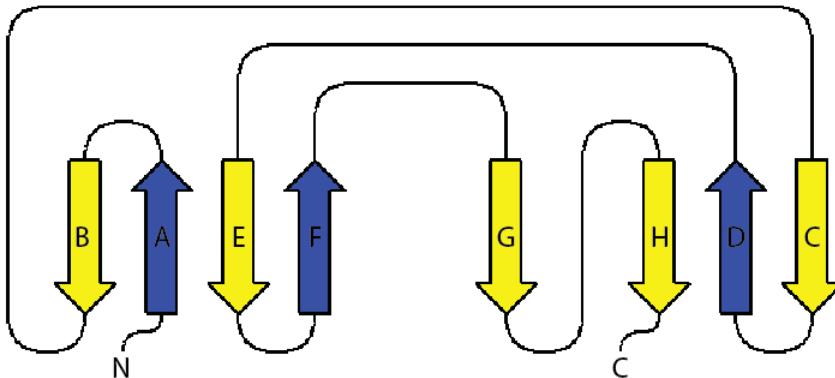
H. sapiens 1 (NP\_001136336.2), D. rerio 1 (XP\_696355.4), R. norvegicus 1 (NP\_001070668.2), M. lucifugus 1 (XP\_006097426.1), O. garnettii 1 (XP\_003800876.1), B. Taurus 1 (XP\_001256011.4), S. harrisii 1 (XP\_003758542.1), G. gorilla 1

(XP\_004058200.1), *C. porcellus* 1 (XP\_003460961.1), *E. caballus* 1 (XP\_005615040.1),  
*P. alecto* 1 (XP\_006927190.1), *M. domestica* 1 (XP\_007477361.1), *L. Chalumnae* 1  
(XP\_006002160.1), *T. rubripes* 1 (XP\_003978351.1), *Anolis carolinensis* 1  
(XP\_008120472.1), *C. lupus* 1 (XP\_005620631.1), *M. gallopavo* 1 (XP\_003209947.1), *S.*  
*partitus* 1 (XP\_008279724.1), *X. tropicalis* 1 (XP\_002933721.2), *C. simum* 1  
(XP\_004437180.1), *H. glaber* 1 (XP\_004843264.1), *O. rosmarus* 1 (XP\_004392217.1), *M.*  
*musculus* 1 (NP\_001032375.1), *O.orca* 1 (XP\_004280184.1), *M. musculus* 2  
(NP\_001034574.4), *G. Gallus* 2 (XP\_419138.4), *H. sapiens* 2 (NP\_071351.2), *B. Taurus* 2  
(XP\_003587835.2), *C. porcelius* 2 (XP\_005001285.1), *C. simum* 2 (XP\_004437180.1), *H.*  
*glaber* 2 (XP\_004843264.1), *M. lucifugus* 2 (XP\_006089356.1), *C. millii* 2  
(XP\_007887509.1), *O. orca* 2 (XP\_004276120.1), *O. aries* 2 (XP\_004020692.1), *G. gorilla*  
2 (XP\_004059237.1), *O. rosmarus* 2 (XP\_004417033.1), *M. lucifugus* 2  
(XP\_006089356.1), *X. tropicalis* 2 (XP\_002937522.2), *C. lupus* 2 (XP\_005623199.1), *M.*  
*gallopavo* 2 (XP\_003205004.1), *A. platyrhynchos* 2 (XP\_005013163.1), *T. guttata* 2  
(XP\_002192627.2), *S. harrisii* 2 (XP\_003760113.1), *A. aegyptii* (XP\_001657818.1), *B.*  
*impatiens* (XP\_003494661.1), *H. saltator* (EFN75267.1), *C. elegans* (CAA92491.3), *C.*  
*sinensis* (GAA51253.1), *D. melanogaster* (AFB77909.1), *M. occidentalis*  
(XP\_003747214.1), *P. humanus* (XP\_002428649.1), *N. vitripennis* (XP\_008202351.1), *E.*  
*histolytica* (XP\_655549.2), *P. tetraurelia* (XP\_001461126.1), *O. trifallax* (EJY84567.1), *T.*  
*cruzi* (EKG00857.1), *S. lycopersicum* (XP\_004247483.1), *G. max* (XP\_006605262.1), *T.*  
*cacao* (XP\_007030785.1), *A. thaliana* (NP\_182327.6), *C. rubella* (XP\_006293550.1), *P.*  
*persica* (XP\_007200947.1), *O. tauri* (XP\_003079754.1)



**Figure S2. Comparison of the secondary structure of the WT *C. elegans* CTL2 from the crystal structure and from the Phyre2 prediction.** The observed secondary structure in the crystal structure of *C. elegans* CTL2 (drawn below the ‘Disorder’ row) generally resembles the secondary structure assignments predicted by Phyre2 [22] (drawn below the ‘Sequence’ row). Amino acid residues that are not visible in the electron density of either the wild type or M31R mutant structures (shaded in purple) coincide with the predicted disordered region. The sequence conservation representation follows the convention used in Fig S1, with gaps in the multiple sequence alignment indicated by the magenta dotted lines.

A



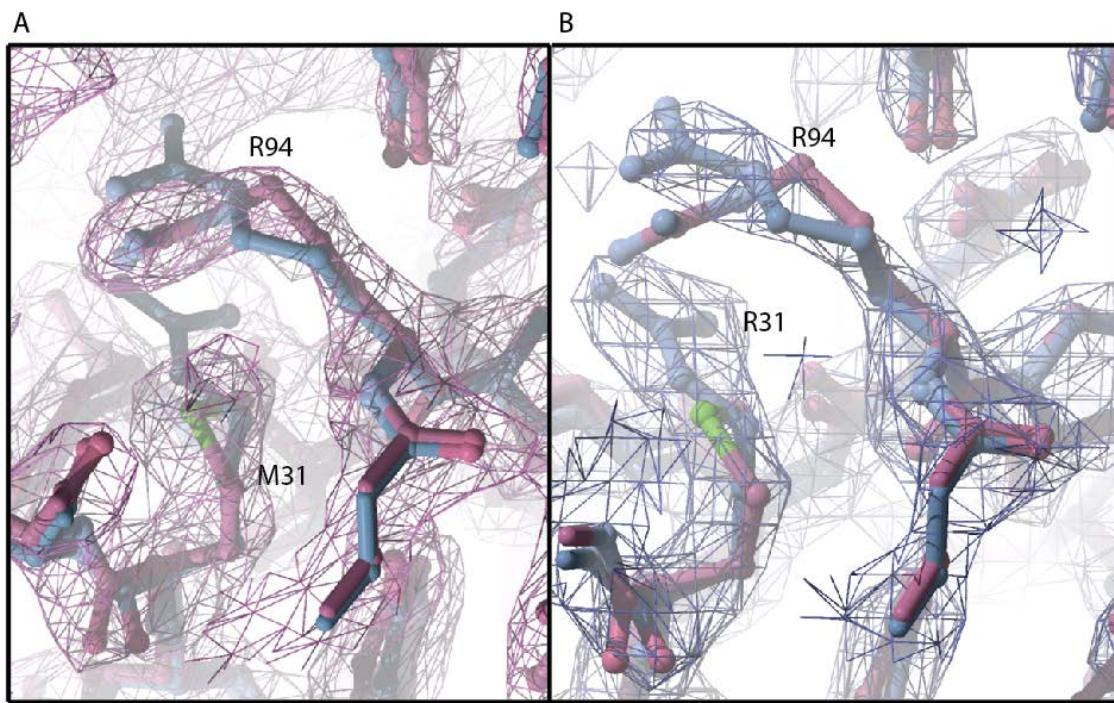
B

Hit	Z-score	rmsd	PDB	Method	Description	CATH #	Sheet 1	Sheet 2
1	3.6	3.3	3N9D	SSM	Monoclinic LigD ( <i>P. aeruginosa</i> )	N/A	ABCD	HGFE
2	2.8	4.1	3QK0	DALI	Phosphatidylinositol-4,5-biphosphate 3-kinase	N/A	GHAD	CBEF
3	2.6	4.1	3L9B	DALI - SSM	Otoferlin	N/A	GHAD	CBEF
4	2.6	3.4	3NSJ	DALI	Perforin-1	N/A	GHAD	CBEF
5	2.6	2.5	2QZ5	DALI	Axin interactor, dorsalization associated	N/A	DEAIH	GFBC
6	2.4	3.4	1TFP	DALI - SSM	Transthyretin	2.60.40.180	DAGH	FEBC
7	2.4	3.8	2NQ3	DALI	ITCH-homolog E3 ubiquitin ligase	N/A	GHAD	CBEF
8	2.3	3.7	3JZY	DALI	Intersectin 2	N/A	GHAD	CBEF
9	2.1	3.4	2CJS	DALI	UNC-13 homolog A	N/A	DAHG	FEBC
10	2.1	4.1	1CYJ	DALI	Cytosolic phospholipase A2	2.60.40.150	GABD	CEF
11	2.1	4.7	1D5R	DALI	Phosphoinositide phosphatase PTEN	2.60.40.150	GABD	CEF
12	2.0	4.2	1UGK	DALI	Synaptotagmin IV	2.60.40.150	GABD	CEF
13	2.0	3.9	1GMI	DALI	Protein kinase C, epsilon type	2.60.40.150	GABD	CEF
14	2.0	3.1	2H6U	DALI	5-Hydroxyisourate hydrolase	2.60.40.180	DAGH	FEBC

**Figure S3. Comparison of the beta strand connectivity of the Piezo CTL2 domain to potentially related domains.** Each beta strand in the connectivity diagram is drawn as an arrow. Beta strands that are parallel to the first beta strand are colored blue, while those that are antiparallel are colored yellow (A). Connectivity diagrams for each of the 14 closest matches to the Piezo CTL2 loop identified by DALI [27] and SSM [26] are summarized in table (B). No exact matches in strand orientation are observed between CTL2 and these candidate structural homologs.

C.elegans	MPEKVTLRISIEGYPPPLY <b>E</b> MEAQGSNHDNAELGMIKPDQLASLNQALTD SYTTRDTNSIL
A.aegypti	LPYDVSVTLRIGPYEPVY <b>V</b> MSAQDSN-----IHGLNDAQWEKFMAPYA-----
Bumblebee	LPYDVSMKIRIGPYEPIY <b>S</b> MSAQSS-----IEYDETFMRFNSNLYA-----
JumpAnts	LPYDVSMKIQIGPYEPIY <b>S</b> MSAQGSS-----IREYTKAEYDDLTNIYT-----
D.rerio1	HPVDVTVTKLGGYEPLFT <b>T</b> MSVQQQS-----IQPFTESRYNQLNNQFS-----
X.tropicalis1	HPIDVTVTFKLGGYEPLFT <b>T</b> MSAQQQS-----IQPFTPQQYEALTYEFE-----
M.musculus1	QPIDVTVTKLGGYEPLFT <b>T</b> MSAQQQPS-----IVPFTPQAYEELSQQFD-----
H.sapiens1	QPIDVTVTKLGGYEPLFT <b>T</b> MSAQQQPS-----IIPFTAQAYEELSQRQFD-----
Microbat1	QPIDVTVTKLGGYEPLFT <b>T</b> MSAQQQPS-----IVPFTQQAYEELSQRQFD-----
C.familiaris1	QPIDVTVTKLGGYEPLFT <b>T</b> MSAQQQPS-----IVPFTQQAYEELSQRQFD-----
Xenopus2	KPLDVSITITLGYYQPIF <b>T</b> MSAQQNQ-----LKGLNDNEFKHLHNLYK-----
M.musculus2	QPLDVSVTITLGYYQPIF <b>T</b> MSAQQSQ-----LKVMNDNSKYNEFLKSFG-----
Microbat2	QPLDVSVTITLGYYQPIF <b>T</b> MSAQQSQ-----LKVMNDQGSFNKFQAFS-----
H.sapiens2	QPLDVSVTITLGYYQPIF <b>T</b> MSAQQSQ-----LKVMDQTRFNKFQAFS-----
C.familiaris2	QPLDVSVTITLGYYQPIF <b>T</b> MSAQQSQ-----LKVMDQTRFNKFQAFS-----
	* . *:: . : * *:: *..* . : :
C.elegans	RSRMSVSYLKGYTYEDILIVEFRPESE <b>I</b> YWPISQDSRNAMIDL-SRNTSVNFVEVSLEFT
A.aegypti	KDKTALTFLSNYESVDVAAVKLGANSTSIWNISPPDKARLLNDL-NTTSTLTCRFRYTIS
Bumblebee	RDRPAVTFLENYIHSDVAAVRLSGFSRKLWSISPPDLDRLITELEDNSTTVVIHVEWTVS
JumpAnts	KEKSAVTFFENYVYSDVAAVRFSGFS <b>K</b> RFWGIPSPPDRERIKSELASNTTVIIHVEWTVS
D.rerio1	KNAVAMQFITMYSYEDIVTANIE <b>G</b> SSSGSVWRISPSSRQELIKELLSSTGDMTLRLDWNFQ
X.tropicalis1	RQPTAMQFITLYSYEDIVTARI <b>E</b> SSSGSVVGISPSSREQMRKELQNGSSDITLRFTWDFO
M.musculus1	PYPLAMQFISQYSPEDIVTAQI <b>E</b> SSSG <b>G</b> ALWRISPSSRAQMKELEYNGTADITLRFTWNFQ
H.sapiens1	PQPLAMQFISQYSPEDIVTAQI <b>E</b> SSSG <b>G</b> ALWRISPSSRAQMKELEYNGTADITLRFTWNFQ
Microbat1	PHPLAMQFISQYSPEDIVTAQI <b>E</b> SSSG <b>G</b> ALWRISPSSRAQMKELEYNGTADITLRFTWNFQ
C.familiaris1	PNPLAMQFISQYSPEDIVTAQI <b>E</b> SSSG <b>G</b> ALWRISPSSRAQMKELEYNGTADITLRFTWNFQ
Xenopus2	GNTPAMQFILESYMQEDITIAGLEGNNSN <del>SLWT</del> ISPPSRMMIERLIQ-EPDFTAVISWSIR
M.musculus2	PNSGAMQFLENYEREDVTVAELEGNSN <del>SLWT</del> ISPPSKQKMIHELLDPNSSFSVVFSWSIQ
Microbat2	SDTGAMQFLENYEKEDITVAELEGNSN <del>SLWT</del> ISPPSKQKMIHELLDPNSSFSVVFSWSIQ
H.sapiens2	RDTGAMQFLENYEKEDITVAELEGNSN <del>SLWT</del> ISPPSKQKMIHELMDPNSSFSVVFSWSIQ
C.familiaris2	RDTGAMQFLENYEKEDITVAELEGNSN <del>SLWT</del> ISPPSKQKMIHELMDPNSSFSVVFSWSIQ
	: : :: * *: . : * * * * . : * . . .

**Figure S4. Multiple sequence alignment of Piezo homologs around the M31 mutation site.** A methionine residue that is mutated into arginine in DHS patients (colored in red) is conserved across Piezo homologs. The position of this methionine (M31 in our *C. elegans* CTL2 construct) is in close proximity to an arginine residue (colored in blue) that is not conserved. There are three glutamate residues (colored in purple) around M31 that provide a net negative charge at this location.



**Figure S5. Substitution of M31 with Arg leads to repositioning of the R94 sidechain.**

The structure of the wild type *C. elegans* PIEZO CTL2 structure is displayed in magenta, while the M31R mutant structure is displayed in blue. A: the electron density of WT Piezo loop is displayed in magenta. B: the electron density of the M31R mutant CTL2 domain is displayed in blue.

# **Chapter 3:**

## **The Pursuit Toward Obtaining a Full Length Piezo Structure**

## *Chapter 3: The Pursuit Toward Obtaining a Full Length Piezo Structure*

### **Historical Context and The Chapter Outline**

The notion of “a picture is worth a thousand words” is certainly true in the field of molecular biology. In the absence of images, it is difficult to deduce a detailed mechanism on how a particular molecular machine works. Consequently, there have been major efforts dedicated to elucidating the structure of biological macromolecules. However, these molecules are much too small for direct visualization by regular light microscopy techniques. Therefore, we have to utilize specialized imaging techniques which employ electromagnetic radiation of much smaller wavelength than visible light to obtain snapshots of these molecules at atomic resolution.

In the late 1895, Wilhelm Conrad Rontgen discovered a new kind of electromagnetic radiation which differentially penetrated objects that were made from different material; he named such radiation ‘x-rays’ [1]. Almost two decades later, x-ray diffraction of copper (II) sulfate crystal was observed by Max von Laue, suggesting that x-rays are waves and crystals can diffract the incoming x-rays. Subsequently, William Lawrence Bragg proposed that two x-ray waves will constructively interfere when the path length difference between these two waves is equal to an integer multiple of their wavelength [2]; this is known as Bragg’s Law. Using this relation, one can deduce the atomic structure of simple crystals from the x-ray diffraction pattern [3]. More generally, the x-ray diffraction pattern and the molecular structure are related by Fourier transformation. Elucidation of the molecular geometries of amino acids and peptides by x-ray crystallography paved the way for the

formulation of alpha helical model of protein secondary structure [4]. Then, another structural biology milestone took place when James Watson and Francis Crick deduced the structure of the DNA double helix from x-ray diffraction images obtained from DNA fibers by Rosalind Franklin [5].

In a typical x-ray crystallography experiment, data is collected as the intensities of x-ray diffraction reflections, containing the amplitudes but missing the phase information. This is commonly referred to as the ‘phase problem’. In 1953, Max Perutz invented a way to solve this problem by utilizing the diffraction pattern difference between the native protein crystal and its heavy metal derivatized isomorphic crystal, a method that is known as isomorphous replacement [6]. Using this method, we can now obtain atomic structures of any well-diffracting protein crystals from which we can deduce detailed mechanisms on a variety of molecular processes; this gave rise to a new biological science field known as the Structural Biology.

X-ray crystallography has been the major tool for obtaining structures of biological macromolecules. As of March 2016, 89.4% of all the structures deposited in the Protein Data Bank (PDB) were obtained by x-ray crystallography. For this reason, we initially aimed to obtain the molecular structure of the full length Piezo by x-ray crystallography. One key requirement for obtaining a suitable protein crystal is to have an abundant amount of protein. Consequently, early efforts to solve protein structures were limited to those proteins which are found abundantly in nature, such as myoglobin [7], haemoglobin [6], hen egg lysozyme [8], ribonucleases [9, 10], chymotrypsin [11], papain [12], carboxypeptidase A [13], and substilisin [14].

With the advancement of recombinant DNA technology in the late 1970s, proteins that are not naturally abundant could finally be produced in cells that were engineered to overexpress the protein of interest [15]. Although progress in solving the structures of soluble proteins experienced a significant boost with the development of overexpression methods, membrane protein structure determination has still trailed behind (Figure 1); the first structure of a membrane protein did not come out until 1985 [16]. This is partly because many membrane proteins failed to be expressed in sufficient amounts using commonly used recombinant systems; this is especially true for many eukaryotic membrane proteins [17].

Furthermore, solubilization of membrane proteins by amphipathic reagents, such as detergents, can be problematic. For this reason, significant efforts have been dedicated to optimize the expression of membrane proteins using different expression systems and different purification conditions. Therefore, **in the first part of this chapter**, I will describe our efforts in exploring several different protein expression systems and detergent-buffer conditions to improve the yield of full length Piezo protein for x-ray crystallographic study.

In December 2013, the first membrane protein structure, that of Transient Receptor Potential V1 (TRPV1), was solved using single particle cryo-electron microscopy (Cryo-EM) by a collaborative effort of the David Julius and Yifan Cheng labs [18, 19]. The major advantage of single particle cryoEM over x-ray crystallography is that protein crystals are not required. Obtaining suitable protein crystals of a membrane protein can be difficult. Therefore, the technology advancements in the single particle cryo-EM field offer an attractive alternative method to obtain protein structures. **In the second part of this**

chapter, I will describe our attempt to look at the full length Piezo protein using single particle EM.

## **Part 1: Exploring Various Protein Expression System for Maximizing the Production of Mammalian Membrane Protein**

### Introduction

Limitations on the availability of naturally-abundant proteins posed significant challenges to the structural biology field. The revolution in DNA manipulation technology circumvented this problem by providing approaches, in principle, to engineer the over-expression of any protein. In the late 1960s and early 1970s, major breakthroughs came about with the isolation of enzymes that can specifically cut DNA molecules (restriction enzymes) [20, 21], and enzymes that can join two double stranded DNA molecules (DNA ligases) [22]. These discoveries enabled the foundational concept that DNA sequences from one species can be copied and inserted into a foreign host. This led the way for the construction of the first few DNA plasmid vectors with selective markers and screening tools to do molecular cloning [23, 24].

In 1977, Sanger et al and Maxam et al independently presented ways to do DNA sequence analysis by the ‘chain-termination method’ and cleavage of partially modified nucleobases, respectively [25, 26]. This allows the resulting DNA product to be verified with single base-pair resolution to detect if there is unwanted incorporation of mutations during the molecular cloning process. In the early 1980s, Polymerase Chain Reaction

(PCR) was invented [27]. In combination with the preceding discovery of reverse transcriptase [28, 29], we are now able to clone eukaryotic protein coding sequences (exonic regions) into cloning vectors for heterologous protein expression using different host cells.

### ***Bacterial Protein Expression Systems as the Major Tool to Produce Large Quantities of Protein***

Bacterial expression systems, particularly *Escherichia coli*, remain the major work force for recombinant protein expression, contributing to more than 96,000 out of 110,000 total protein structures deposited at PDB as of February 2016. Factors that make *E. coli* the initial choice for most laboratories are the low cost and rapid time scale. The time it takes from the initial cell inoculation until the time for harvest is between two to five days, depending on the expression method being used. Protein production with 60 L *E. coli* culture costs less than \$100; the biomass obtained from a 60 L culture of *E. coli* grown in a bioreactor can exceed 700 g of cells. In contrast, protein production in 1 L suspension-adapted mammalian culture costs roughly \$100, and yields only 10 g of biomass. These considerations, along with the ease-of-use, make *E. coli* the major workforce for large-scale protein synthesis for structural studies. Indeed, we used an *E. coli* expression system to express the Piezo soluble domain, CTL2 (discussed previously in Chapter 2).

There is a concern, however, that bacterial protein expression system may not be suitable for the mass production of many eukaryotic proteins for several reasons. First, bacterial systems lack the sophisticated downstream post-translational protein modification machinery that may be required for the proper folding of many eukaryotic proteins.

Furthermore, lipid composition is crucial for the proper folding as well as for the function of membrane protein [17, 30]; difficulties in obtaining sufficient amounts of functional eukaryotic membrane proteins may be attributed to differences between the lipid compositions of the bacterial membrane and the membrane of the native species. Although several advances to mitigate these issues have been made [31], many labs and pharmaceutical companies have turned to eukaryotic protein expression system for the synthesis of eukaryotic membrane proteins.

### ***Pichia – Yeast Expression System***

*Pichia pastoris* and *Saccharomyces cerevisiae* are the two most common yeast hosts used for the overexpression of recombinant proteins. Although *S. cerevisiae* has produced several successes, protein expression using the *P. pastoris* system offers two major advantages. First, *Pichia* is able to utilize simple alcohols, such as methanol, as the carbon source. This is relatively cheaper compared to sugars that are the required carbon source for *S. cerevisiae*. Further, *Pichia* can be grown to a very high density (up to 100g of cells/L) [32], which is scalable for a large-scale growth using a Bioreactor [33]. Unlike *E. coli* expression systems, with *Pichia*, the vector carrying the gene of interest is integrated into the genomic DNA through homologous recombination [34]. The genomic integration of the vector can produce either a single copy or multiple copies of the gene of interest. Since the gene of interest is placed under the alcohol oxidase (AOX1) promoter, protein synthesis can be initiated by the addition of methanol once the alternative carbon source has been depleted; initial growth was done using glycerol as a carbon source to repress the premature activation of the AOX1 promoter.

During my first year in the Rees lab, I worked with the *Pichia* system to express mammalian ion channels of the Transient Receptor Potential (TRP) channel family. One of the challenges I encountered was to administer sufficient aeration to a high-density *Pichia* culture, since a minimum level of 20% of dissolved oxygen is required for optimal protein expression; 100% dissolved oxygen level is defined as a complete oxygen saturation condition. For medium scale protein expression, I assembled a 2 – 3 L bioreactor where temperature control was done by gentle heating of a water bath using a hot plate (Figure 2b). Aeration was done by blowing sterile-filtered air into the growing culture with vigorous stirring. Typical yields were around 20 – 25 g of biomass per liter of culture, which was about twice as much as we obtained using culture flasks (albeit significantly lower than 100 g/L). Biomass yield could be improved by growing *Pichia* in a commercial bioreactor (Figure 2c). Under these conditions, *Pichia* could reach up to 80 g of cells per L of culture. Such a bioreactor provides precise monitoring of crucial parameters, such as temperature, pH, and dissolved oxygen (DO).

Despite the concern that *Pichia* only shares the core glycosylation step of (Man)<sub>8</sub>-(GlcNAc)<sub>2</sub> with mammalian cells [35], there are several examples of membrane proteins expressed in *Pichia* that were crystallized and their structures solved by X-ray crystallography [36-40]. This makes *Pichia* an attractive alternative to the bacterial system for expressing eukaryotic membrane protein.

## The Journey Begins – Exploring Various Eukaryotic Protein Expression Systems For Full Length Piezo

### ***The Decision to Switch to a Mammalian Protein Expression System***

As promising the *Pichia* expression system is, there are concerns that differences in membrane composition [41] cause the final protein product to be improperly folded (i.e. non-functional or not in a physiologically relevant state) [17]. Furthermore, Coste et al were able to express full length Piezo using a mammalian expression system which yielded functional protein [42]. These considerations, along with the lack of appropriate infrastructure for *Pichia* expression in our lab, shifted our efforts to mammalian expression systems. However, we were concerned about the protein expression level. Thus, my first objective was to optimize Piezo expression level.

We decided to use a suspension-adapted HEK293 cell line to express full length Piezo for several reasons. First of all, suspension-adapted mammalian cells can be grown to a high culture density in a shaker flask. This can provide 10 gram of biomass from one 1 L shaker flask mammalian culture. Also, maintaining a suspension adapted culture requires less work compared to the maintenance of adherent mammalian cells. This greatly reduces the amount of labor required to maintain the cells as well as to do large scale (beyond 4L) protein expression. Furthermore, suspension adapted HEK293 cells are also compatible with several transfection methods. Finally, suspension adapted cell lines are available and have been regularly used at the Protein Expression Center (PEC) facility at Caltech to

express many soluble proteins. Thus, high density culture growth, ease of use, and the well-developed infrastructure drove us to use suspension-adapted HEK293 cell culture to express the full length Piezo.

### ***Construction of Fluorescence Size Exclusion Chromatography (FSEC) Piezo Constructs***

Initially, we received six full length Piezo constructs from the Patapoutian lab: human PIEZO1, mouse PIEZO1, mouse PIEZO1-GST, mouse PIEZO2, *Drosophila* PIEZO, and *C. elegans* PIEZO. As a start, I expressed mouse PIEZO1-GST (MmP1-GST) in a 1L culture of suspension-adapted HEK293 cells in an attempt to look at the protein yield of this construct when expressed in suspension-adapted HEK cells. MmP1-GST was purified using glutathione sepharose beads in a detergent (CHAPS) – lipid (phosphatidyl choline / PC) buffer. This detergent-lipid buffer was chosen because mouse PIEZO1 had been shown to exhibit some channel activity when reconstituted into liposomes after being purified using this buffer condition [42]. From this, we were able to obtain 0.1 – 0.2 mg of pure protein from a 1L mammalian culture (Figure 3). This protein yield was still low for x-ray crystallography study.

Sample homogeneity is crucial to increase the likelihood for protein crystallization. For this reason, protein constructs are typically subjected to pre-crystallization screening to look at the monodispersity of the purified protein under several detergent buffer conditions. Two common methods to accomplish this task are Size Exclusion Chromatography (SEC) and Blue Native Poly Acrylamide Gel Electrophoresis (BN-PAGE). Traditionally, these methods require microgram to milligram amounts of pure protein. Since we could only obtain 100 – 200 microgram of mouse PIEZO1-GST protein from 1L suspension-adapted

mammalian culture, it would be difficult to obtain enough protein to test out different detergent-buffer conditions. Therefore, we were prompted to utilize Fluorescent Size Exclusion Chromatography (FSEC) [43], which is a method where we could screen the buffer condition using a low amount of protein. In this approach, a target protein is fused to Green Fluorescence Protein (GFP), and the behavior of the protein in different detergent-buffer condition is investigated using SEC by looking at the GFP fluorescence signal instead of the protein UV absorption at 280nm. Using this approach, protein can be analyzed directly from the crude extract after detergent solubilization. As a proof of principle, we characterized crude extracts from cells expressing a GFP-fused membrane protein construct, MscL-GFP, on the SEC and observed that the GFP fluorescence peak exhibited similar characteristics as the absorbance 280nm for both peak shape and retention volume (data not shown).

For my FSEC studies, I made several Piezo GFP-fusion (FSEC-PIEZO) constructs where GFP is fused at either the N-terminal or the C-terminal for each Piezo homolog (Figure 4). Once the constructs had been successfully made and verified by DNA sequencing, we tested the level of protein expression and monitored the GFP fluorescence signal to determine the optimal protein expression time. HEK cells transfected with Piezo FSEC constructs using Lipofectamine 2000 (Life Technology Inc.) started to show GFP fluorescence by 12 hours. Expression of Piezo FSEC constructs beyond 30 hours, however, resulted in increased green background fluorescence (Figure 5). For this reason, we chose 24 – 30 hours to express Piezo FSEC constructs when Lipofectamine 2000 was used. Then, we did small scale protein expression tests for these FSEC constructs to see their behavior in several mild detergents. However, most of these mild detergents showed no fluorescence

signal with the exception of C<sub>12</sub>E<sub>9</sub>-containing buffer (Figure 6). This indicates that the amount of Piezo FSEC protein that can be successfully solubilized by the set of detergents we used was very little.

To search for detergent(s) that can effectively solubilize Piezo and keep the protein from aggregating, we screened over 70 detergents. Briefly, a Piezo-containing membrane suspension was incubated at 4°C with constant shaking in the presence of 1% (w/v) of the detergent of interest. The mixture was then spun down to remove insoluble material before the supernatant was collected for analysis using SDS PAGE and BN PAGE. SDS PAGE results provided information on the extent of solubilization of Piezo by that detergent, whereas BN PAGE results showed whether the protein can migrate uniformly as a single band (monodisperse) in that detergent-buffer condition; although not all detergents are compatible with BN PAGE. Detergent screen results showed that most detergents are unable to efficiently extract full length Piezo from the membrane with the exception of sodium deoxycholate, sodium cholate, C<sub>12</sub>E<sub>8</sub>, C<sub>12</sub>E<sub>9</sub>, C<sub>12</sub>E<sub>10</sub>, and sodium dodecanoyl sarcosine (Figure 7). C<sub>12</sub>E<sub>9</sub> and C<sub>12</sub>E<sub>8</sub> were more effective than CHAPS used previously for membrane extraction. Between these six detergents, however, only in the polyoxyethylene detergent C<sub>12</sub>E<sub>8</sub>, and C<sub>12</sub>E<sub>9</sub> did the full length Piezo migrate as a single band on BN PAGE (Figure 8). For this reason, we used C<sub>12</sub>E<sub>9</sub> as detergent of choice in addition to the previously used detergent, CHAPS.

### ***Exploring Other Ways of Expressing Full Length Piezo***

Although transient transfection by chemical reagents (Lipofectamine 2000) is relatively

easy to do, there is batch-to-batch variability on the amount of protein expression. Furthermore, the transfection efficiency for large plasmids, such as Piezo-containing plasmid (15 – 20 kbp), can be quite low [41]. Thus, we attempted to explore several other ways to express full length Piezo: construction of a stably-expressing Piezo cell line, as well as Baculoviral transduction of both mammalian and insect cells.

In collaboration with the Protein Expression Center (PEC), we established a stable mammalian cell line for expressing the mouse PIEZO1-GST construct. One major advantage of using stable cell lines is elimination of batch-to-batch variation in protein expression because all cells in the culture are genetically identical (i.e. the plasmid carrying PIEZO1-GST gene is integrated into the cellular genome). To do this, HEK293 cells were transfected with mouse PIEZO1-GST IRES EGFP plasmid using Lipofectamine 2000. Several Individual cells with GFP fluorescence were isolated into 384-well plates and supplemented with Geneticin antibiotics to select for cells that had integrated the plasmid into their genome; cells that failed to integrate the plasmid into their genome would not be able to propagate efficiently since any without the plasmid would be eliminated by Geneticin. After several months of selection, two cell lines, line 4-1 and line 4-7, were obtained that exhibited stable expression of GFP fluorescence; line 4-7 showed a stronger GFP fluorescence. We then ran the cell lysate for both cell lines on SDS PAGE and analyzed the protein expression level by Western blots using anti-GST antibody. Overall, line 4-1 showed a stronger signal when blotted with anti-GST antibody, and the band was observed at the approximate molecular weight (Figure 9). Despite having a stronger GFP fluorescence signal, line 4-7 showed a much weaker signal and the band ran at a molecular weight smaller than expected for full length Piezo. For this reason, we attempted to do a

medium-scale protein expression and purification of Piezo protein from the line 4-1.

However, it did not yield any usable amount of full length Piezo protein (data not shown).

We also tried a viral-based method of gene delivery to mammalian cells, the BacMam system. Here, the gene of interest is introduced into mammalian cells using Baculovirus although the precise viral entry mechanism is unclear [44]. This method has been used to produce sufficient recombinant soluble protein for crystallization trials [45]. We cloned full length mouse Piezo1-GST into the BacMam vector (pVLAD6) that was constructed and provided by the Garcia lab [46]. 250mL of viral titer was then produced in Sf9 insect cells, and the Ratio Of Infection (ROI), defined as the volumetric ratio between the viral particle suspension to the mammalian cells, for that viral titer was determined in a small scale test expression. A non-specific histone deacetylase, sodium butyrate, was added to enhance the expression level [47, 48]. We tested the expression levels of two different cell lines: HEK293F cells and the restricted N-glycosylation HEK293 GnTI<sup>-</sup> cells [49]. Overall, the two cell lines exhibited similar expression patterns for all tested ROI (Figure 10). The protein expression was comparable to 5ng of GST loading control at the highest viral titer (almost 50% of total volume of the culture); this translates to about 50 – 100 µg of unpurified total full length mouse Piezo1-GST protein in 1L of mammalian cell culture. The combination of high amount of viral titer usage and modest amount of final protein yield prompted us to put this pursuit on hold.

Although we focused primarily on mammalian protein expression systems, we could not ignore the fact that many labs have had great successes expressing eukaryotic proteins using the Baculoviral-insect cells system; proteins expressed using this system contribute the largest number of molecular structures deposited in the Protein Data Bank (PDB)

compared to all other eukaryotic protein expression systems (Figure 11). For this reason, we cloned the full length *Drosophila* PIEZO FSEC construct, which has been modified from the construct provided by the Patapoutian lab [50], into the pVL1393 Baculoviral vector. We then tested the expression of this construct in two insect cell lines: Sf9 and High Five (Hi-5) cells. Full length *Drosophila* PIEZO protein seems to be degraded quite readily in both Hi5 and Sf9 cell lines (Figure 12). Expression of *Drosophila* PIEZO was observed at 48 hours, but not 24 hours, in both Sf9 and Hi-5 insect cell lines. Interestingly, expression of *Drosophila* PIEZO in Sf9 cells was better at lower ROI, and it gradually decreased as we increased the viral titer; there was an increase in the intensity of the two lower bands (green arrows) starting at an ROI of 2. The expression level of *Drosophila* PIEZO in Hi-5 cells, on the other hand, exhibited positive correlation with the viral titer.

Next, we performed a detergent screen to identify the detergents that efficiently extract the full length protein from insect cell membranes. Detergent screen results showed that most detergents we tested were not able to extract full length *Drosophila* PIEZO, with the exception of Fos Choline 16 and sodium dodecanoyl sarcosine (Figure 13). Therefore, we chose these detergents to perform an initial purification test with the full length *Drosophila* PIEZO FSEC protein (Figure 14a). Small scale purification of this protein was done in 100 mL insect cell culture in a batch method using 0.5 mL NiNTA column (GE healthcare). The sample was washed with fifteen column volumes of 40 mM imidazole wash buffer prior to elution using 250 mM imidazole.

It appears that the full length Piezo was absent from the elution fraction (Figure 14a); instead, we obtained more of the 75 kDa peptide fragment in both detergents. Full length *Drosophila* PIEZO protein binds poorly to the NiNTA beads (Figure 14a).

Curiously, the 75 kDa fragment was reactive to the anti-His antibody (Figure 14b). At the time, we did not attempt to investigate the identity of this fragment. However, if the 75 kDa band was indeed a peptide fragment of the full length *Drosophila* PIEZO, it must be the C-terminal Piezo fragment since the His-tag is located at the C-terminal end of the protein. It is reasonable to suspect that overexpression of functional full length Piezo is toxic for cells. As a result, the survivor cells are the ones which are able to minimize the existence of full length Piezo protein in the plasma membrane. One strategy to achieve this is by proteolysis of full length Piezo into a form which is no longer functional. Further characterization of Piezo C-terminal fragment (mouse PIEZO1) will be discussed in the last chapter (Chapter 4) of this thesis.

In addition to these efforts, we also set up a collaboration with the Stowell lab at the University of Colorado, at Boulder, to look at expression of mouse PIEZO1-GST using the Viral-Like Particles (VLP) expression system. In this expression system, sub-viral particles assemble to form vesicles from the lipid raft region of the host plasma membrane. Any membrane proteins present in that lipid raft will be packaged into vesicles which are subsequently released into the media. One major advantage of this system is that the population of protein that is packaged into the vesicles consists only of protein molecules that are folded properly and are able to localize to the plasma membrane. However, we found that VLP expression system also failed to produce full length Piezo protein in sufficient amounts for x-ray crystallography.

## **Part 2: Visualizing Full Length Mouse PIEZO1 Using Single Particle Electron**

### **Microscope**

#### ***Transition from X-Ray Crystallography to Single Particle EM***

Obtaining well-diffracting crystals for the structure determination of biological macromolecules is not trivial. Success requires a sufficient amount of good quality starting material. Proteins are typically concentrated to about 10 mg/mL for crystallization screens. For each crystallization condition, 0.5 µL of 10 mg/mL protein is mixed with an equal volume of precipitant in a hanging drop crystallization screen set-up. The typical crystallization screen kit consists of 96 different crystallization conditions. Thus, about 50 µL of 10 mg/mL (0.5 mg) of protein is needed to set up one crystallization kit at one crystallization temperature. Typically, we set up screens using several crystallization kits at different temperatures (4°C, 10°C, or 20°C). Once we obtained an initial crystal hit, a grid screen was typically performed to optimize the crystallization conditions around the initial hit. Thus, it is common for crystallographers to go through 10 – 100 milligrams of well-behaving protein from the inception to the acquisition of well-diffracting crystals. Automated liquid handlers, such as the Mosquito (TTP Labtech Inc.) and the Gryphon (Art Robbins Instruments Inc) robotic systems, enable us to set up smaller protein drops (0.2 µL) for each crystallization condition in a 96-well plate format (Figure 15). Nevertheless, this still requires a significant amount of protein. Although there is always the possibility of success by simply scaling up protein expression to obtain the required amount of protein, we decided to put these efforts on hold considering that our typical protein yield was 0.2 -

0.3 mg per 1 L of mammalian culture; at this amount, I would need about 50 L of mammalian culture to obtain the necessary amount of protein for crystallization screen.

In late 2013, the first structure of a membrane protein, the TRPV1 ion channel, was solved at high resolution using single particle cryo-electron microscopy (cryo-EM) by the collaboration of David Julius' and Yifan Cheng's labs at UCSF [18, 19]. This milestone was made possible by several recent technological breakthroughs in the field, including the direct electron detector camera [51], motion-induced image blurring correction technology [51-53], and improved maximum likelihood-based particle classification [54, 55]. Cryo-EM quickly became an attractive alternative to x-ray crystallography for obtaining the molecular structure of large proteins. One primary reason for this is that single particle cryo-EM eliminates a major bottleneck for solving protein structures by x-ray crystallography, namely the requirement of well-diffracting crystals. Thus, in principle, one needs a much smaller amount of purified protein for single particle EM imaging than for x-ray crystallography. For this reason, the structure determinations of proteins that are poorly expressed, such as Piezo, can now be attempted. It is important to note, however, that a high quality of purified protein is still required to obtain a high-resolution structure using this method.

### ***Initial Observation of Full Length Piezo by Negative Stain EM***

For electron microscopy imaging, we decided to use the mouse PIEZO1-GST construct since we had the most success with this construct in getting pure protein. Mouse PIEZO1-GST was expressed in suspension-adapted HEK293T cells. These cells were lysed and protein was purified according to the protocol that was previously described [42]. Protein

was brought to a final concentration of 3  $\mu\text{g/mL}$  and was stained using 1% uranyl acetate on a Formvar-Carbon supported copper EM grid (TED PELLA, Inc.). At this concentration, we were able to see distinct and uniform particles that adopt the shape of butterflies (Figure 16). We further optimized the staining procedure to improve the image quality by altering several parameters, such as the type of negative stain, concentration of the stain, and the type of EM grids. The significant improvement in the image quality was obtained by using a holey carbon EM grid. From here, we saw that there was a mix of butterfly and half-butterfly particles (Figure 17). However, we were not sure whether these particles were really half-sized particles or if they were different orientations of the same type of particle. Therefore, we used 2994 of half-butterfly and butterfly particles to create an initial 2D class averages (Figure 18). From the 2D class averages, we noticed limited density between the two butterfly wings. Moreover, the two butterfly wings look identical, which provides a hint that the butterfly particles may be a dimer of an intact Piezo (Figure 18). A diameter of  $\sim 15$  nm may be estimated for a spherical particle with the molecular weight of full length Piezo (1.2 MDa). Since the butterfly particle has rough dimensions of 40 nm across the two butterfly wings and 20 nm spanning each wing, it is reasonable to hypothesize that our mouse PIEZO1-GST sample contained a mix between the intact mouse PIEZO1-GST molecules (the half-butterfly particles), and dimer of the intact mouse PIEZO1-GST (the butterfly particles).

### ***Full Length Mouse PIEZO1 is a Trimer***

Dimensional analysis of purified mouse PIEZO1-GST in solution using Dynamic Light Scattering/ DLS (DynaPro NanoStar, Wyatt technology Inc.) showed that the purified protein solution contains a distribution of particles with the average radius of 7 nm. We then ran the purified protein sample on blue native PAGE (BN-PAGE) and observed that purified mouse PIEZO1-GST sample yielded 2 bands: one ran at the same position as the 1.2 MDa molecular weight marker, the other ran between the 1 MDa and the 720 kDa molecular weight marker (Figure 19a). This suggested that we had a mixed population of Piezo species, or that Piezo might partially dissociate under these conditions. These bands were also observed by Coste et al [42], although not discussed by them. We ran mouse PIEZO1-GST protein sample purified in CHAPS-PC (phosphatidyl choline) buffer on a Superose6 gel filtration column (GE Healthcare, Inc.). Here, we saw two peaks: one peak came out at the void volume (8 mL) while the other peak had retention volume of 11.3 mL (Figure 19d). Peak 1 and peak 2 were analyzed by BN-PAGE (Figure 19b). We noticed that the void peak (peak 1) contained the high molecular-weight protein smear, which is a common indicator for protein aggregates. Peak 2, on the other hand, was enriched in the band that ran between 720 kDa and 1 MDa molecular weight protein marker. Both peak 1 and peak 2 consisted of mostly one protein band on the SDS PAGE, which eliminated a possibility that the lower molecular weight peak was contaminant or degradation product of full length Piezo protein (Figure 19c).

Although the two protein bands were not completely separated, we were curious to see how the sample would look like under negative stain EM. Peak 1 looked identical to the total mouse Piezo1-GST sample directly after affinity purification procedure (Figure 20a);

it mostly contained butterfly particles. Non-specific protein aggregation could also be observed in this sample, which was consistent with the BN-PAGE result. Peak 2, on the other hand, consisted of mostly half-butterfly particle (Figure 20bc) with some butterflies. Our result suggested that the lower protein band on the BN-PAGE may correspond to the half-butterfly particles and the upper protein band is the dimer of these particles.

After this analysis was conducted, a medium resolution structure of mouse PIEZO1, solved using cryo electron microscopy (cryo-EM) [56], revealed that mouse PIEZO1 forms a trimer. The authors demonstrated that the top protein band on the BN-PAGE corresponds to the dimer of intact trimeric Piezo channel, while the lower band is a trimer. In retrospect, we should have considered a trimeric organization more carefully since the wild-type *C. elegans* CTL2 protein forms a three-fold symmetry in our crystals [57] (Figure 21). The three-fold arrangement observed in our wild-type CTL2 structure resembles the structure of the corresponding domain, C-terminal External Domain (CED), determined for the trimeric full length mouse PIEZO1 structure. CTL2 variant which contains a clinically-relevant mutation (M31R), corresponding to the M2225R mutation in human PIEZO1, did not adopt a three-fold arrangement in the crystal. This conserved methionine residue is located at the interface between the monomers and substitution of this residue into arginine may disrupt the crystal contact (Figure 21). We did not more seriously consider the possibility that CTL2 was a trimer since both the wild-type and mutant protein behaved like monomer on a Superdex S200 gel filtration column under our purification conditions (data not shown).

We then attempted to generate a low resolution 3D model for the full length mouse PIEZO-GST using the negative stain EM images that we had collected from peak 2

fraction. From these images, we manually selected 2488 half-butterfly particles to construct an initial model using C3 symmetry (Figure 22a). Our initial model is at the same approximate dimension compared to the trimeric mouse PIEZO1 structure (Figure 23).

### **Chapter Summary and Future Direction**

In this chapter, I have described our quest to obtain the structure of full length Piezo channel. We first explored several eukaryotic protein expression methods, such as transient chemical transfection of suspension-adapted mammalian culture, Baculoviral transduction of mammalian culture, construction of mammalian stable cell line, Baculoviral transduction of insect cells, and the Viral-Like Particle (VLP) method. We then searched for the optimal detergent-buffer condition for the purification of these constructs. However, the maximum protein yield we obtained (0.2 mg of protein / 1 L of mammalian culture) was less than ideal for producing the amount of protein typically used for an x-ray crystallography study.

I also described our short attempt to obtain the molecular structure of the full length mouse PIEZO1 using single particle electron microscopy. We were at the stage of analyzing our negative stain images when the structure of mouse PIEZO1 came out. The structure of mouse PIEZO1 is currently the only available model for the Piezo family. However, many of the molecular features are still missing from this model. Indeed, CED/CTL2 is the only region resolvable in high resolution. Furthermore, among 1000 amino acids residues which were modelled into the medium-resolution Piezo structure, less than 600 of them have reasonable connectivity. Thus, there is still need to obtain a higher resolution and a more complete molecular model of intact Piezo channel. Moreover, there are also other Piezo homologs that have different electrophysiological characteristics than

mouse PIEZO1, such as Piezo2 and invertebrate Piezo proteins. It would be interesting to see how the structures of these Piezo channels differ from mouse PIEZO1, and how these differences are manifested in electrophysiological differences among Piezo homologs. Last but not least, we have yet to see a compilation of Piezo structures illustrating how mechanical signals can activate these channels. Therefore, although we started our quest and explored a variety of promising routes to understand the gating mechanism of this enigmatic channel, there is still much to do to answer the fundamental question of how mechanosensing molecules, such as Piezo channel, work.

## **Methods**

### ***Cloning***

In general, the cloning was done using the In-Fusion cloning kit (Clontech Inc.) [58]. Briefly, vector DNA was linearized either by restriction digest or polymerase chain reaction (PCR). Next, Insert DNA was amplified by PCR using forward and reverse primers which contain an overhang sequence that is complementary to the DNA sequence of the linearized vector. Both insert and vector DNA were incubated together in the proprietary In-Fusion mix for 15 minutes at 55°C. The plasmid DNA mixture was transformed into Stellar *E.coli* (Clontech Inc.), and the sequence of the resulting plasmid DNA was verified by DNA sequencing.

Fluorescence Size Exclusion Chromatography (FSEC) constructs were assembled in two steps. First, His-GFP- and -GFP-His were cloned into pcDNA3.1 mammalian vector to create a backbone vector for N-terminal and C-terminal FSEC constructs, respectively.

Each Piezo gene was then inserted behind the His-GFP- for the N-terminal FSEC constructs, and in front of the -GFP-His for the C-terminal FSEC constructs. BacMam constructs were created by inserting several full length Piezo genes into PVLAD6 vector provided by Christopher Garcia's lab at Stanford University. Constructs for Piezo expression in insect cells were made by inserting a C-terminally fused Drosophila PIEZO FSEC construct into pVL1393 vector provided by the Protein Expression Center (PEC) at Caltech.

### ***Suspension Adapted Cell Culture Maintenance***

Frozen stocks of Human Embryonic Kidney (HEK) 293 were provided by the Protein Expression Center (PEC) at Caltech: HEK293F, HEK293T, and HEK293 GnTI. Cells were grown in Freestyle<sup>TM</sup> 293 expression medium (Gibco®) and maintained at roughly  $1 \times 10^6$  cells/ mL with at least 95% cell viability. Cells were diluted into  $0.5 \times 10^6$  cells/ mL when it reached  $2 - 2.5 \times 10^6$  cells/ mL. Media change / culture dilution was done every two days considering the typical doubling time for these cell lines is 24 hours.

### ***Small-Scale Test Expression***

Cells were seeded at 4 mL of  $5 \times 10^5$  cells/mL in 24-wells plates. Transfection was performed approximately 24 hours after cell seeding. Typical cell density was around  $1 \times 10^6$  cells/ mL. 4 µg of plasmid DNA was transfected into 4 mL of suspension-adapted HEK293 cells using Lipofectamine 2000 as transfection reagent (Invitrogen<sup>TM</sup>).

Immediately, 0.5 mL of cells were collected as time 0 hr. Subsequently, 0.5 mL samples were collected every 6 or 12 hours.

Collected samples were spun down, and resuspended in 100  $\mu$ L PIPES buffer. 20  $\mu$ L of 6x SDS loading buffer was added to the cell suspension, and the sample was loaded onto SDS PAGE (BioRad Inc.). Protein expression was analyzed by western blot using either anti-his or anti-GST antibodies primary antibody at 1:10,000 dilution (Rockland Immunochemicals Inc.) and secondary antibody at 1:5,000 dilution.

### ***Stable Cell Line Construction***

100 ml of suspension adapted HEK293-GnTI<sup>-</sup> Cells growing in FreeStyle<sup>TM</sup> 293 Expression Medium were transfected with the mouse PIEZO1-GST IRES EGFP construct using Lipofectamine<sup>®</sup> 2000 Transfection Reagent (L2K) in a reduced serum medium. Cells were incubated at 37°C and 8% CO<sub>2</sub> overnight. The next day, cells were supplemented with Geneticin to a final concentration of 100  $\mu$ g/mL and 1% fetal bovine serum. Cells were plated on 15 cm tissue culture dishes (20 ml per plate) and were grown for 14 days. Dead cells were washed off using fresh media and cells were incubated for another week. GFP-positive cells were selected using a cloning cylinder (plastic ring cut from 1 ml pipette). The cloning cylinder was sealed against the plate bottom using silicone grease. Cells were collected in a well of a 96-well plate and were propagated by scaling up from 96-well to 24-well plates. Cells were sequentially transferred to T25 and to T75 tissue culture flask. Serum was gradually removed from the medium. Cells were scraped off the

bottom of the flask, transferred into a 125 mL shaker flask, and grown in FreeStyle<sup>TM</sup> 293 Expression Medium.

### ***The BacMam System***

The BacMam protein expression method was performed according to the previously-described protocol [46]. Briefly, the mouse Piezo1-GST construct was cloned into the BacMam vector, pVLAD6. Expression constructs were co-transfected with linearized Baculovirus backbone into sf9 insect cells. Virus particles were harvested around 4 – 5 days, after sufficient GFP fluorescence signal was detected. This virus suspension would be used to transduce suspension-adapted mammalian culture.

HEK293 GnTI<sup>-</sup> Cells were grown in Freestyle<sup>TM</sup> 293 expression medium (Gibco®) at  $2 \times 10^6$  cells/ mL with at least 95% cell viability on the day of viral transduction. Appropriate amount of Baculovirus solution was added to 1 L mammalian culture, and a final concentration of 10 mM sodium butyrate was added to the culture. The culture was left shaking at mammalian cell culture incubator (37°C, humidified, and 5% CO<sub>2</sub>) for 42 hours before harvest.

### ***Insect Cell Protein Expression***

The full length *Drosophila* PIEZO FSEC construct was cloned into pVL1393 vector. Monolayer Sf9 insect cells were prepared at  $10^6$  cells in T25 tissue culture flasks on the day of transfection. 4 µg of Piezo in pVL1393 was mixed with 5 µg of ProGreen Baculovirus linearized DNA backbone in 250 µL Grace's insect media for 15 minutes at room

temperature. This mix was combined and incubated for 15 minutes at room temperature with another mix containing 30  $\mu$ L Lipofectin in 250  $\mu$ L Grace's insect media. DNA-lipid mixture was then added to the Sf9 cells in T25 flasks and incubated at 27°C for 18 hours. The following day, 2 – 3 mL of media was replaced with TNM-FH media, and the culture was further incubated for 5 days or until sufficient GFP fluorescence was observed. This is the P0 Baculovirus stage. One to two mL of P0 Baculovirus was used to transduce Sf9 cells in T75 tissue culture flasks at  $10^6$  cells. P1 Baculoviral stock was harvested after 5 days. 5 mL of P1 Baculovirus stock was used to transduce 200 mL of Sf9 cells at  $2 \times 10^6$  cells/ mL density. After day 5, the Baculoviral suspension was harvested as the P2 stock. Ratio of infection (ROI) for each P2 Baculoviral stock was experimentally determined on a small-scale protein test expression in 24-well deep well blocks. Finally, the appropriate amount of Baculovirus was added into 1 L of insect cells grown in shaker flasks at a density of 2 X  $10^6$  cells/ mL for large-scale protein expression.

#### ***Large Scale Protein Expression in Mammalian (HEK 293) Cells***

Suspension-adapted HEK293 cells were maintained at  $1 \times 10^6$  cells/ mL with at least 95% cell viability on the day of transfection. Polyethyleneimine (PEI) solution was prepared by dissolving 10 mg of branched PEI (average  $M_w \sim 25,000$ ) in ~ 8mL of water with 1 – 2 drops of concentrated hydrochloric acid. The solution was vortexed until all PEI powder was dissolved. Sodium hydroxide was added slowly to bring the pH back to pH 7.4 before additional water was added to bring the volume up to 10 mL.

2 mL of freshly-made PEI solution and one mg of DNA were mixed in sterile-filtered

20 mM HEPES buffer, pH 7.5, to make a complete transfection solution. This solution was gently mixed and incubated at room temperature for 5 – 10 minutes before it was slowly added to a 1 L mammalian culture. Based on the test protein expression, the optimum expression time for most Piezo constructs was determined to be between 48 to 60 hours when transfected using PEI.

### ***Membrane Fraction Isolation***

Cells were lysed in lysis buffer at a ratio of 10 mL buffer/ 1 g of cells for *E. coli* cells and 5 mL buffer/ 1 g of cells for mammalian cells. Cells were sonicated twice at 50% amplitude for 1 minute using the ‘1 sec pulse – 1 sec rest’ setting (Misonix, S 4000). Then, the sample was spun down using a tabletop centrifuge at 14,000 rpm for 15 minutes in the cold room. The supernatant was collected and spun down using an ultracentrifuge at 150,000x g for one hour to collect membrane fraction.

### ***Fluorescence Size Exclusion Chromatography (FSEC)***

Fluorescence Size Exclusion Chromatography (FSEC) was done according to a previously described protocol [43]. Briefly, the membrane fraction was collected and re-suspended in 25 mM PIPES buffer, pH 7.2, containing the detergent of interest (DDM or C<sub>12</sub>E<sub>9</sub>). The sample was shaken gently at 4°C for one hour before it was spun down using an ultracentrifuge at 150,000x g for one hour to remove any insoluble material. 150 µL of the supernatant was injected into a pre-equilibrated sephacryl S400 gel filtration column, and GFP fluorescence was monitored.

### **Detergent Screen**

The membrane fraction was collected and re-suspended in PIPES buffer (25 mM NaPIPES pH 7.2, 140 mM NaCl, and protease inhibitor cocktail) such that the membrane fraction collected from 10 g of mammalian cells was resuspended in 5 mL buffer (10x membrane suspension). The screen was performed in a 96-well plate allowing us to screen over 70 detergents plus controls. 10 µL of membrane suspension was mixed with 10 µL of 10% detergent stocks and 80 µL of PIPES buffer in the cold room for one hour. Then, the reaction mix was spun down using a microultracentrifuge (Beckman Optima TLX, TLA100 rotor) at 150,000 xg for one hour. The supernatant was collected for analysis by SDS PAGE and BN PAGE.

### **NiNTA Purification of DmPIEZO**

Insect cells were resuspended in PIPES buffer in a ratio of 1 g of insect cells in 5 mL of buffer. Cells were lysed by sonication (Misonix, S 4000). The membrane fraction was isolated according to the protocol described above. Membranes were solubilized using 1% Fos Choline 16, 1% Sodium Dodecanoyl Sarcosine, or 1% Triton X-100 by shaking in the cold room for one hour. Upon ultracentrifugation to remove insoluble material, samples were loaded onto 0.5 mL pre-equilibrated NiNTA column (GE healthcare). Beads were washed with fifteen column volume of 40 mM imidazole wash buffer prior to elution using 4 column volume of 250 mM imidazole.

### Purification of the Full Length Mouse PIEZO1-GST

Purification was done according to the previously described protocol [42]. Mammalian cells were resuspended in 25 mM PIPES buffer, pH 7.2, in a ratio of 1 g of cells in 5 mL of buffer. Cells were lysed using a Dounce homogenizer. The cell lysate was spun down using a tabletop centrifuge at 14,000 rpm for 10 minutes in the cold room. The supernatant was collected and spun down using an ultracentrifuge at 150,000x g for one hour to collect the membrane fraction. Membranes were washed two more times with PIPES buffer. After the third wash, membranes were resuspended and solubilized in detergent-lipid PIPES buffer (25 mM NaPIPES pH 7.2, 140 mM NaCl, 0.6% CHAPS, 0.14% PC, PIC, 2.5 mM DTT), and shaken for one hour in cold room. Insoluble material was removed by ultracentrifugation. Samples were incubated with 0.5 mL pre-equilibrated glutathione sepharose 4b (GE healthcare) beads overnight in the cold room. The following day, beads were washed with 30 column volume of detergent-lipid PIPES buffer prior to elution with 2 mL of 100 mM glutathione. Glutathione was removed by running the elution fraction on a desalting column that had been pre-equilibrated with detergent-lipid PIPES buffer.

### Amphipols Exchange

Purified mouse PIEZO1-GST protein was brought to a concentration of 0.1 mg/mL. Purified Piezo was supplemented with Amphipol A8-35 to a final concentration of 2.5 mg/mL. Samples were mixed by gentle rotation for 4 hours at 4°C. Then, 40 mg/ mL of activated Biobeads SM-2 resin (Biorad, Inc.) was added. Amphipol-detergent exchange was allowed to occur over night at 4°C. The following day, samples were spun down to

remove the Biobeads. Supernatant was collected and loaded on a Superose 6 gel filtration column that had been pre-equilibrated with buffer without detergent to remove excess amphipols. Non-void protein peak was collected and used for the EM imaging.

### **Negative Staining and Electron Microscopy Imaging**

Mouse PIEZO1-GST was brought to a concentration of 30 µg/ mL. 4 µL of protein sample was spotted onto a glow-discharged holey carbon TEM grid (TED PELLA, Inc.) for one minute. Excess liquid was removed by gently blotting it with a filter paper. 4 µL of 1% uranyl acetate was applied to the EM grid and incubated at room temperature for one minute. Excess uranyl acetate was blotted using filter paper, and the TEM grid was allowed to dry out. Once dried, this TEM grid is ready for observation under the TEM.

### **Data Collection and Image Processing**

Three hundred negative stain images of full length mouse PIEZO1-GST were taken at 26,000x (4.1 Å / pixel) using a 120kV Tecnai T12 electron microscope (FEI, Inc.) at 2µm defocus. From these images, 2488 half-butterfly particles were manually selected using EMAN2 [59], and an initial 2D classification was used to create an initial reference model.

**Author Contribution**

A.K. led the project, designed the experiments, performed most of the experiments, analyzed the results, and wrote the manuscript. H.W. performed data processing downstream of the particle selection for the three-dimensional electron microscopy (3D EM) reconstruction. J.V. developed a method to isolate a stable cell line where mouse PIEZO1-GST plasmid is integrated into the cell's genome, and A.H. maintained the resulting stable cell line. P.N. was involved in the maintenance of cell culture. A.M. provided technical expertise on single particle EM data collection. D.C.R. provided the mentorship, funding, and editorial services.

**References:**

1. Rontgen, W.C., *On a New Kind of Rays*. Science, 1896. **3**(59): p. 227-31.
2. Bragg, W.H. and W.L. Bragg, *The Reflection of X-rays by Crystals*. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 1913. **88**(605): p. 428-438.
3. Bragg, W.L., *The structure of some crystals as indicated by their diffraction of x-rays*. Proceedings of the Royal Society of London Series a-Containing Papers of a Mathematical and Physical Character, 1913. **89**(610): p. 248-277.
4. Pauling, L., R.B. Corey, and H.R. Branson, *The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain*. Proc Natl Acad Sci U S A, 1951. **37**(4): p. 205-11.
5. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid*. Nature, 1953. **171**(4356): p. 737-8.
6. Perutz, M.F., et al., *Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis*. Nature, 1960. **185**(4711): p. 416-22.
7. Kendrew, J.C., et al., *A three-dimensional model of the myoglobin molecule obtained by x-ray analysis*. Nature, 1958. **181**(4610): p. 662-6.
8. Blake, C.C., et al., *Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution*. Nature, 1965. **206**(4986): p. 757-61.
9. Kartha, G., J. Bello, and D. Harker, *Tertiary structure of ribonuclease*. Nature, 1967. **213**(5079): p. 862-5.
10. Wyckoff, H.W., et al., *The structure of ribonuclease-S at 3.5 A resolution*. J Biol Chem, 1967. **242**(17): p. 3984-8.

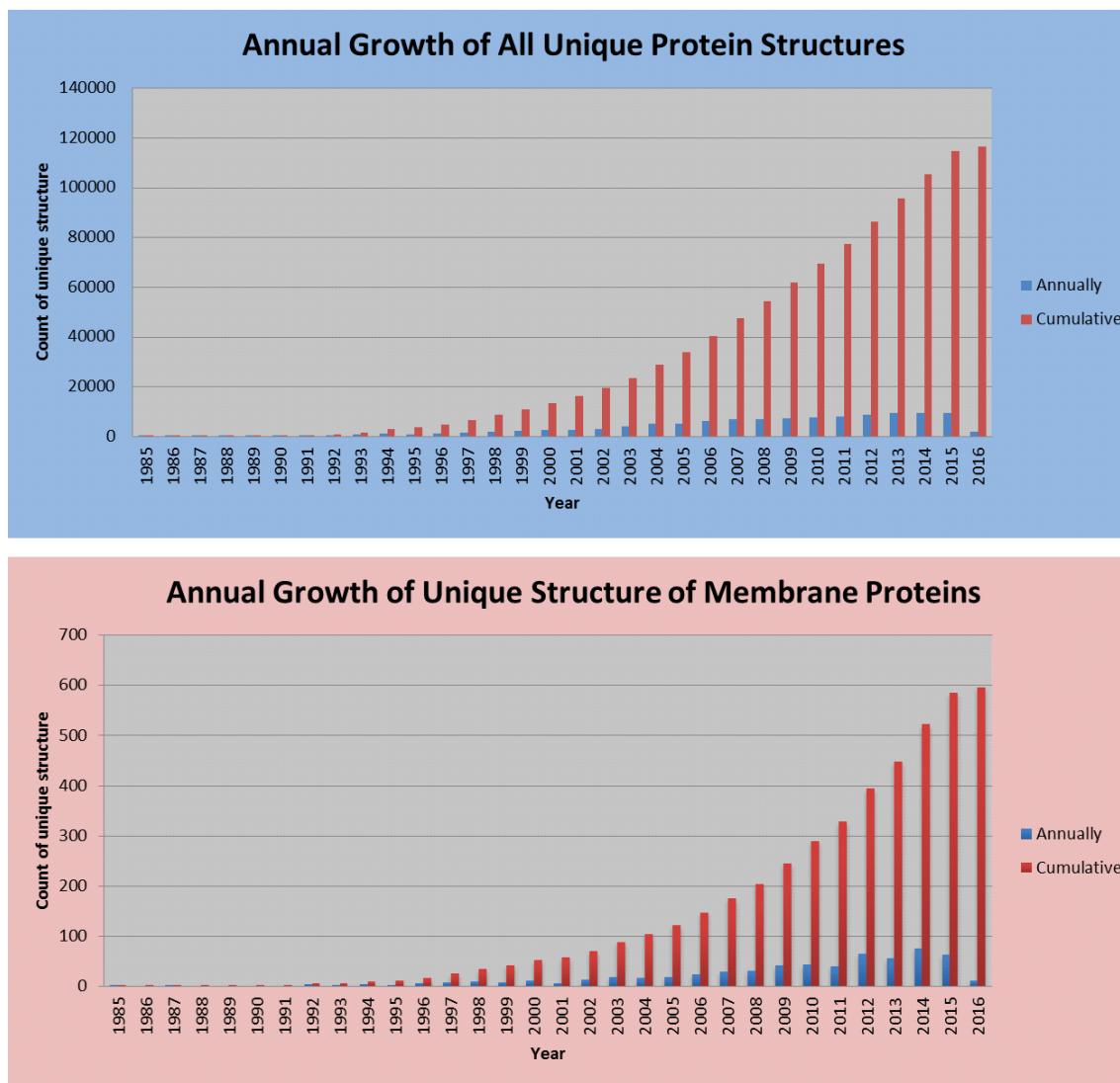
11. Matthews, B.W., et al., *Three-dimensional structure of tosyl-alpha-chymotrypsin*. Nature, 1967. **214**(5089): p. 652-6.
12. Drenth, J., et al., *Structure of papain*. Nature, 1968. **218**(5145): p. 929-32.
13. Lipscomb, W.N., et al., *The structure of carboxypeptidase A. IX. The x-ray diffraction results in the light of the chemical sequence*. Proc Natl Acad Sci U S A, 1969. **64**(1): p. 28-35.
14. Wright, C.S., R.A. Alden, and J. Kraut, *Structure of subtilisin BPN' at 2.5 angstrom resolution*. Nature, 1969. **221**(5177): p. 235-42.
15. Shi, Y., *A glimpse of structural biology through X-ray crystallography*. Cell, 2014. **159**(5): p. 995-1014.
16. Deisenhofer, J., et al., *Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3A resolution*. Nature, 1985. **318**(6047): p. 618-24.
17. Tate, C.G., *Overexpression of mammalian integral membrane proteins for structural studies*. FEBS Lett, 2001. **504**(3): p. 94-8.
18. Cao, E., et al., *TRPV1 structures in distinct conformations reveal activation mechanisms*. Nature, 2013. **504**(7478): p. 113-8.
19. Liao, M., et al., *Structure of the TRPV1 ion channel determined by electron cryo-microscopy*. Nature, 2013. **504**(7478): p. 107-12.
20. Linn, S. and W. Arber, *Host specificity of DNA produced by Escherichia coli, X. In vitro restriction of phage fd replicative form*. Proc Natl Acad Sci U S A, 1968. **59**(4): p. 1300-6.
21. Smith, H.O. and K.W. Wilcox, *A restriction enzyme from Hemophilus influenzae. I. Purification and general properties*. J Mol Biol, 1970. **51**(2): p. 379-91.

22. Jackson, D.A., R.H. Symons, and P. Berg, *Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli*. Proc Natl Acad Sci U S A, 1972. **69**(10): p. 2904-9.
23. Bolivar, F., et al., *Construction and characterization of new cloning vehicles. II. A multipurpose cloning system*. Gene, 1977. **2**(2): p. 95-113.
24. Yanisch-Perron, C., J. Vieira, and J. Messing, *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors*. Gene, 1985. **33**(1): p. 103-19.
25. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
26. Maxam, A.M. and W. Gilbert, *A new method for sequencing DNA*. Proc Natl Acad Sci U S A, 1977. **74**(2): p. 560-4.
27. Mullis, K.B. and F.A. Faloona, *Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction*. Methods Enzymol, 1987. **155**: p. 335-50.
28. Baltimore, D., *RNA-dependent DNA polymerase in virions of RNA tumour viruses*. Nature, 1970. **226**(5252): p. 1209-11.
29. Temin, H.M. and S. Mizutani, *RNA-dependent DNA polymerase in virions of Rous sarcoma virus*. Nature, 1970. **226**(5252): p. 1211-3.
30. Opekarova, M. and W. Tanner, *Specific lipid requirements of membrane proteins--a putative bottleneck in heterologous expression*. Biochim Biophys Acta, 2003. **1610**(1): p. 11-22.
31. Rosano, G.L. and E.A. Ceccarelli, *Recombinant protein expression in Escherichia coli: advances and challenges*. Front Microbiol, 2014. **5**: p. 172.

32. Cregg, J.M., et al., *Expression in the yeast Pichia pastoris*. Methods Enzymol, 2009. **463**: p. 169-89.
33. Cereghino, G.P., et al., *Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris*. Curr Opin Biotechnol, 2002. **13**(4): p. 329-32.
34. Naatsaari, L., et al., *Deletion of the Pichia pastoris KU70 homologue facilitates platform strain generation for gene expression and synthetic biology*. PLoS One, 2012. **7**(6): p. e39720.
35. Byrne, B., *Pichia pastoris as an expression host for membrane protein structural biology*. Curr Opin Struct Biol, 2015. **32**: p. 9-17.
36. Ho, J.D., et al., *Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance*. Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7437-42.
37. Aller, S.G., et al., *Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding*. Science, 2009. **323**(5922): p. 1718-22.
38. Brohawn, S.G., J. del Marmol, and R. MacKinnon, *Crystal structure of the human K2P TRAAK, a lipid- and mechano-sensitive K<sup>+</sup> ion channel*. Science, 2012. **335**(6067): p. 436-41.
39. Hino, T., et al., *G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody*. Nature, 2012. **482**(7384): p. 237-40.
40. Hou, X., et al., *Crystal structure of the calcium release-activated calcium channel Orai*. Science, 2012. **338**(6112): p. 1308-13.
41. Andrell, J. and C.G. Tate, *Overexpression of membrane proteins in mammalian cells for structural studies*. Mol Membr Biol, 2013. **30**(1): p. 52-63.
42. Coste, B., et al., *Piezo proteins are pore-forming subunits of mechanically activated channels*. Nature, 2012. **483**(7388): p. 176-81.

43. Kawate, T. and E. Gouaux, *Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins*. Structure, 2006. **14**(4): p. 673-81.
44. Kost, T.A. and J.P. Condreay, *Recombinant baculoviruses as mammalian cell gene-delivery vectors*. Trends Biotechnol, 2002. **20**(4): p. 173-80.
45. Scott, M.J., et al., *Efficient expression of secreted proteases via recombinant BacMam virus*. Protein Expr Purif, 2007. **52**(1): p. 104-16.
46. Dukkipati, A., et al., *BacMam system for high-level expression of recombinant soluble and membrane glycoproteins for structural studies*. Protein Expr Purif, 2008. **62**(2): p. 160-70.
47. Condreay, J.P., et al., *Transient and stable gene expression in mammalian cells transduced with a recombinant baculovirus vector*. Proc Natl Acad Sci U S A, 1999. **96**(1): p. 127-32.
48. Ramos, L., et al., *Rapid expression of recombinant proteins in modified CHO cells using the baculovirus system*. Cytotechnology, 2002. **38**(1-3): p. 37-41.
49. Reeves, P.J., et al., *Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line*. Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13419-24.
50. Kim, S.E., et al., *The role of Drosophila Piezo in mechanical nociception*. Nature, 2012. **483**(7388): p. 209-12.
51. Li, X., et al., *Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM*. Nat Methods, 2013. **10**(6): p. 584-90.
52. Brilot, A.F., et al., *Beam-induced motion of vitrified specimen on holey carbon film*. J Struct Biol, 2012. **177**(3): p. 630-7.

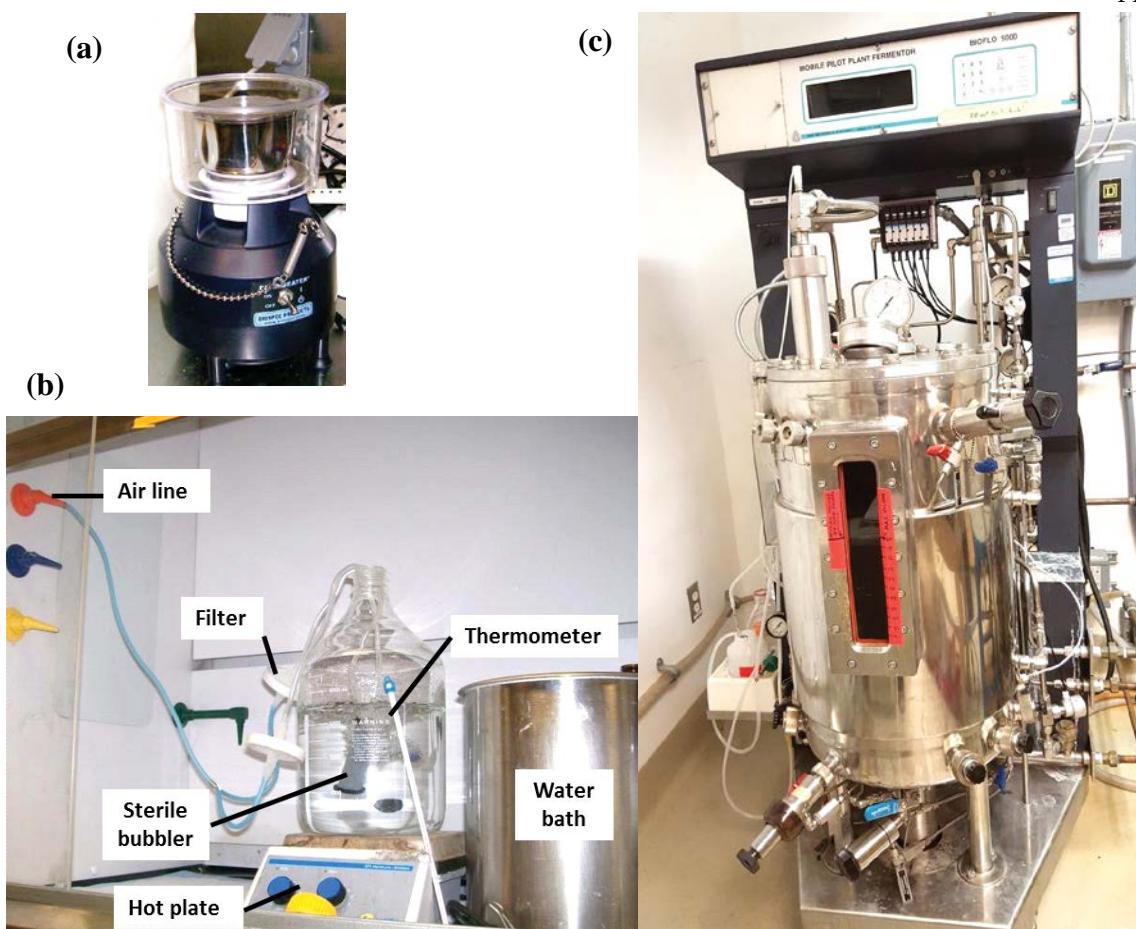
53. Campbell, M.G., et al., *Movies of ice-embedded particles enhance resolution in electron cryo-microscopy*. Structure, 2012. **20**(11): p. 1823-8.
54. Sigworth, F.J., et al., *An introduction to maximum-likelihood methods in cryo-EM*. Methods Enzymol, 2010. **482**: p. 263-94.
55. Lyumkis, D., et al., *Likelihood-based classification of cryo-EM images using FREALIGN*. J Struct Biol, 2013. **183**(3): p. 377-88.
56. Ge, J., et al., *Architecture of the mammalian mechanosensitive Piezo1 channel*. Nature, 2015.
57. Kamajaya, A., et al., *The structure of a conserved piezo channel domain reveals a topologically distinct beta sandwich fold*. Structure, 2014. **22**(10): p. 1520-7.
58. Sleight, S.C., et al., *In-Fusion BioBrick assembly and re-engineering*. Nucleic Acids Res, 2010. **38**(8): p. 2624-36.
59. Tang, G., et al., *EMAN2: an extensible image processing suite for electron microscopy*. J Struct Biol, 2007. **157**(1): p. 38-46.



**Figure 1. Annual growth of unique protein structures deposited in the Protein Data Bank (PDB).** Statistics as of March 3rd 2016. Notice that the number of membrane protein structure solved and deposited on the PDB is fewer than a thousand membrane proteins, whereas the total number of protein structures in the PDB is over one hundred thousand.

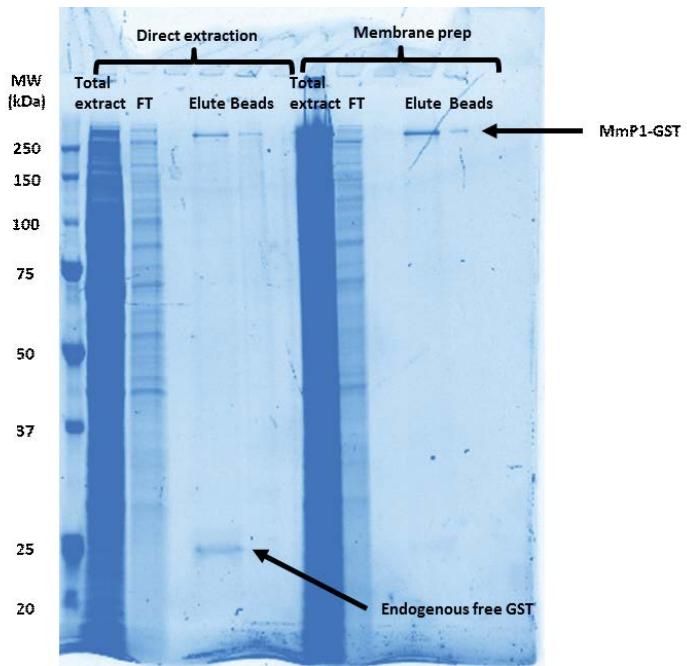
Source:

- PDB statistics ([http://www.rcsb.org/pdb/static.do?p=general\\_information/pdb\\_statistics/index.html](http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html))
- (<http://blanco.biomol.uci.edu/mpstruc/query>)

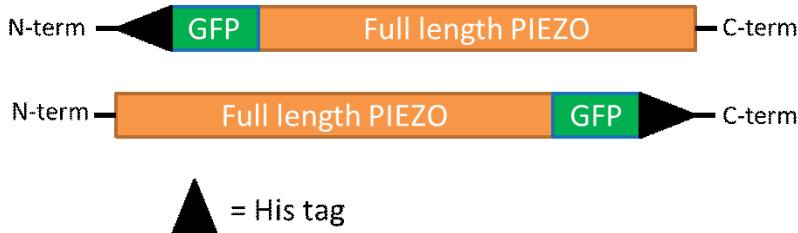


**Figure 2. Equipment set-ups for protein expression using the Pichia system:**

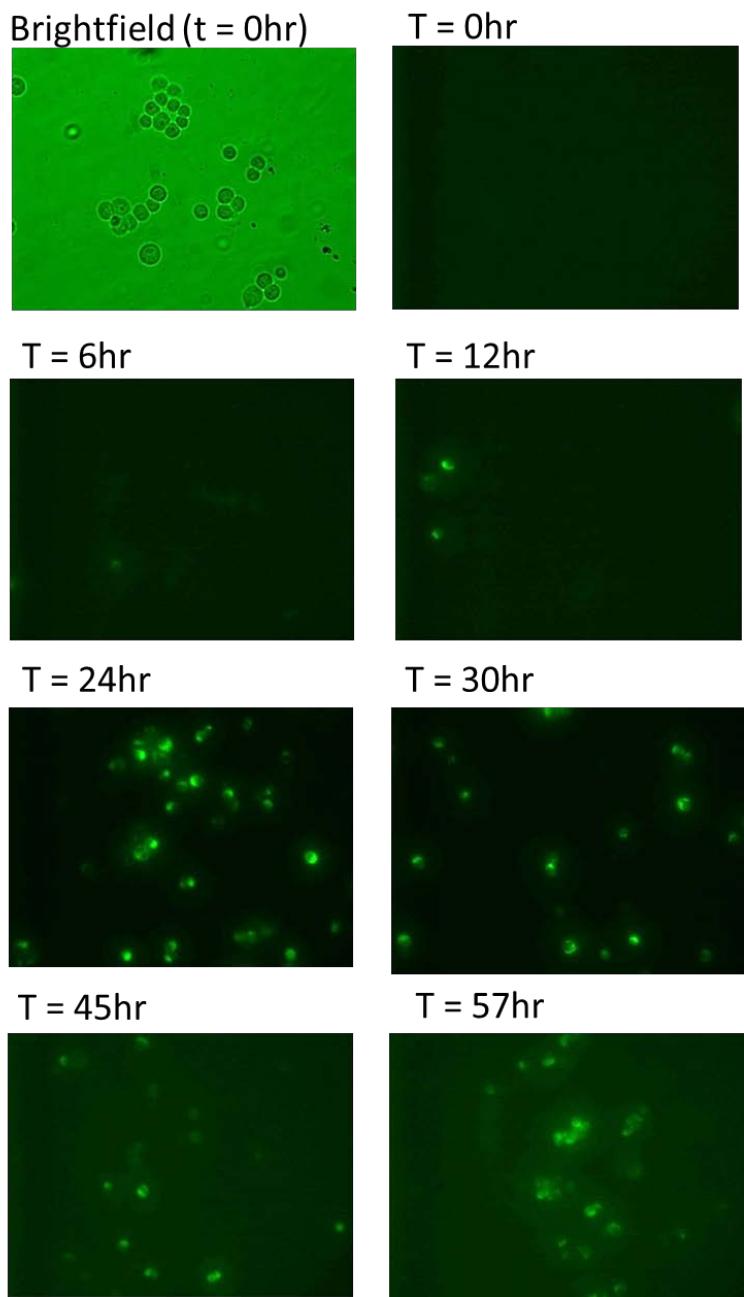
- (a) Bead beater, an apparatus that we used to lyse Pichia cells for protein purification
- (b) Mini-bioreactor that we assembled for medium-scale (2 – 3 L) protein expression using the Pichia expression system
- (c) Bioreactor that was used for large scale (>30 L) protein expression in Pichia



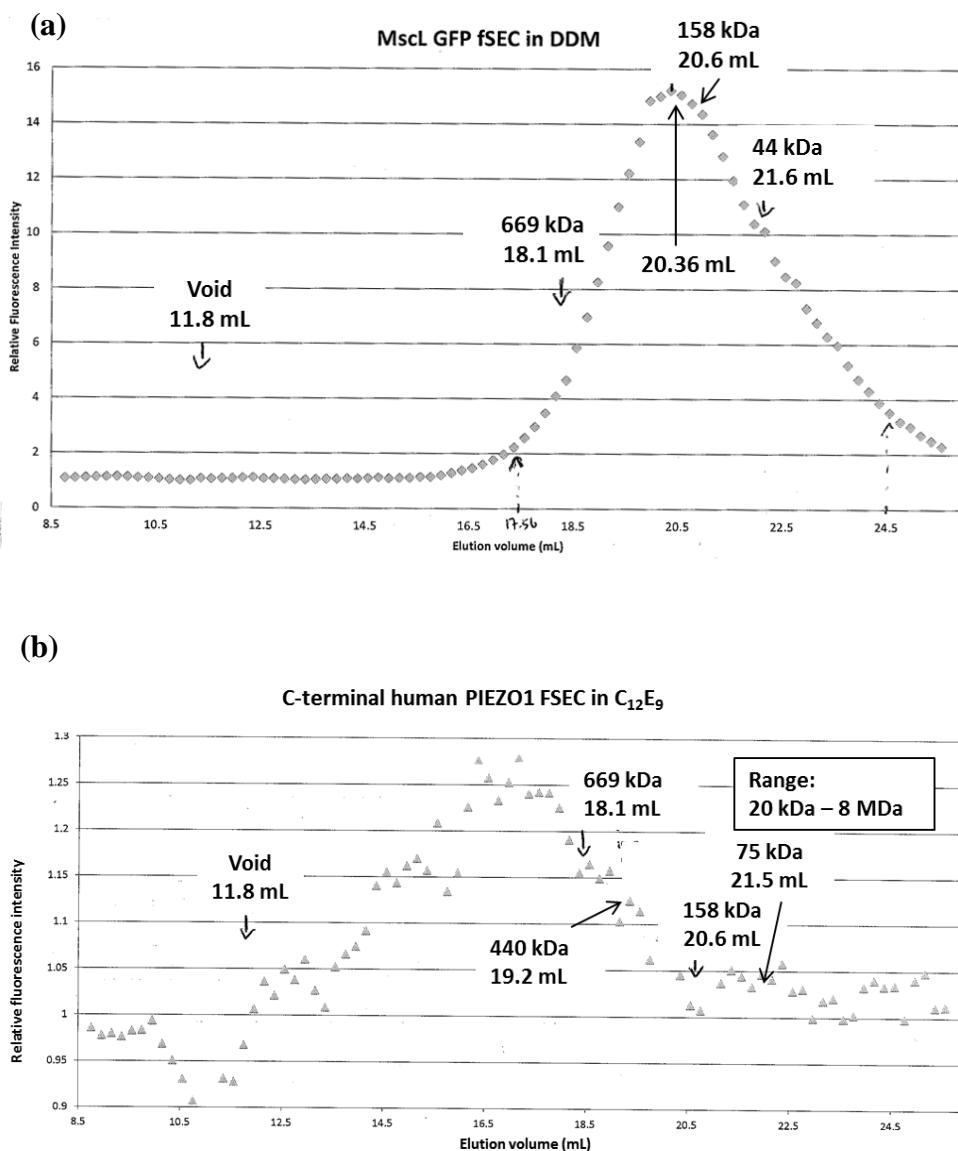
**Figure 3. Representative SDS PAGE gel for the purification of mouse PIEZO1-GST.** Two purification methods were compared: direct membrane protein extraction after cell lysis and protein extraction from the isolated membrane fraction. Elution fraction for direct membrane protein extraction method contains a 25 kDa band (the endogenous free GST). This band is no longer present when an additional membrane isolation procedure was incorporated into the purification protocol.



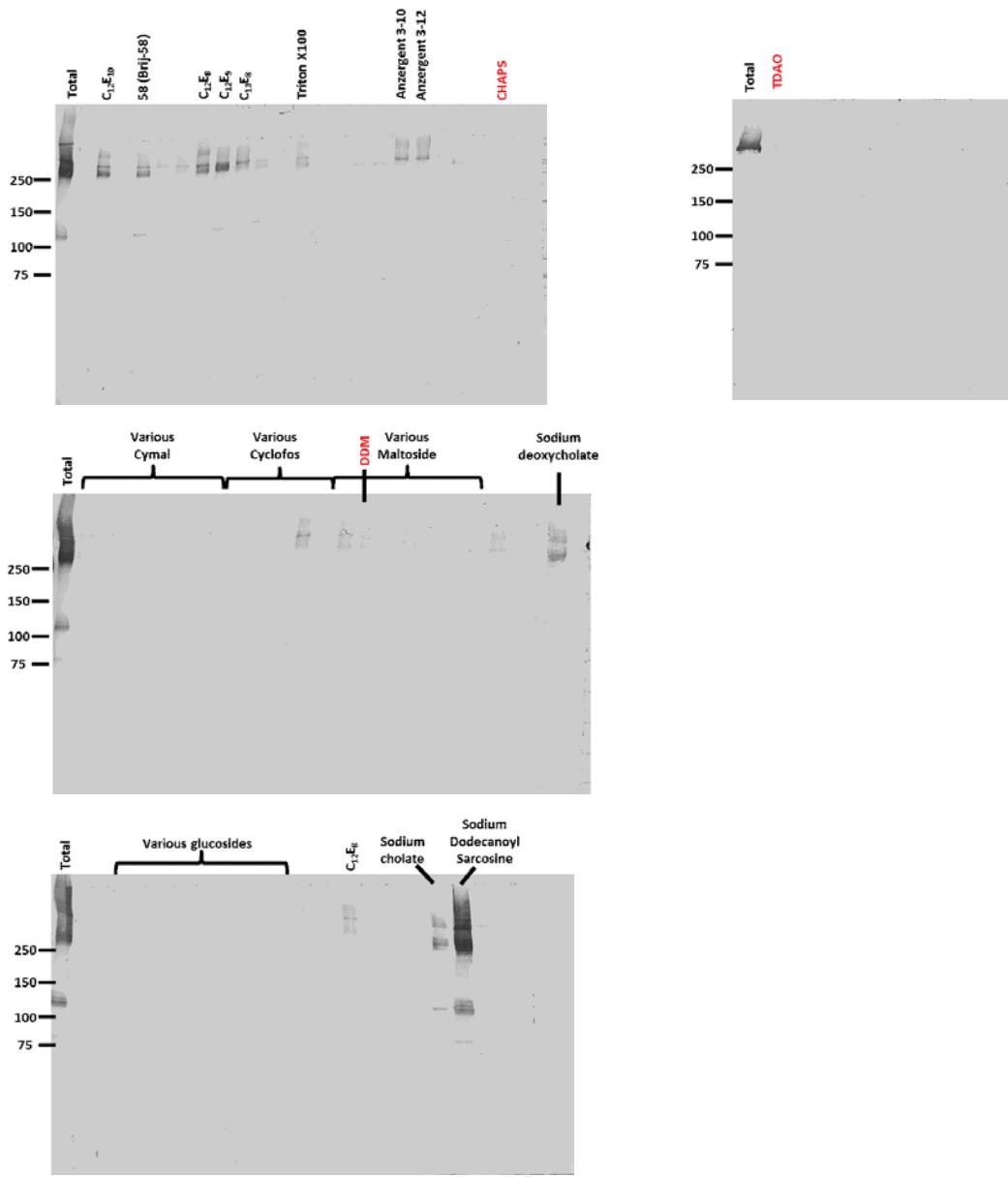
**Figure 4. Fluorescence Size Exclusion Chromatography (FSEC) Piezo constructs.** Two variants of FSEC constructs were made for each Piezo homolog: N-terminal GFP fusion construct and C-terminal GFP fusion construct. Both constructs contain poly-His affinity tag.



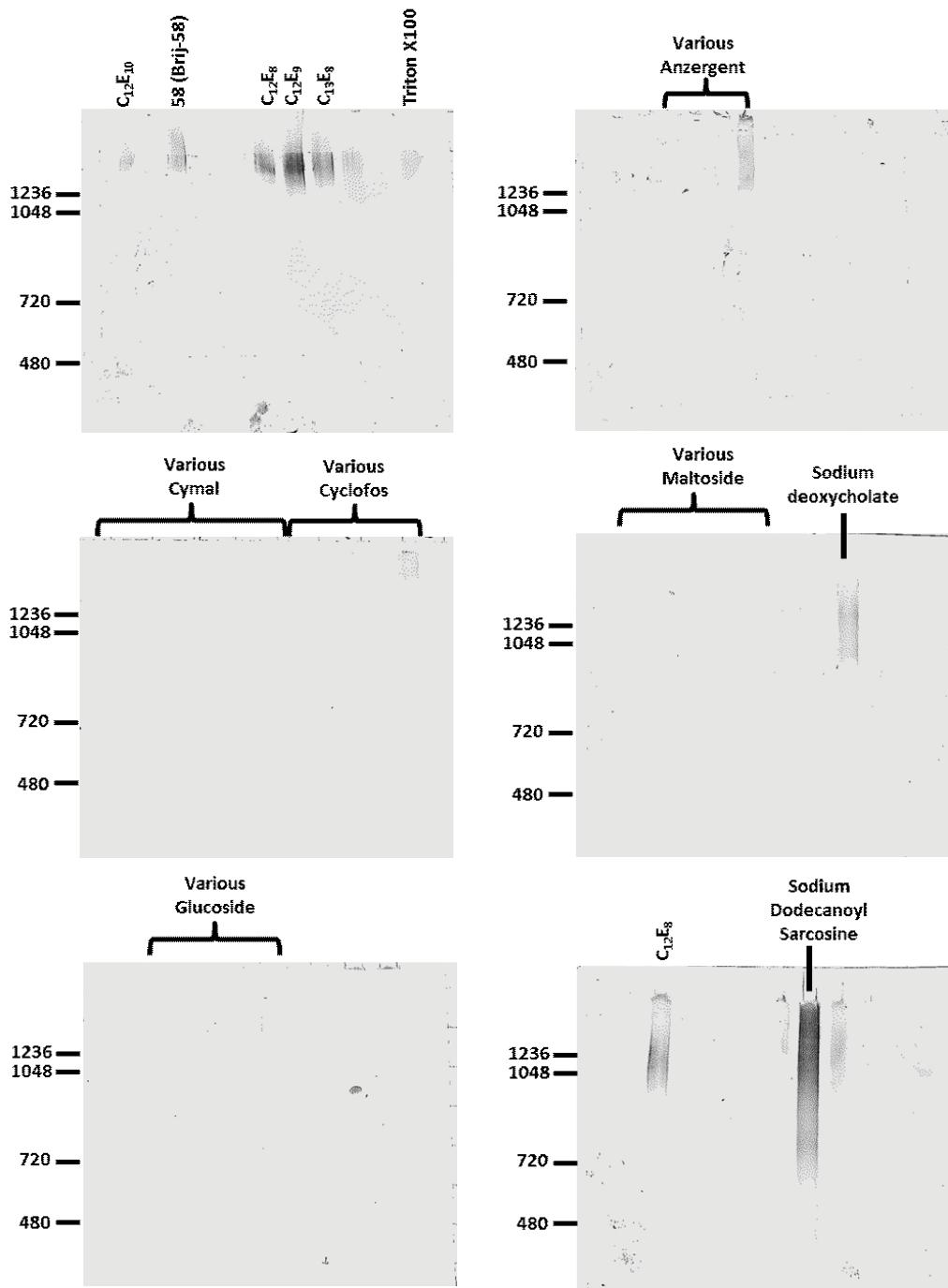
**Figure 5. GFP fluorescence of HEK293 cells transfected with C-terminal GFP human PIEZO1 FSEC construct using Lipofectamine 2000.** After 45 hours, there was prominent background GFP fluorescence. Presumably, it was due to a combination of cell lysis and degradation of the FSEC construct, which releases free GFP molecules.



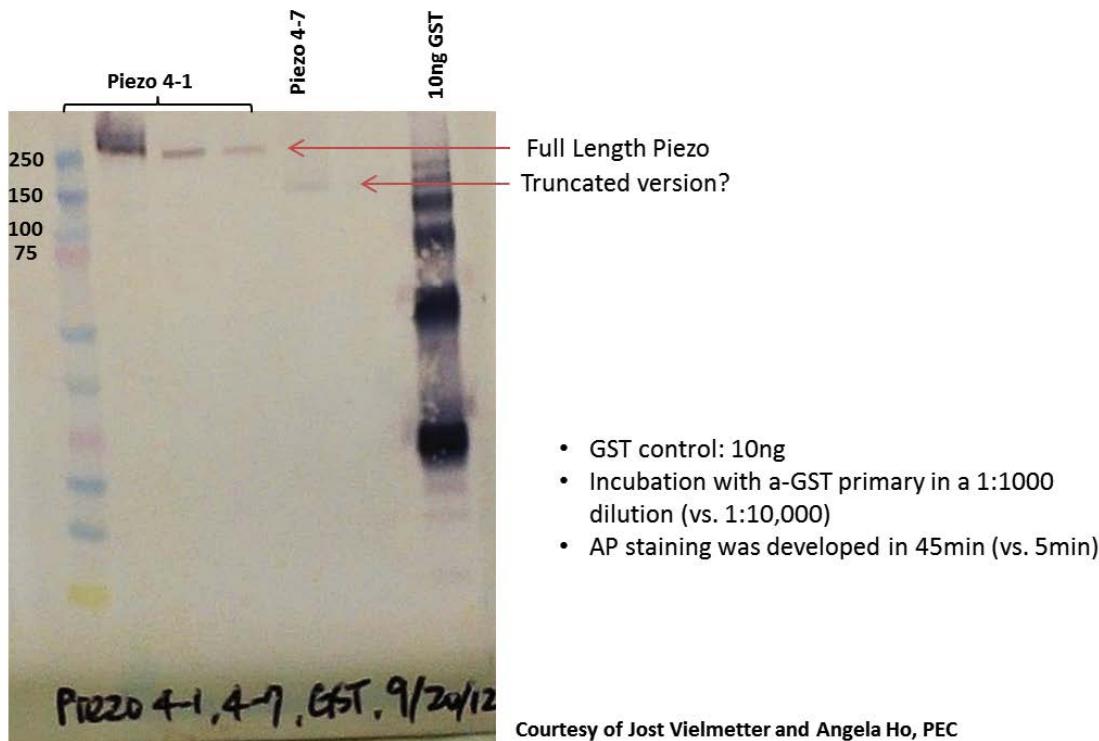
**Figure 6. Chromatogram of proteins on a Sephadryl S400 gel filtration column.** (a) Chromatogram profile of MscL-GFP which served as control. DDM-solubilized MscL GFP FSEC showed a peak at 20.3 mL, which coincided with the absorbance profile of the purified protein at 280 nm. (b) Chromatogram profile of C-terminal human PIEZO1 FSEC construct which was solubilized in C<sub>12</sub>E<sub>9</sub>. Arrows indicate retention volume of molecular weight standard (GE Healthcare): Thyroglobulin (669 kDa), Ferritin (440 kDa), Aldolase (150 kDa), Conalbumin (75 kDa), and Ovalbumin (44 kDa).



**Figure 7. Representative western blots (from SDS-PAGE) of the detergent screen for Piezo expressed in suspension-adapted mammalian culture.** This particular result is from the C-terminal FSEC full length human PIEZO1 construct. Sodium deoxycholate, sodium cholate, C<sub>12</sub>E<sub>8</sub>, C<sub>12</sub>E<sub>9</sub>, C<sub>13</sub>E<sub>10</sub>, and Sodium dodecanoyl sarcosine were the only detergents which could efficiently extract full length Piezo. CHAPS, which was used previously by Coste et al. [42], as well as commonly used mild detergents (**red**) did not seem to extract Piezo constructs.

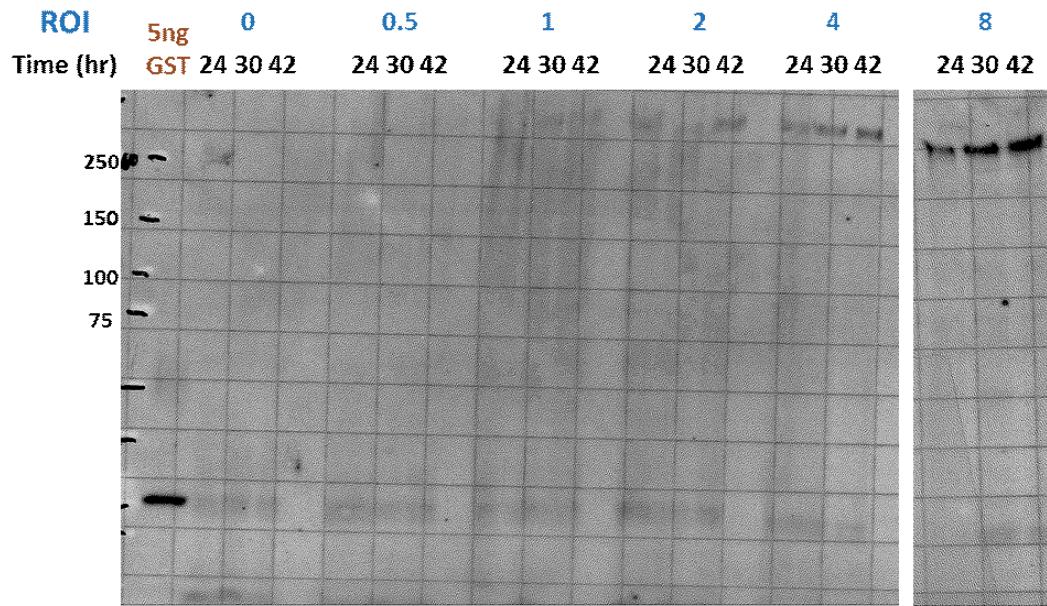


**Figure 8. Representative BN PAGE western blot images of the detergent screen for human PIEZO1 expressed in suspension-adapted mammalian culture.** Among detergents which could efficiently extract full length Piezo, only C<sub>12</sub>E<sub>8</sub>, C<sub>12</sub>E<sub>9</sub>, and C<sub>13</sub>E<sub>10</sub> kept the protein monodispersed.

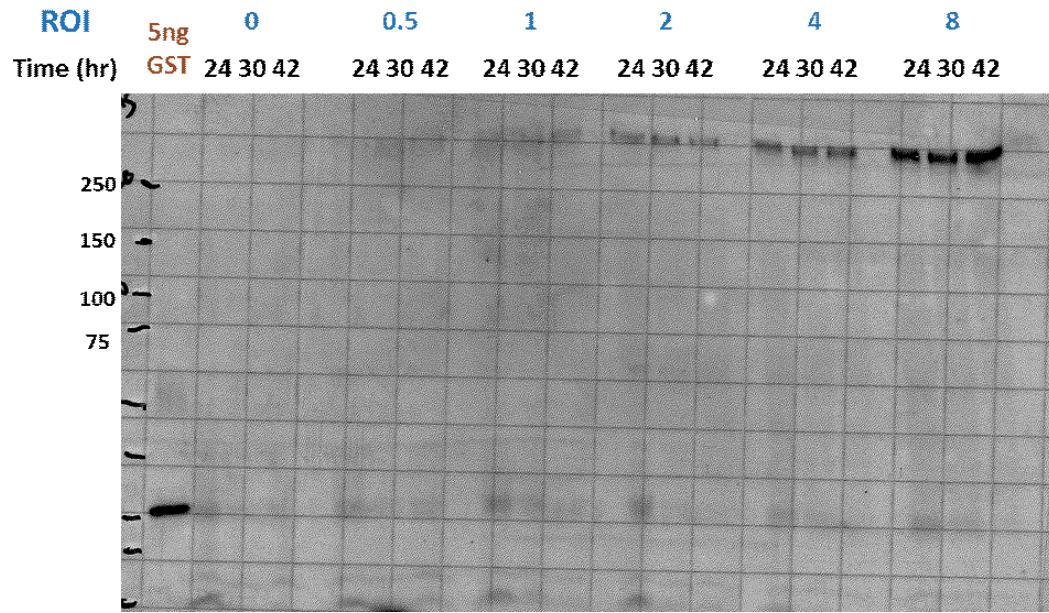


**Figure 9. Western blot result of protein test expression for the stable cell lines.** Piezo 4-1 cell line showed a much stronger signal than Piezo 4-7 despite the weaker GFP fluorescence signal. Nevertheless, the Piezo signal was still very weak compared to a 10ng GST loading control. This data was obtained by Angela Ho (Protein Expression Center, PEC) and used under the permission of the PEC.

### BacMam Piezo test expression on HEK293 F cells

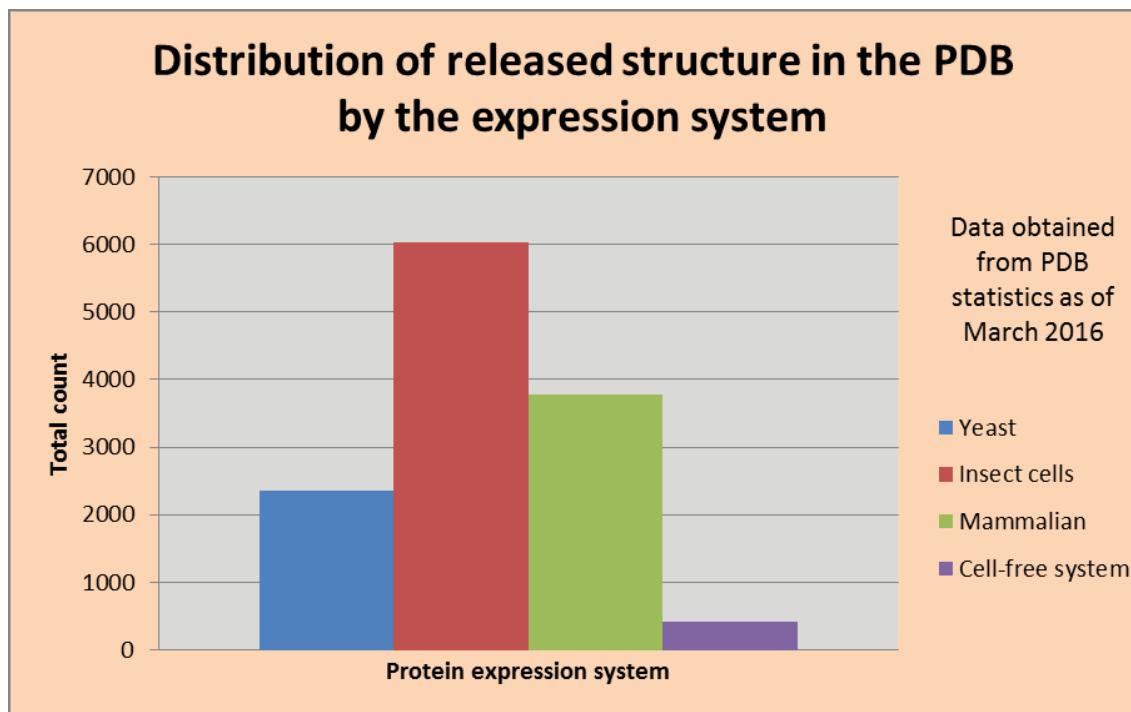


### BacMam Piezo test expression on HEK293 GnTI<sup>-</sup> cells



**Figure 10. Western blot images of small scale test expression of mouse PIEZO1-GST in HEK293F and HEK293 GnTI- cells using the BacMam system.**

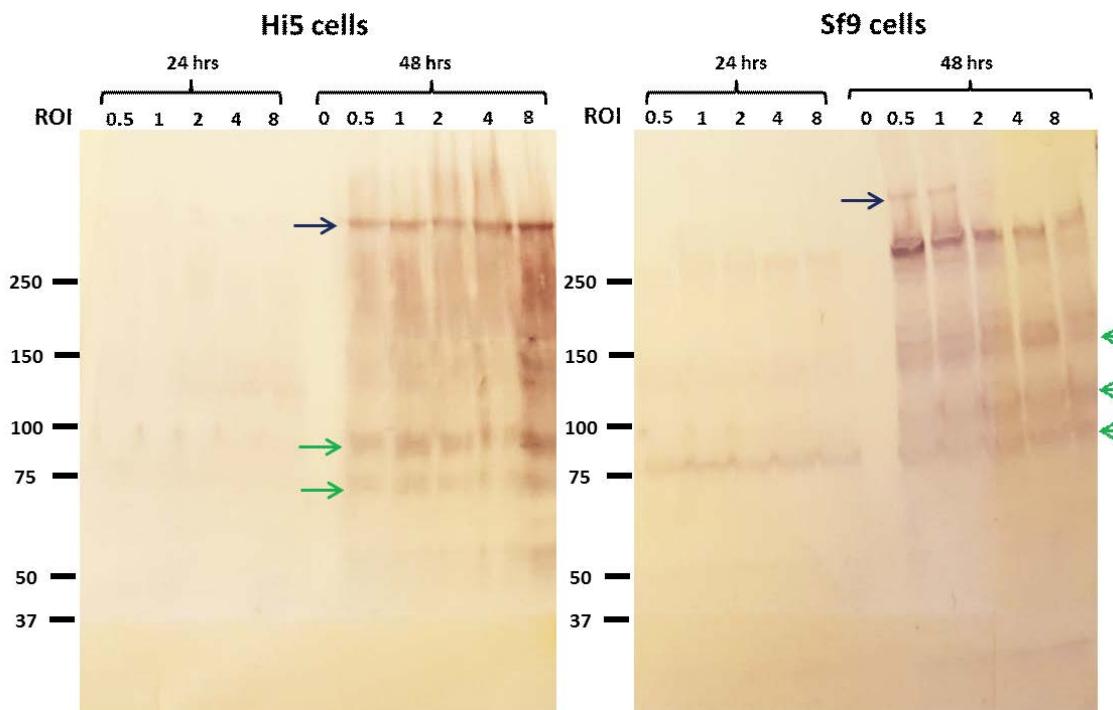
Both cell lines showed similar protein expression level at most ROI (ratio of infection). At ROI of 8, the viral titer is almost 50% of the total culture volume.



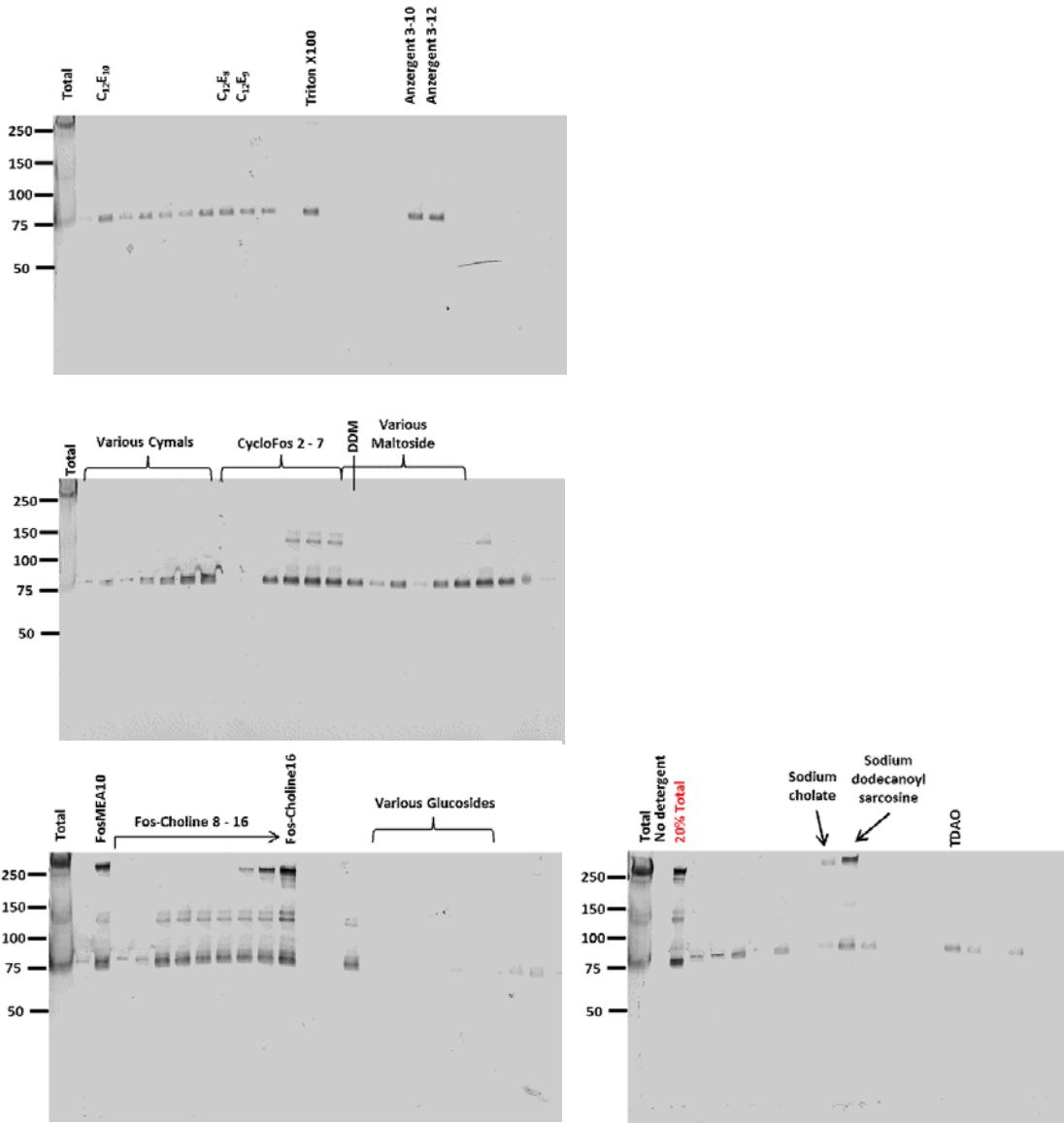
**Figure 11. Counts of all structures deposited on PDB by host expression organisms.** Statistics as of March 2016. The statistics only compare nonbacterial protein expression systems. As a comparison, total structures deposited on PDB obtained from the bacterial protein expression systems are more than 98,000 entries.

Source:

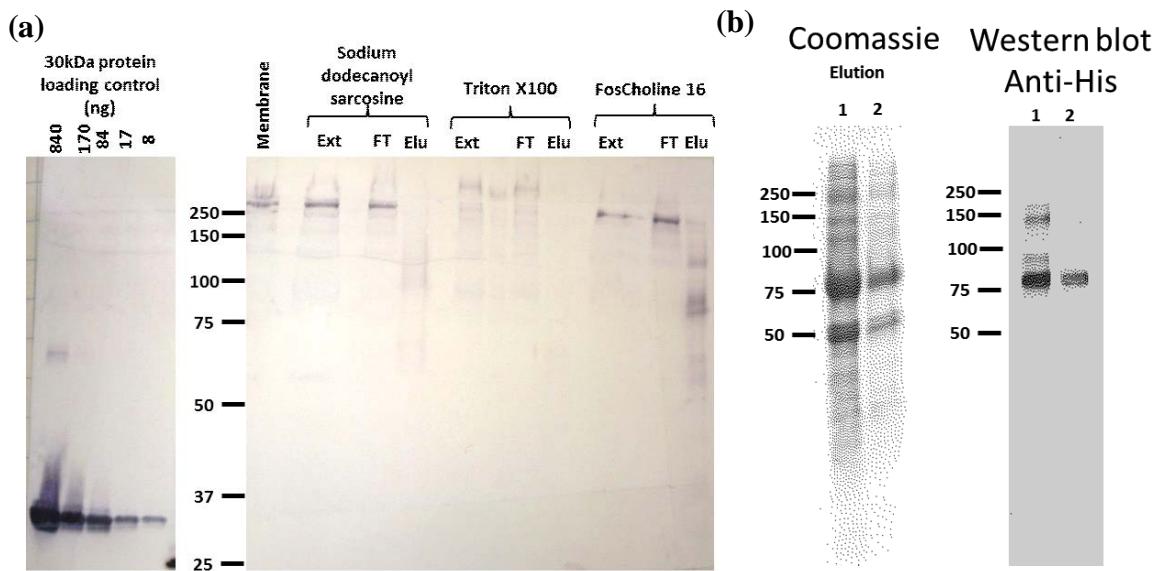
PDB statistics ([http://www.rcsb.org/pdb/static.do?p=general\\_information/pdb\\_statistics/index.html](http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html))



**Figure 12. Western blot result of Drosophila PIEZO FSEC test expression in Hi5 and Sf9 insect cell lines.** Drosophila PIEZO expression was absent at 24 hours. Protein expression was observed at 48 hours in both Sf9 and Hi-5 insect cell lines.

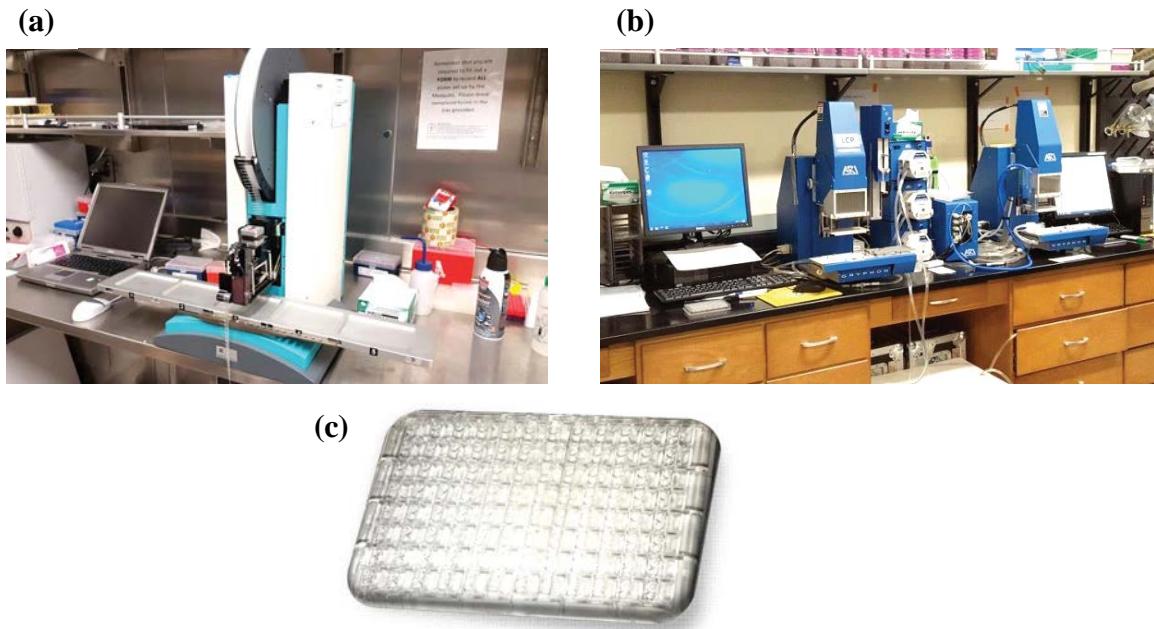


**Figure 13. SDS-PAGE Western blot results for detergent extraction screen of full length Drosophila PIEZO in insect cells.** Most detergents were not able to efficiently extract full length Drosophila PIEZO with the exception of Fos-Choline-16 and sodium dodecanoyl sarcosine. Most detergents were able to extract the 75 kDa peptide fragment. Comparing extraction efficiency among Fos-Choline detergents, the longer the hydrocarbon chain, the more efficient it is in extracting the larger full length Piezo protein; extraction efficiency for the 75kDa fragment, however, is comparable across all Fos-Choline detergents.

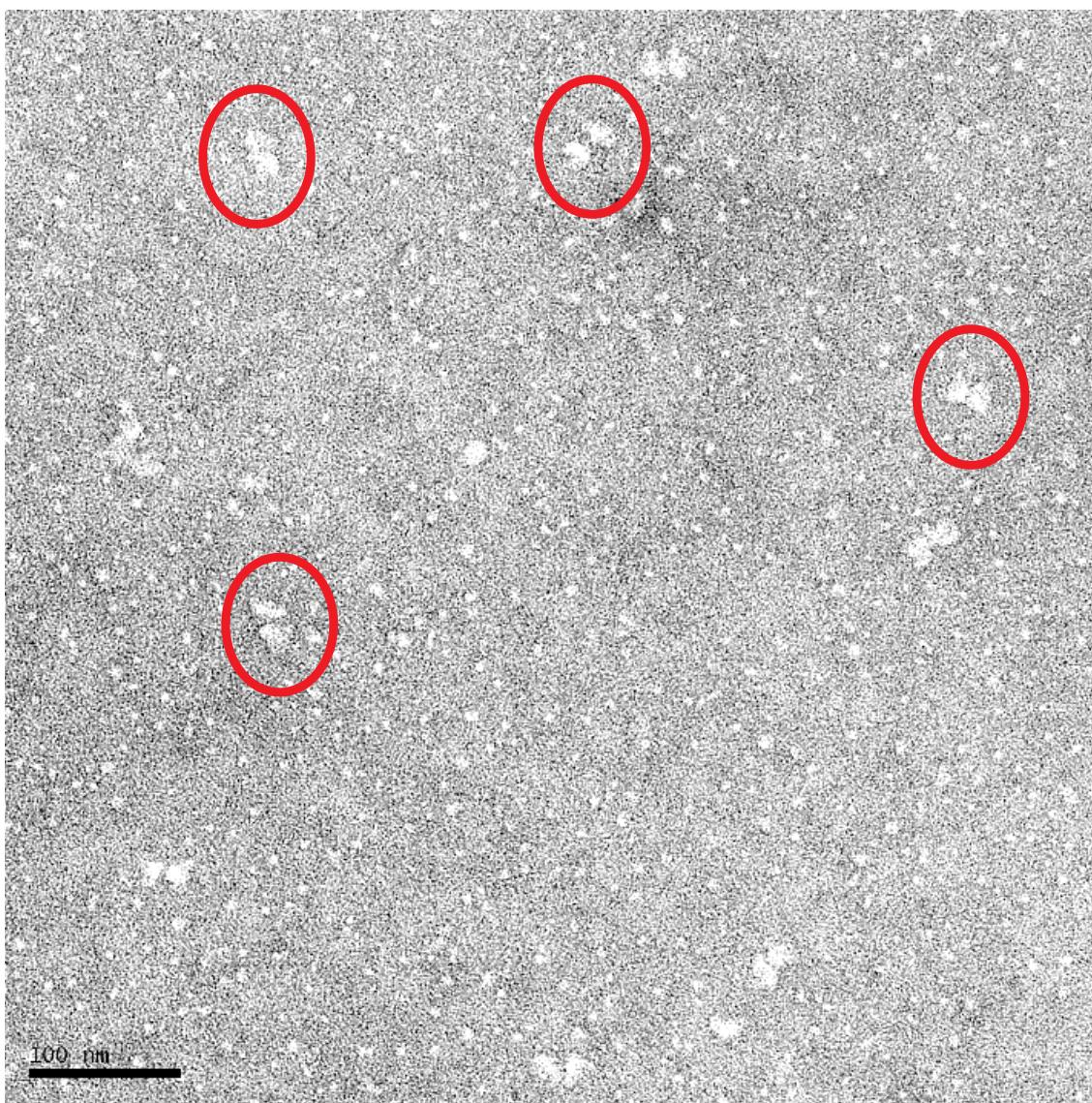


**Figure 14. Medium scale purification of full length Drosophila PIEZO.** (a)

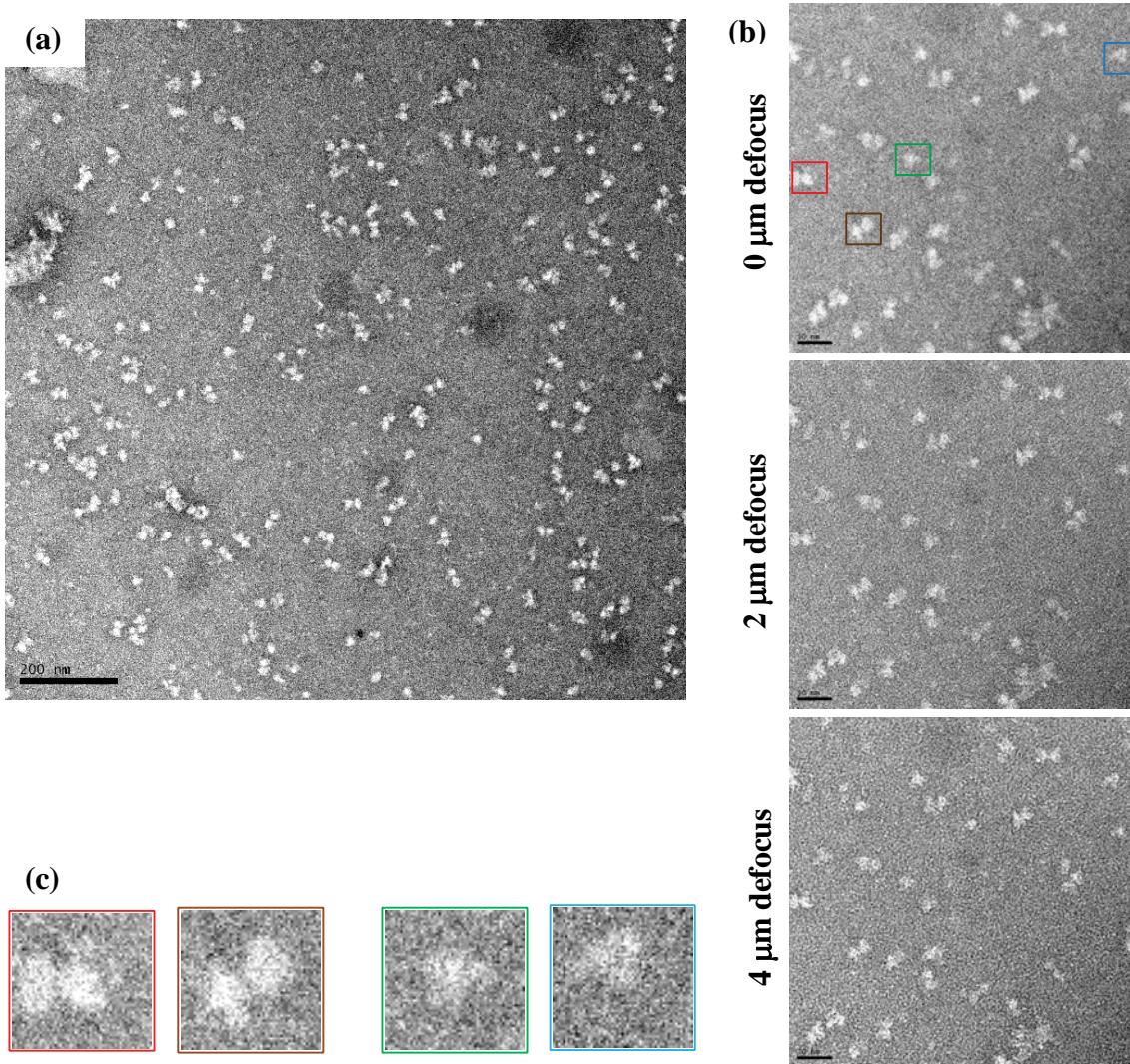
Western blot results for full length Drosophila PIEZO purified in different detergents. It appears that most of the extracted full length Piezo (Ext) failed to bind to the NiNTA beads under our purification conditions; most of them came out in the flow through (FT) fraction. (b) Coomassie and western blot images of the concentrated elution fraction of full length *Drosophila* PIEZO purified in Fos-Choline-16 detergent. Elution was done by applying elution buffer twice as the following: Beads were incubated in two column-volume of elution buffer for 5 minutes; this was collected as the elution fraction #1 (lane 1). This step was repeated one more time using fresh elution buffer; this was collected as elution fraction #2 (lane 2).



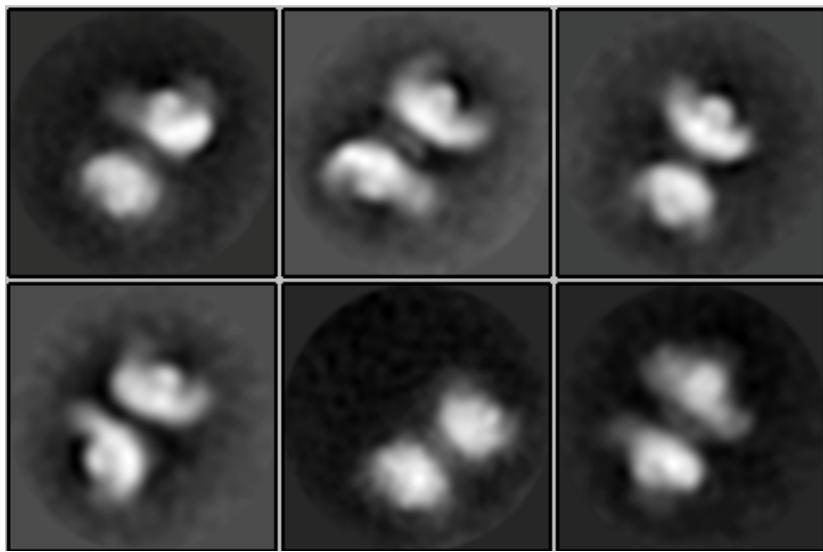
**Figure 15. Crystallization screen set-up at the Molecular Observatory (MOLOBS).** (a) Mosquito (TTP Labtech, Inc.) (b) Gryphon (ARI, Inc.) (c) intelli-plate® 96 for sitting drop crystal screen set-up (ARI, Inc.)



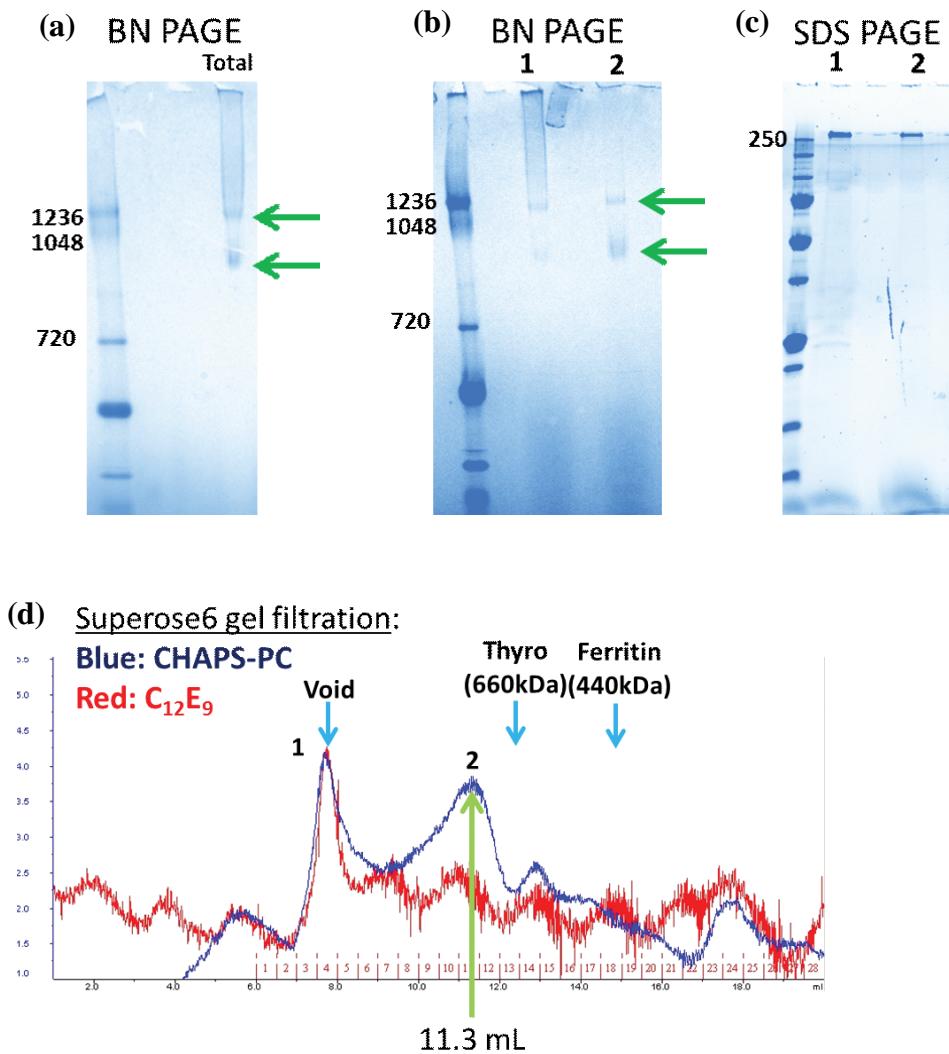
**Figure 16. Initial observation of mouse PIEZO1-GST on Formvar-Carbon supported copper EM grid.** Negative stain EM image of 1:100 dilution mouse PIEZO1-GST (3 µg/mL). Scale bar length is 100 nm. This was the first time we observed mouse PIEZO1-GST particle as butterfly particles (red circles).



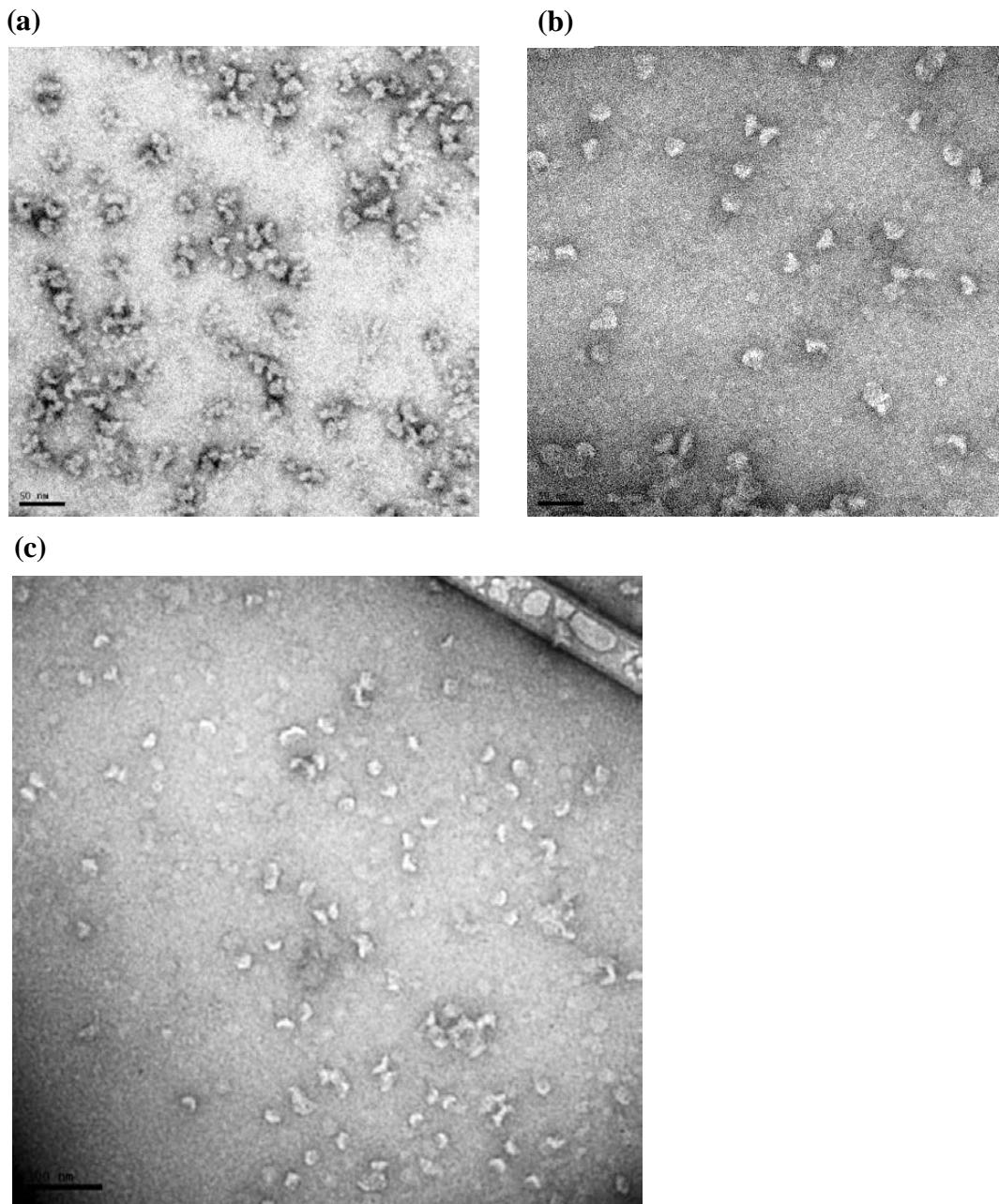
**Figure 17. Representative view of mouse PIEZO1-GST particles on holey carbon EM grid negatively stained with 1% uranyl acetate.** These images show that most of mouse PIEZO1-GST molecules adopt butterfly shape although some particles look like half-butterfly. (a) Image taken at lower magnification showing the overall distribution of Piezo particles at a larger field of view. Scale bar is 200 nm. (b) Images of the same field of view were taken at different defocus level at 42,000x magnification. Scale bar is 50 nm. (c) Representative images of the butterfly particles (**red** and **brown**) and half-butterfly particles (**green** and **blue**).



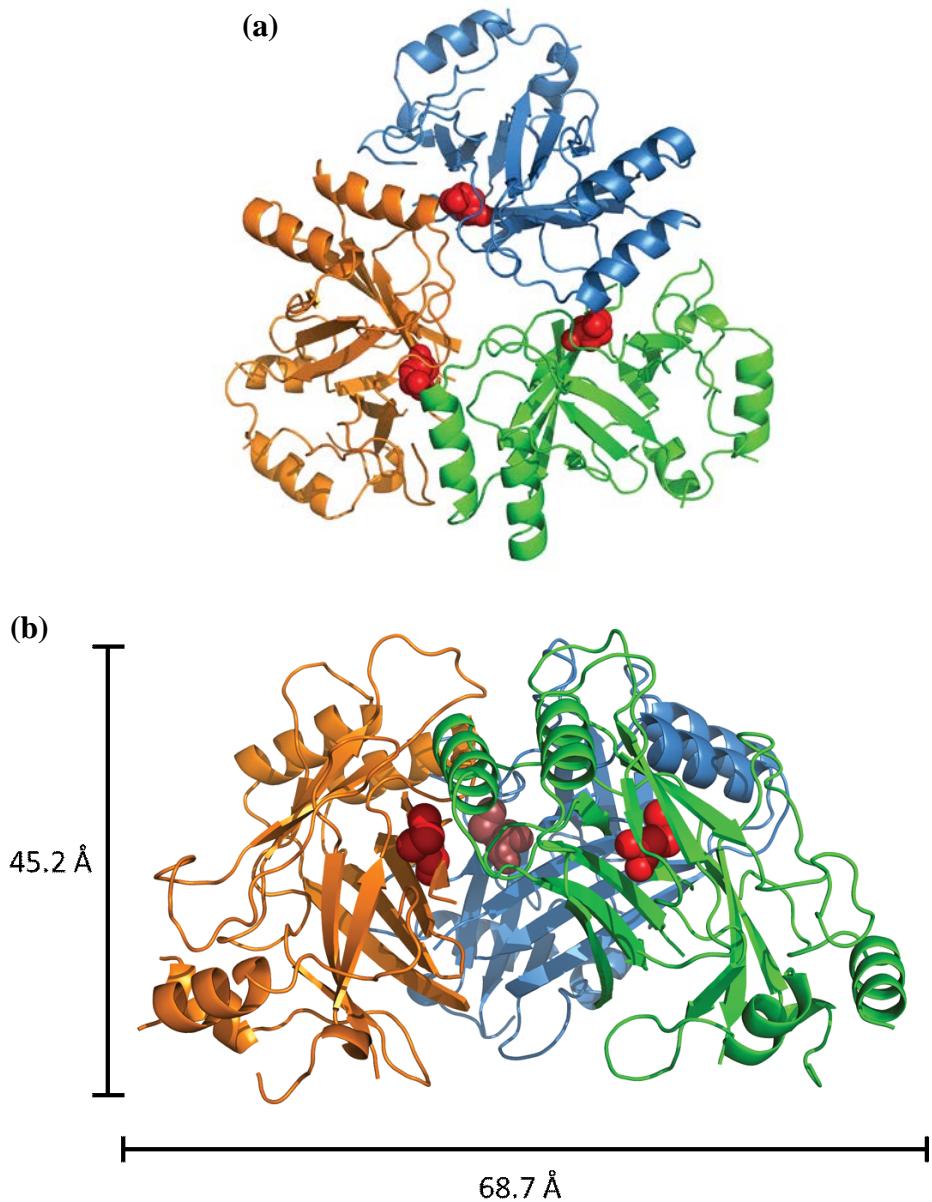
**Figure 18. 2D class average of the butterfly particles from negative stain EM images.** 2994 of butterfly and half-butterfly particles were manually picked to generate this 2D class average. Out of a total 30 classes (the total number of particles divided by 100), 6 classes showed reasonable signals. From these classes the two butterfly wings look identical, which may suggest a dimeric arrangement. Furthermore, it appears that the relative spatial orientation of the two butterfly wings is not uniform, indicating that there is heterogeneity in the connection between the two subunits.



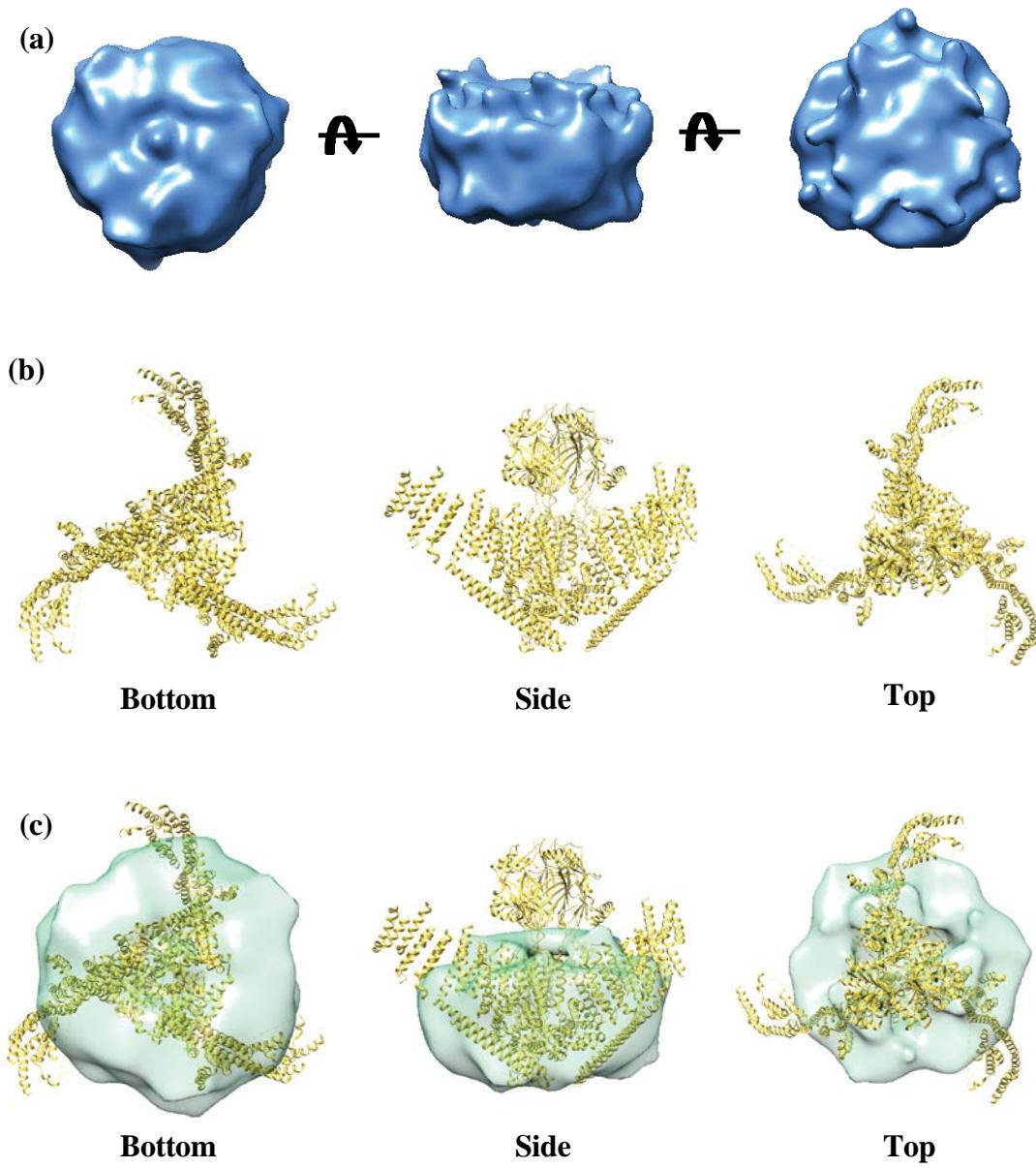
**Figure 19. Purification of mouse PIEZO1-GST for EM imaging.** (a) A representative BN-PAGE result of purified mouse PIEZO1-GST prior to the gel filtration analysis. (b) Mouse PIEZO1-GST on BN PAGE after being separated into two fractions (peak 1 and 2) using a Superose 6 gel filtration column. Peak 1 mainly consisted of the higher molecular weight band and the aggregation, whereas peak 2 was enriched in the lower molecular weight species. (c) Mouse PIEZO1-GST on SDS PAGE for peak 1 and 2, showing that both fractions contained primarily full length Piezo protein. (d) Superose 6 gel filtration profile of mouse PIEZO1-GST. Peak 1 and peak 2 were collected separately for gel analysis.



**Figure 20. Negative stain images of mouse PIEZO1-GST after separation using Superose 6 gel filtration.** Purified mouse PIEZO1-GST was separated into two different fractions: peak 1 and peak 2. (a) Peak 1 mainly consisted of the butterfly particles. Scale bar is 50 nm (b) Peak 2 was enriched in the half-butterfly particles and some butterfly particles. Scale bar is 50 nm. (c) Negative stain image of peak 2 at lower magnification to show a larger field of view. Scale bar is 100 nm.



**Figure 21. Arrangement of *C. elegans* PIEZO soluble domain, CTL2, in the crystal lattice.** The clinically relevant mutation in human PIEZO1 (M2225R) corresponds to M31R in the *C. elegans* CTL2 construct [57]. This methionine residue, displayed as **red spheres**, is located at the interface between two monomers. (a) Top view of *C. elegans* CTL2. (b) Side view of *C. elegans* CTL2. Trimeric CTL2 arrangement is similar to the CED trimer in the full length mouse PIEZO1 EM structure [56].



**Figure 22. 3D reconstruction of mouse PIEZO1-GST trimer from negative stain EM data.** (a) The initial model was built with EMAN2 using a total of 2488 half-butterly particles. (b) Structure of the full length mouse PIEZO1 trimer [56]. (c) Superposition of the mouse PIEZO1 structure into our low resolution initial model. Mouse PIEZO1 structure has similar approximate dimension as the initial model, which we generated from negative stain mouse PIEZO1-GST data.

# **Chapter 4:**

## **Isolation of a Stable C-terminal Piezo Fragment**

## **Background and Motivation**

One ideal goal that structural biologists strive to achieve is to provide mechanistic insights into complex biological processes through the acquisition of molecular snapshots. Typically, one would obtain molecular structures of biomolecules in different functional states, and from these snapshots, mechanistic model(s) can be generated to explain how structural changes of these molecules enable them to carry out their function. One great example to illustrate such a scientific feat is the *alternating access model* of substrate transport by the ATP Binding Cassette (ABC) transporter. The *alternating access model* of substrate transport has actually been around for half a century [1]. BtuCD-F, an ABC transporter which transports vitamin B<sub>12</sub>, has been extensively investigated structurally and biochemically. In addition to the individual structures of the transporter (BtuCD) and the binding protein (BtuF) [2, 3], the structures of the transporter-binding protein complex (BtuCD-F) have been solved in several different states, illustrating the full cycle of the *alternating access model* of substrate transport [4-6]; this was made possible by introducing a few key molecules to the protein mix prior to the crystallization step in order to trap the transporter molecules in a desired conformation.

Finding conditions to trap a mechanosensitive (MS) channel in different conformations for structural study is not trivial; currently, there is not a good way to apply a constant mechanical perturbation to MS channels during sample preparation or the image acquisition process. Several MS channels can be modulated by chemical ligands so that one

can use these ligands to trap these channels in different functional states. Such efforts are exemplified by the two pore-domain (K<sub>2</sub>P) potassium channels, TRAAK and TREK-2. These channels were crystallized in both their native states as well as with known modulators [7-10]. The question of whether chemical activation causes equivalent conformational changes as mechanical activation in these MS channels remains unclear, however. Alternatively, two different states of the bacterial MS channel, MscL, have been determined. The first MscL structure came from the *Mycobacterium tuberculosis* homolog (MtMscL) where this channel assembles as a homopentamer with a narrow pore opening diameter (~ 2 Å) [11]. Subsequently, the structure of the C-terminally truncated MscL homolog from *Staphylococcus aureus* (SaMscL) was determined. Truncated SaMscL assembles as a homotetramer with a pore diameter of 6 Å [12]. The transmembrane helices on the SaMscL structure are more laterally tilted relative to the helices on the MtMscL structure. More recently, using different detergent combinations, structures of an archaeal MscL (*Methanosarcina acetivorans* MscL – MaMscL) were determined in two different conformations [13]; form 1 and form 2 of MaMscL resemble the close-state MtMscL and the expanded intermediate state- SaMscL, respectively. Two distinct conformations have also been observed for another bacterial MS channel, MscS [14, 15]. Together with functional studies, this structural information has been used to infer a general gating model for bacterial MS channels [16].

Across its homologs, the Piezo polypeptide chain consists of at least 2000 amino acids, typically over 2500 amino acids. The Piezo family is an extreme outlier when it comes to the number of amino acids that it possesses compared to most transport proteins as well as most proteins in the known genomes (Figure 1) [17, 18]. At the beginning of my journey,

there was no structural information on Piezo beyond bioinformatics predictions and characterization of the oligomeric state [19, 20]. In 2014, I successfully generated a conserved soluble domain construct of Piezo (CTL2) from the *C. elegans* homolog and determined the first molecular glimpse of this mysterious channel (Chapter 2) [21]. A report appeared showing that removing the entire C-terminal end of this channel, including the CTL2 domain, produced functional channels with altered inactivation kinetics [22]. This suggested that the C-terminal segment of Piezo may function to modulate conductance properties rather than provide the permeation pathway. Multiple sequence alignment of Piezo homologs showed that there is a high degree of sequence conservation at the C-terminal segment [21], suggesting an important role for this segment on the general channel function.

It is not uncommon for a large protein to have a modular construction where different domains have several specialized functions. One of many examples is the bacterial methionine ABC transporter, MetNIQ. MetNIQ is composed of several domains which handle different tasks required for function: the transmembrane domains (TMD), nucleotide binding domains (NBD), and C2 trans-inhibition domains, as well as the substrate binding protein (MetQ) [23, 24]. The C2 trans-inhibition domain provides feedback inhibition which regulates the methionine import process depending on the intracellular concentration of methionine [23, 25]. Furthermore, all members of the Transient Receptor Potential (TRP) super family possess six transmembrane segments with a pore domain located between helix 5 and helix 6 [26]. However, different types of TRP channels possess distinct domain(s) which allows each member to perform its unique function [27].

With a presumptive size of 1.2 MDa, Piezo would be the largest ion channel known to date. We reasoned that Piezo consists of multiple domains with specific functions and so we explored conditions for generating stable fragments of this channel. Here, I will discuss our attempt to isolate such a construct, and outline the potential next steps to further characterize this Piezo fragment.

### ***C-terminal Piezo Constructs Expression and Detergent Extraction Screen***

The degree of sequence conservation across Piezo homologs was used as a guide in choosing the appropriate boundaries to create several C-terminal Piezo constructs. Initially, we made twelve different constructs (numbered #1 - #12) from the Piezo homologs available to us (human PIEZO1, mouse PIEZO1, mouse PIEZO2, *Drosophila* PIEZO, and *C. elegans* PIEZO). The approximate starting residues of each of these constructs are indicated by blue arrows in Figure 2a. In general, each construct differs from each other by approximately one predicted transmembrane segment; we made a few additional variants to either include or exclude regions that are predicted to be large or highly conserved soluble domains. The shortest C-terminal constructs contained the entire CTL2 domain, the only region with known structure at the time. The sizes of constructs we made are listed in Table 1. Subsequently, the Patapoutian group reported a study of chimeric Piezo channel which suggested that the C-terminal segment of Piezo determines the pore properties of the channel [28]. This work took advantage of the differences in conductance properties between *Drosophila* PIEZO and mouse PIEZO1 [20]; mouse PIEZO1 has slower channel inactivation kinetics and is sensitive to inhibition by Ruthenium Red (RR), whereas *Drosophila* PIEZO has faster channel inactivation kinetics and is unaffected by RR. Fusing

the N-terminal region of mouse PIEZO1 up to residue 1973 with the *Drosophila* PIEZO starting at residue 1930 ( $mP1^{1-1973}/dP^{1930-2548}$ ) resulted in a chimeric channel with pore properties resembling the *Drosophila* PIEZO [28]. This suggests that the C-terminal residues (1930 – 2548) may possibly form the core ion permeation pathway. Interestingly, this region ( $dP^{1930}$ ) is very close to the site of our C-terminal construct #6 (starting at  $dP^{1928}$ ).

Initially, we looked at protein expression levels for constructs # 5 to # 8 of human PIEZO1, mouse PIEZO1, and mouse PIEZO2 constructs (Figure 3). In general, the longest of the four constructs (construct # 5) showed much less protein expression relative to the three shorter constructs. In the absence of construct #7 of mouse PIEZO2 (C7), Construct #6 of mouse PIEZO2 (C6) showed the highest protein expression level on a western blot; subsequent western blot analysis revealed that C7 showed comparable protein expression levels to C6. Next, we performed detergent screens using several mild detergents to find the detergent which could optimally solubilize the protein. However, it appeared that none of these detergents could efficiently solubilize the C-terminal constructs (Figure 4). We then tested several more detergents, including Fos-Cholines (Figure 5). These results showed that Fos-Choline-16 solubilized the C-terminal constructs from mouse PIEZO2 (C6 and C7) efficiently. Effective solubilization by Fos-Choline-16 was also observed for other C-terminal Piezo constructs (data not shown). Subsequently, we performed medium scale purification where purified C6 and C7 proteins from the elution fractions of an affinity column chromatography (NiNTA) were immediately analyzed on a gel filtration (Superdex S200) column. Overall, purification of C6 and C7 in Fos-Choline-16 detergent yielded polydisperse mixtures as seen by the gel filtration chromatogram (Figure 6ac). When Fos-

Choline-16 was exchanged with DDM during the purification procedure, the protein yield is somewhat lower and the gel filtration chromatogram profile shifts to a slightly longer retention volume (Figure 6bd). Analysis of these chromatogram peaks revealed that there are multiple bands, which can be attributed to potential contamination or protein degradation (Figure 7a). Interestingly, purification of C7 resulted in a higher molecular weight band that did not yield a signal on the western blot (Figure 7b).

We also investigated the protein expression levels of the C-terminal constructs created from *Drosophila* PIEZO (D6, D7, D8) and *C. elegans* PIEZO (E7 and E8). The relative protein expression levels were assessed based on the relative signal intensity on the western blot in comparison to the expression levels of C6, C7, and C8. The expression level of one *C. elegans* C-terminal construct (E8) caught our attention immediately since it was much higher than any other constructs (Figure 8a). From the detergent screen results, we learned that E8 can also be extracted efficiently by Fos-Choline detergents (Figure 8b).

Overall, we learned that C-terminal Piezo constructs can be extracted efficiently from the *E. coli* membrane using Fos-Choline detergents, especially Fos-Choline-16. We also learned that extraction and purification of C-terminal Piezo constructs in Fos-Choline-16 detergent does not drive the protein to large non-specific aggregates as evidenced by the absence of a peak at the void volume on the gel filtration chromatogram. However, purification of these constructs is still far from optimum since there are multiple bands on the gel. We have not systematically explored the identities of these bands to determine whether they are degradation products or non-Piezo contaminant.

### Isolation of a Stable Fragment of Mouse PIEZO1

In one sample preparation of full length mouse PIEZO1-GST, we observed that the purified protein stored at 4 °C appeared to be proteolyzing as evidenced by the appearance by day 2 of fragments with reduced molecular weights (Figure 9a). We followed the degradation progress by taking small aliquots of protein each day and running these samples on SDS PAGE. By day 7, full length mouse PIEZO1-GST was completely cleaved into a peptide fragment with an approximate molecular mass of 95 kDa. This 95 kDa fragment was relatively stable even after another week of incubation at 4 °C. Western blot analysis using an anti-GST antibody suggested that this peptide fragment corresponded to the C-terminal Piezo region since the GST tag is attached to the C-terminal end of the protein (Figure 9b). After several attempts to replicate the natural degradation of mouse PIEZO1-GST, we were unable to get a consistent degradation results; we later noticed that the stock solution used to make the purification buffers was contaminated with what appeared to be biofilm-like material (picture not shown). We performed a limited tryptic digest to see if we could obtain a similar fragment under controlled conditions. At a molar ratio of 1:10 (trypsin:protein), a C-terminal fragment was generated with a comparable molecular weight to the one we obtained from the natural proteolysis (Figure 10). N-terminal peptide sequencing of the isolated peptide fragment revealed that this fragment started at residue H1924, directly after an arginine (R1923) trypsin cleavage site. Without a GST tag, this peptide fragment alone consists of 623 amino acids with calculated molecular weight of 70 kDa.

## C-terminal Fragment in the Context of the Full Length Mouse PIEZO1 Structure

The recently solved structure of mouse PIEZO1 revealed a trimeric assembly of Piezo subunits (Figure 11a) [29]. Overall, this medium (4.8 Å) resolution Piezo structure provided a general model for an intact Piezo channel molecule. However, the model is incomplete. Only ~1000 amino acid residues of mouse PIEZO1 could be modelled of the total ~2500 residues; the region corresponding to the CTL2 domain best fits the electron density (i.e. this region has the highest local resolution). Connectivity could be established for what is assigned as the last 600 amino acids of Piezo (less than 25% of total number of amino acids). This region includes the C-terminal domain, an inner helix, 250 residues of CTL2 domain (CED), an outer helix, anchor helix, and a few more helices in front of the anchor helix. The rest of the helices in the model were built into low resolution electron density without clear connection to each other (marked in black on Figure 11b). Whether these helices are connected to the C-terminal piece remains unclear. The end of the connectivity for the mouse PIEZO1 model is located directly upstream of Tm 35 in our membrane topology prediction (Figure 11c). On the same membrane topology prediction, the stable C-terminal mouse PIEZO1 fragment that we isolated starts at approximately 37 amino acids upstream of the F1961, the first residue where connection can be traced into the C-terminal end of the full length mouse PIEZO1 structure. According to the full length Piezo model, our C-terminal Piezo fragment makes up the central core of an intact Piezo channel, which includes the ion permeation pathway. Given the presence of these regions, we asked the question of whether this C-terminal fragment alone is functional.

## Mouse PIEZO1 C-terminal Fragment can be Expressed and Localized to the Plasma Membrane

Functional characterization of a putative mechanosensing molecule such as the mouse PIEZO1 C-terminal fragment necessitates benefits from localization in the plasma membrane. Accordingly, I designed Piezo constructs with a C-terminal GFP-6xHis tag under a CMV promoter for constitutive mammalian expression system. In Human Embryonic Kidney (HEK293T) cells, expression of the GFP-fused C-terminal mouse PIEZO1 fragment showed a similar green fluorescence localization pattern to the full length mouse PIEZO1-GFP (Figure 12a). The GFP fluorescence forms a ‘crescent moon’ shape, which may indicate the cluster of fluorescence protein molecules that are trapped within the endomembrane due to the overexpression of this protein. GFP fluorescence can be seen around the cellular border of some cells. This resembles the localization pattern of a non-mechanosensitive membrane protein control, GFP-fused ATP Binding Cassette (ABC) exporter protein (Atm1). Meanwhile, expression of free GFP alone results in a green fluorescence signal uniformly distributed throughout the cell. Cellular fractionation experiments further revealed that both the C-terminal and full length mouse PIEZO1 are found in the membrane (mem) fraction, but not the cytosolic (cyt) fraction (Figure 12b). Together, these observations suggest that the C-terminal Piezo fragment is localized to a membrane. However, it is unclear whether this membrane localization includes the plasma membrane.

To further explore the possibility that our C-terminal construct can be localized to the plasma membrane, we performed live-cell immunocytochemistry (ICC) staining. For this purpose, we created Piezo constructs with an extracellular epitope tag created by inserting a

Myc tag directly after residue 2336 (Myc 2336), located in the CTL2 domain (Figure 13a); the constructs also have a C-terminal His tag, which serves as an intracellular epitope control, and GFP. The Myc tag can be probed with an anti-Myc primary antibody, followed by a secondary antibody that is conjugated to the Alexa 647 (red) fluorescence dye, as previously described [28]. Thus, if the C-terminal Piezo fragment is localized to the plasma membrane, we expect to see a red signal after ICC staining, assuming that cellular permeability is not compromised during this procedure. To ensure that our assumption holds true, we also performed ICC staining against the intracellular His epitope to test whether cell membrane was compromised during the live-cells ICC staining procedure. In addition to the two Piezo constructs with both Myc 2336 and GFP, we also stained the wildtype Piezo-GFP-His construct which lacks the Myc epitope. In this experiment, cells were grown on cover slips that had been coated with gelatin in a 24-well plate, and the ICC staining was done in triplicate as indicated in Figure 13b.

As expected, cells that did not receive a Piezo plasmid (mock) failed to show either GFP or Alexa 647 fluorescence signals (Figure 14a). This indicates that the ICC staining does not produce an off-target red fluorescence signal on the membrane of native HEK293T cells. In cells expressing the wildtype full length mouse PIEZO1-GFP-His construct, no red fluorescence was observed after staining with an anti-Myc antibody since the protein construct does not have the Myc epitope. Probing with an anti-His antibody, however, showed a red signal that is comparable to the GFP signal only when the cells were permeabilized with 0.4% Triton X-100 (Figure 14b), consistent with the cytoplasmic localization of the His-tag. Full length mouse PIEZO1 Myc 2336 showed a red fluorescence signal when probed with an anti-Myc antibody even without the

permeabilization procedure (Figure 14c). This confirms that full length Piezo can be localized to the plasma membrane as previously shown [28]. The red signal intensity is increased when cells were permeabilized prior to the ICC staining procedure because the antibody molecules can access the additional pool of Piezo molecules trapped within the endomembrane. Significantly, the C-terminal mouse PIEZO1 Myc 2336 construct shows red fluorescence signal even when cells were not permeabilized (Figure 14d). Considering that we did not see any red fluorescence signal in the live-cells (non-permeabilized) upon ICC staining with an anti-His antibody, it is unlikely that the cell membrane is compromised during the ICC staining procedure. Therefore, these results suggest that the C-terminal fragment of mouse PIEZO1 can be localized to the plasma membrane.

Additionally, through the Beckman Institute Biological Imaging Facility at Caltech, we used confocal microscopy to image cells after ICC staining. These studies revealed that full length Piezo Myc 2336 and C-terminal Piezo Myc 2336 were localized to the plasma membrane since red fluorescence signal is only found at the cellular border when non-permeabilized cells were probed with the anti-Myc antibody (Figure 15). Cells permeabilized prior to the ICC staining procedure showed that the localization of red fluorescence signals is comparable to the GFP fluorescence signal in all of the z-stack images. Collectively, these observations indicate that the C-terminal Piezo fragment can be localized to the plasma membrane.

## **Future Directions**

In this chapter, we have described the preparation of a C-terminal fragment from mouse PIEZO1. The cryoEM structure of mouse PIEZO1 indicates that this C-terminal Piezo fragment makes up the central core of the intact Piezo channel, which includes the ion permeation pathway. Cellular fractionation, GFP fluorescence localization, and immunocytochemistry staining results suggest that this fragment can be localized to the plasma membrane. Thus, we have a reason to suspect that the C-terminal fragment of mouse PIEZO1 is properly folded into the native 3D structure which makes up the central core of the intact Piezo channel. Consequently, a reasonable next step would be to investigate whether this C-terminal Piezo fragment is functional. Thus, it will be interesting to see if our C-terminal mouse PIEZO1 fragment can exhibit channel activity comparable to the full length mouse PIEZO1. A recent report concludes that the ‘propeller’ domains of Piezo channel serve as the mechanosensing module [30]. As this region is lacking in our C-terminal fragment, it provides a useful system for further exploring the features of the Piezo channel required for mechanosensitivity.

Furthermore, it will be interesting to establish the oligomeric state of the C-terminal Piezo fragment to assess whether it forms a trimer like the full length Piezo channel. To address this question, we are working to optimize conditions for expression and purification of this fragment. Briefly, C-terminal mouse PIEZO1-GST construct was expressed in suspension adapted Human Embryonic Kidney (HEK293T) cells. After cell lysis and membrane isolation, we screened detergents to optimize extraction and purification of this construct. Again, we observed that Fos-Choline-16 detergent could efficiently extract the C-terminal Piezo protein from the membrane (Figure 16a). In

general, Fos-Choline detergents are considered a harsh detergent since they are often found to extract membrane protein at much higher levels than any other detergents [31-33]. Despite the concern that membrane protein solubilized in the Fos-Choline detergents may be partially denatured, Fos-Choline-12 and Fos-Choline-14 have been used in the structural study of *Escherichia coli* OmpF and MscS, respectively [14, 34]. Furthermore, recent structures of MscS from *E. coli* and *H. pylori*, solubilized in DDM, resemble the previous MscS structure [35].

After extraction and purification using Fos-Choline-16, the C-terminal Piezo-GST construct forms a discrete band on a blue native PAGE (BN-PAGE) (Figure 16b). Purification of this construct resulted in a sample giving one major band at the appropriate molecular mass on a SDS PAGE with minor contaminant bands (Figure 17b). However, we were not able to obtain sufficient amounts of protein for more detailed characterization. As a result, we were not able to obtain unambiguous gel filtration chromatogram profile for this construct, which would help establish the oligomeric state (Figure 17a). When the two peaks were taken for analysis by BN-PAGE, we observed one major band that ran between the 720 and 480 kDa molecular weight markers, as well as one very faint band that ran just slightly below the 480 kDa molecular weight marker (Figure 17c). Considering that the full length mouse PIEZO1-GST protein was previously shown to form prominent dimers of the intact trimeric Piezo channel [29], one may suspect that the top major band may correspond to the dimer of the oligomeric C-terminal Piezo fragment; the single trimeric form of this construct and the dimer of this trimer are expected to have a molecular weight of 291 kDa and 583 kDa, respectively. This offers hope that our C-terminal construct can fold into the native conformation of Piezo central core.

In conclusion, isolation of the stable C-terminal Piezo came to us serendipitously. We have evidence that this construct localizes to the plasma membrane and therefore may assemble into the native structure of the Piezo channel central core. Thus, it will be interesting if one can obtain the structure of this fragment and unequivocally show that this fragment can assemble into the native oligomeric form which resembles the arrangement of the full length Piezo channel. In the end, there is one burning question that awaits an answer: “is this fragment functional?”

### **Methods:**

#### ***Bioinformatics and Cloning of the C-terminal Constructs***

Multiple sequence alignment was done using MAFFT [37], and membrane topology of mouse PIEZO1 was predicted by TopCons [36]. Cloning was done using the In-Fusion cloning kit (Clontech Inc.) [38]. Briefly, vector DNA was linearized either by restriction digest method or polymerase chain reaction (PCR). Next, Insert DNA was amplified by PCR using the forward and reverse primers which contain the overhang sequence that are complementary to the DNA sequence of the linearized vector. Both insert and vector DNA were incubated together in the proprietary In-Fusion mix for 15 minutes at 55°C. The plasmid DNA mixture was transformed into Stellar E.coli strain (Clontech Inc.), and the sequence of the resulting plasmid DNA was verified by DNA sequencing.

### ***Protein Expression Using E. coli System***

Frozen cell stock was inoculated in a 0.5 mL starting culture one day prior to the induction in a 96-well plate (2 mL per well). 50 µL of the starter culture was transferred to 4 mL TB in a 24-well plate. Cultures were grown for 4 – 5 hours or until the OD<sub>600</sub> reached 1.0 to 1.5. For 37°C induction, 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added and the plate was shaken at 37°C for 2 hours. For room temperature induction, cells were first put on wet ice for 10 minutes to bring the temperature down before 1 mM IPTG was added. Upon IPTG addition, the plate was shaken at room temperature overnight. Cells were then harvested at the appropriate time by centrifugation and stored at -80°C after being flash-frozen using liquid nitrogen.

### ***Medium Scale NiNTA Purification***

Cells were resuspended in Tris lysis buffer (100 mM Tris-Cl pH 7.5, 25 mM NaCl, protease inhibitor, lysozyme, and DNase) in a ratio of 10 mL buffer per 1 g of cells. Cells were sonicated twice at 50% amplitude for 1 minute using ‘1 sec pulse – 1 sec rest’ setting (Misonix, S 4000). Then, sample was spun down using tabletop centrifuge at 14,000 rpm for 15 minutes in the cold room. Supernatant was collected and was spun down using ultracentrifuge at 150,000x g for one hour to collect membrane fraction. Membrane was solubilized using the appropriate 1% detergent for 1 hour. Membrane detergent suspension was spun down to remove insoluble materials. The resulting supernatant was injected into 5 mL NiNTA column (GE Healthcare, Inc.). Column was washed for 15 column volume with appropriate detergent-buffer containing 40 mM imidazole. Elution of this purification

was automatically collected and injected into a Superdex S-200 16/600 gel filtration column (GE Healthcare, Inc.).

The same type of detergent was used throughout extraction and purification procedure. In some purification, however, one detergent is replaced with another during the purification step. Briefly, membrane protein was extracted in 1% of detergent X. The protein extract was then incubated with the affinity beads that have been pre-equilibrated with detergent Y. Subsequent washes, elution, and gel filtration procedure were carried out in the presence of detergent Y.

#### ***Limited Tryptic Digest and Sample Preparation for the N-terminal Sequencing***

50 µg of Mouse PIEZO1-GST (160 pmole) was mixed with 10 pmole of trypsin in a 75 µL volume. Reaction was incubated at 4°C and 10 µL samples were taken periodically. Reaction was stopped by mixing the collected samples with the SDS loading buffer which contains 2% SDS and 5% DTT. After the last time point at 4°C, reaction was driven to completion by incubating the reaction mix at 37°C for 25 minutes. Each 10 µL samples were loaded into SDS PAGE gel. Limited proteolysis product was shown as both stain-free gel image and the western blot results.

For N-terminal peptide sequencing, samples were mixed with SDS running buffer and were ran on NuPAGE Novex 3 - 8% Tris-Acetate gel (ThermoFisher). Gel was subsequently blotted on Immobilon-P PVDF membrane (Millipore). The resulting blot was stained using 0.2% Ponceau S (in 3% acetic acid) for one minute before it was rinsed with

water until the red band was apparent. The protein band on the PVDF membrane was sent to the PPMAL facility at Caltech for the N-terminal peptide sequencing.

### ***Cell Fractionation***

The cell pellet was resuspended in PIPES lysis buffer (25 mM NaPIPES pH 7.2, 140 mM NaCl, 1.0 mM EGTA, and protease inhibitor) in ratio of 5 mL buffer per 1 g of mammalian cell pellet. Cells were gently lysed using Dounce homogenizer. Cells were spun down using tabletop centrifuge at 14,000 rpm for 15 minutes in the cold room. Take the supernatant as the ‘total’ fraction (Tot). The total fraction was then spun down using ultracentrifuge at 150,000x g for one hour. The resulting supernatant was collected as the cytoplasmic fraction (Cyt). The resulting pellet was resuspended in 1 mL PIPES lysis buffer per 1 g of original cell pellet mass, resulting in 5x concentration. Then, samples were mixed with the SDS loading buffer, and were loaded onto SDS PAGE. Finally, gels were blotted onto a nitrocellulose membrane for Western blot using anti-His antibody.

### ***Immunocytochemistry (ICC) Staining***

HEK293T cells were plated on coverslips coated with 0.1% gelatin. When cells reached 70% confluence, they were transfected with the appropriate DNA plasmid using Lipofectamine 3000 (ThermoFisher). Thirty-six hours after transfection, *live-cells ICC staining* was performed by incubating the transfected cells with either Myc 9E11 antibody (Santa Cruz Biotechnology, at 1:100 dilution) or HIS.H8 antibody (Santa Cruz Biotechnology, at 1:200 dilution) for one hour at 37°C. Cells were then washed five times with warm media and incubated with the secondary antibodies conjugated to Alexa Fluor

647 (Life Technologies, at 1:500 dilutions) for one hour at room temperature. Finally, cells were washed five times with PBS buffer and fixed with 2% PFA in PBS for 20 minutes.

For *ICC staining on permeabilized cells*, fixation procedure with 4% PFA in PBS was carried out for 10 minutes followed by five washes with PBS. Next, cells were permeabilized with 0.4% Triton X-100 for 10 minutes at room temperature. Then, cells were washed three times with PBS and blocked with 5% FBS in PBS for one hour at room temperature. Following fixation and permeabilization, ICC staining procedure was done by incubating the permeabilized cells with either Myc 9E11 antibody (at 1:100 dilution) or HIS.H8 antibody (at 1:200 dilution) in blocking buffer followed for one hour at room temperature. Cells were then washed three times with PBS and secondary antibodies conjugated to Alexa Fluor 647 dye (at 1:500 dilution) was applied for one hour at room temperature. Cells were then washed with PBS three times. Mounting media ProLong Gold (Life Technologies) was then used to mount the coverslips on slides.

Results were visualized by widefield fluorescence microscope provided by the Protein Expression Center at California Institute of Technology (Caltech). The ICC staining results were also visualized using confocal fluorescence microscope (Zeiss LSM 710 NLO confocal microscope and Plan-Apochromatic f1.4 63x objective) provided by the Biological Imaging Center at Caltech. Confocal images were obtained for every 0.5  $\mu\text{m}$  increment along the z-axis.

**Author Contribution**

A.K. designed the experiments, performed most of the experiments, analyzed the results, and wrote the manuscript. P.N. performed the immunocytochemistry staining and confocal imaging, as well as some detergent screen procedures and protein purification. J.V. offered useful advice on mammalian cell work and provided the mammalian cell culture facilities and widefield fluorescence microscope for localization imaging. D.C.R. provided mentorship and funding which made this project possible.

**References:**

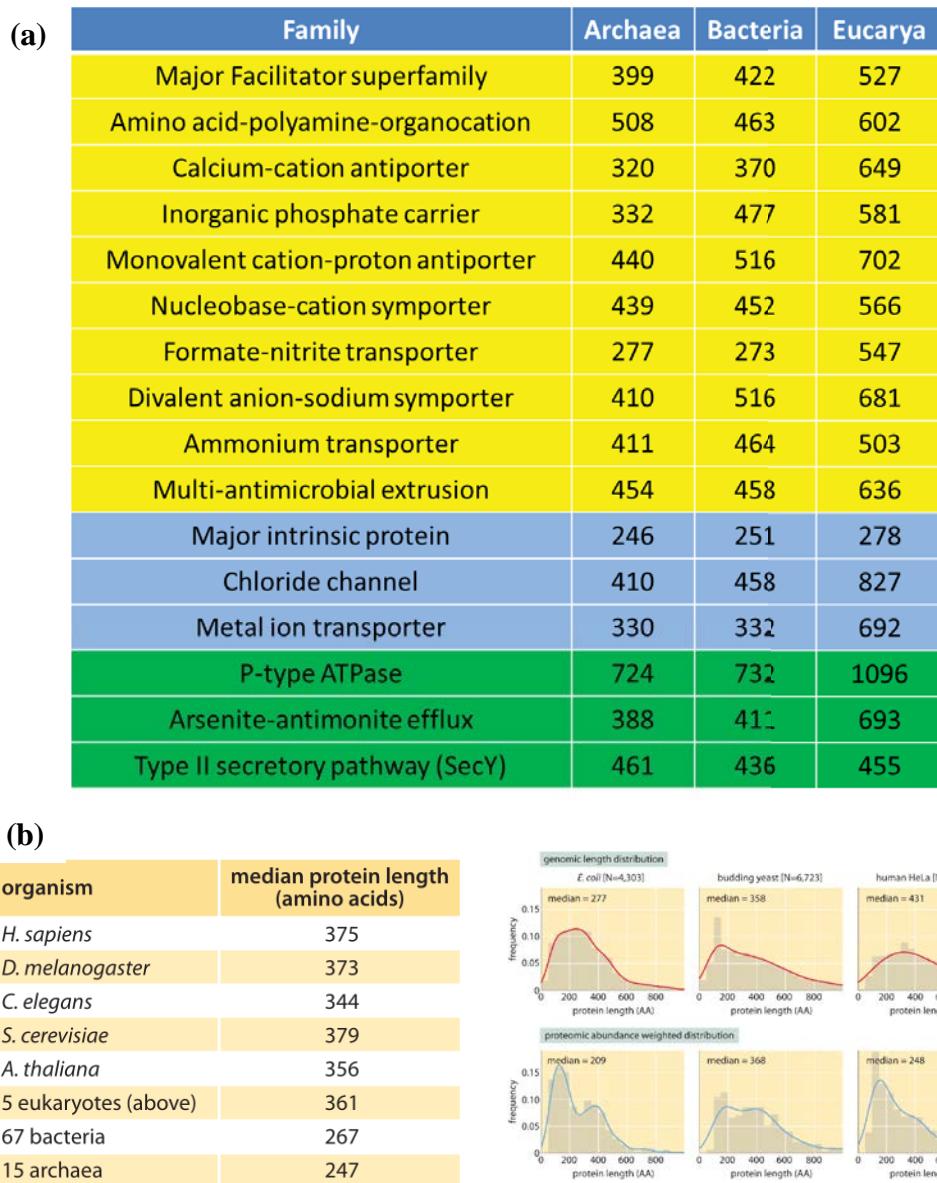
1. Jardetzky, O., *Simple allosteric model for membrane pumps.* Nature, 1966. **211**(5052): p. 969-70.
2. Locher, K.P., A.T. Lee, and D.C. Rees, *The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism.* Science, 2002. **296**(5570): p. 1091-8.
3. Borths, E.L., et al., *The structure of Escherichia coli BtuF and binding to its cognate ATP binding cassette transporter.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16642-7.
4. Hvorup, R.N., et al., *Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF.* Science, 2007. **317**(5843): p. 1387-90.
5. Korkhov, V.M., S.A. Mireku, and K.P. Locher, *Structure of AMP-PNP-bound vitamin B12 transporter BtuCD-F.* Nature, 2012. **490**(7420): p. 367-72.
6. Korkhov, V.M., et al., *Structure of AMP-PNP-bound BtuCD and mechanism of ATP-powered vitamin B12 transport by BtuCD-F.* Nat Struct Mol Biol, 2014. **21**(12): p. 1097-9.
7. Dong, Y.Y., et al., *K2P channel gating mechanisms revealed by structures of TREK-2 and a complex with Prozac.* Science, 2015. **347**(6227): p. 1256-9.
8. Brohawn, S.G., J. del Marmol, and R. MacKinnon, *Crystal structure of the human K2P TRAAK, a lipid- and mechano-sensitive K<sup>+</sup> ion channel.* Science, 2012. **335**(6067): p. 436-41.

9. Brohawn, S.G., E.B. Campbell, and R. MacKinnon, *Domain-swapped chain connectivity and gated membrane access in a Fab-mediated crystal of the human TRAAK K<sup>+</sup> channel*. Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2129-34.
10. Brohawn, S.G., E.B. Campbell, and R. MacKinnon, *Physical mechanism for gating and mechanosensitivity of the human TRAAK K<sup>+</sup> channel*. Nature, 2014. **516**(7529): p. 126-30.
11. Chang, G., et al., *Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel*. Science, 1998. **282**(5397): p. 2220-6.
12. Liu, Z., C.S. Gandhi, and D.C. Rees, *Structure of a tetrameric MscL in an expanded intermediate state*. Nature, 2009. **461**(7260): p. 120-4.
13. Li, J., et al., *Mechanical coupling of the multiple structural elements of the large-conductance mechanosensitive channel during expansion*. Proc Natl Acad Sci U S A, 2015. **112**(34): p. 10726-31.
14. Bass, R.B., et al., *Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel*. Science, 2002. **298**(5598): p. 1582-7.
15. Wang, W., et al., *The structure of an open form of an E. coli mechanosensitive channel at 3.45 Å resolution*. Science, 2008. **321**(5893): p. 1179-83.
16. Haswell, E.S., R. Phillips, and D.C. Rees, *Mechanosensitive channels: what can they do and how do they do it?* Structure, 2011. **19**(10): p. 1356-69.
17. Milo, R. and R. Phillips, *Cell Biology by the Numbers*, 2015, Garland Science. p. 400.

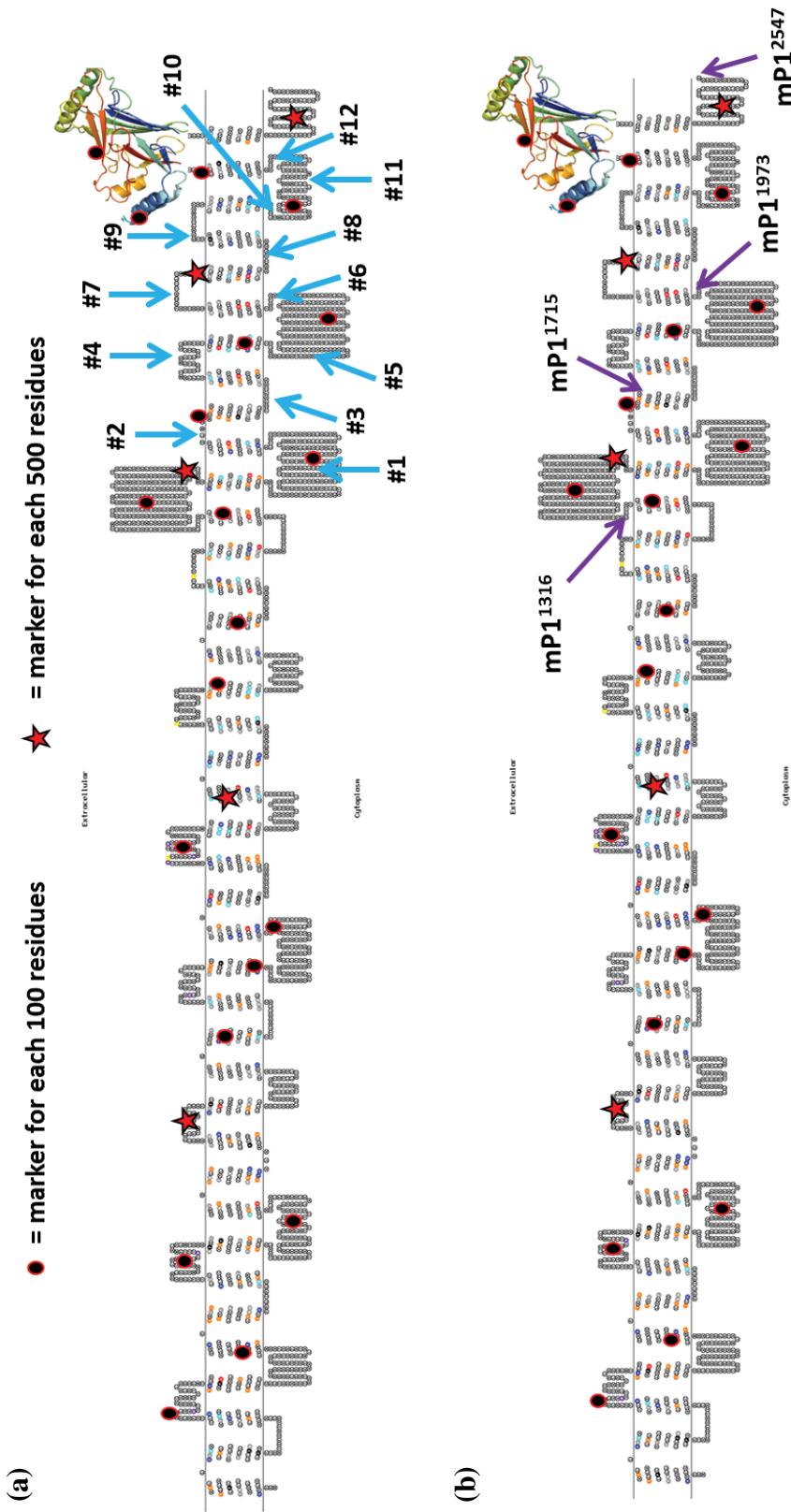
18. Chung, Y.J., et al., *Size comparisons among integral membrane transport protein homologues in bacteria, Archaea, and Eucarya*. J Bacteriol, 2001. **183**(3): p. 1012-21.
19. Coste, B., et al., *Piezol1 and Piezo2 are essential components of distinct mechanically activated cation channels*. Science, 2010. **330**(6000): p. 55-60.
20. Coste, B., et al., *Piezo proteins are pore-forming subunits of mechanically activated channels*. Nature, 2012. **483**(7388): p. 176-81.
21. Kamajaya, A., et al., *The structure of a conserved piezo channel domain reveals a topologically distinct beta sandwich fold*. Structure, 2014. **22**(10): p. 1520-7.
22. Bae, C., et al., *Xerocytosis is caused by mutations that alter the kinetics of the mechanosensitive channel PIEZO1*. Proc Natl Acad Sci U S A, 2013. **110**(12): p. E1162-8.
23. Kadaba, N.S., et al., *The high-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation*. Science, 2008. **321**(5886): p. 250-3.
24. Johnson, E., et al., *Inward facing conformations of the MetNI methionine ABC transporter: Implications for the mechanism of transinhibition*. Protein Sci, 2012. **21**(1): p. 84-96.
25. Kadner, R.J., *Regulation of methionine transport activity in *Escherichia coli**. J Bacteriol, 1975. **122**(1): p. 110-9.
26. Ramsey, I.S., M. Delling, and D.E. Clapham, *An introduction to TRP channels*. Annu Rev Physiol, 2006. **68**: p. 619-47.
27. Montell, C., *The TRP superfamily of cation channels*. Sci STKE, 2005. **2005**(272): p. re3.

28. Coste, B., et al., *Piezol ion channel pore properties are dictated by C-terminal region.* Nat Commun, 2015. **6**.
29. Ge, J., et al., *Architecture of the mammalian mechanosensitive Piezol channel.* Nature, 2015.
30. Zhao, Q., et al., *Ion Permeation and Mechanotransduction Mechanisms of Mechanosensitive Piezo Channels.* Neuron, 2016. **89**(6): p. 1248-63.
31. Lewinson, O., A.T. Lee, and D.C. Rees, *The funnel approach to the precrystallization production of membrane proteins.* J Mol Biol, 2008. **377**(1): p. 62-73.
32. Chaudhary, S., et al., *Efficient expression screening of human membrane proteins in transiently transfected Human Embryonic Kidney 293S cells.* Methods, 2011. **55**(4): p. 273-80.
33. Lebon, G., et al., *Thermostabilisation of an agonist-bound conformation of the human adenosine A(2A) receptor.* J Mol Biol, 2011. **409**(3): p. 298-310.
34. Kefala, G., et al., *Structures of the OmpF porin crystallized in the presence of foscholine-12.* Protein Sci, 2010. **19**(5): p. 1117-25.
35. Lai, J.Y., et al., *Open and shut: crystal structures of the dodecylmaltoside solubilized mechanosensitive channel of small conductance from Escherichia coli and Helicobacter pylori at 4.4 Å and 4.1 Å resolutions.* Protein Sci, 2013. **22**(4): p. 502-9.
36. Bernsel, A., et al., *TOPCONS: consensus prediction of membrane protein topology.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W465-8.

37. Katoh, K., et al., *MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.* Nucleic Acids Res, 2002. **30**(14): p. 3059-66.
38. Sleight, S.C., et al., *In-Fusion BioBrick assembly and re-engineering.* Nucleic Acids Res, 2010. **38**(8): p. 2624-36.



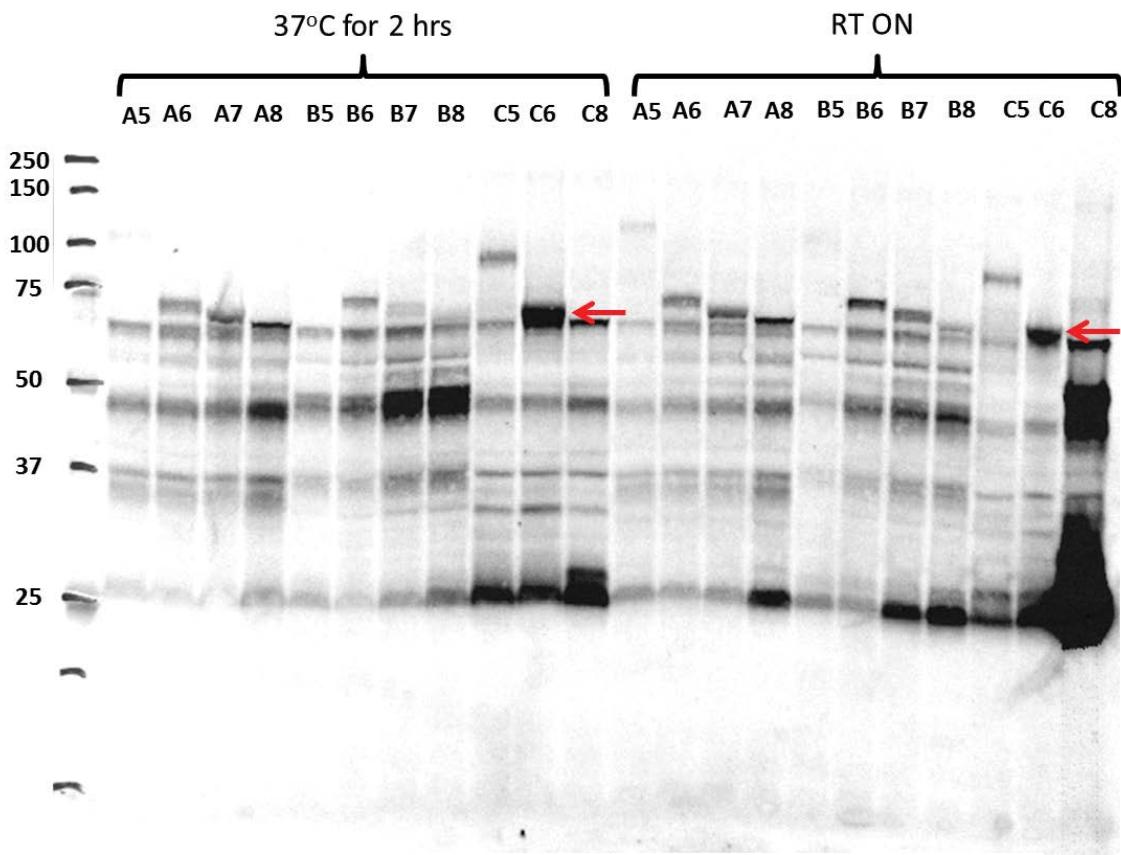
**Figure 1. Average number of amino acids in known genomes across three domains of life.** (a) Average number of amino acids composition for major classes of transporters and ion channels. Yellow shading represents carriers, blue shading represents channels, and green shading represents primary active transporters. This table is adapted from [18]. (b) Distribution of protein length in term of the number of amino acids from representative organisms from all three domains of life. Images are directly taken from [17], and are used under the permission of Rob Phillips and Ron Milo.



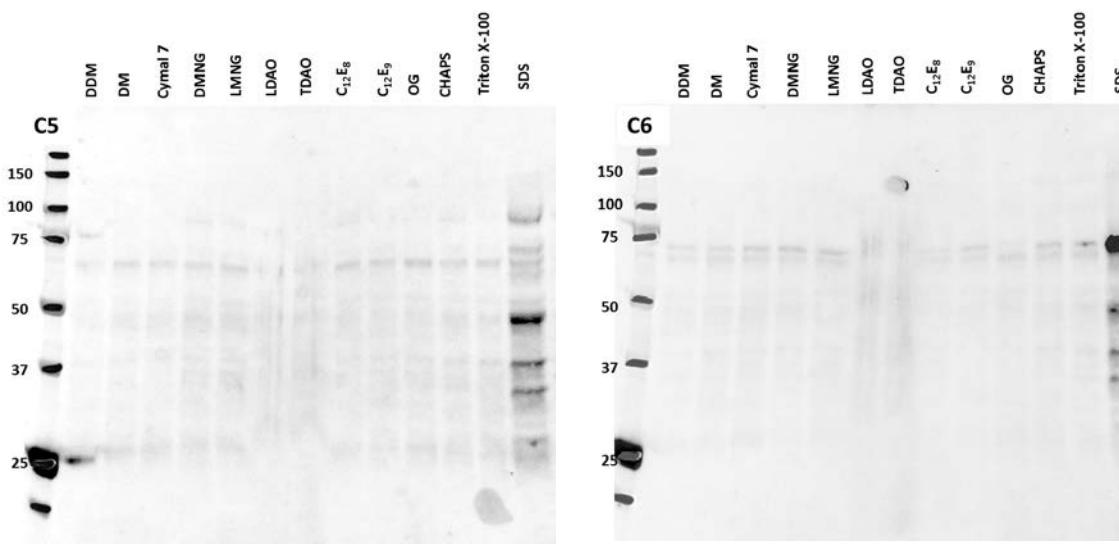
**Figure 2. Diagram of mouse PIEZO1.** Membrane topology prediction was created by Topcons [36], and approximate location of residues are marked every 100 (circle) and 500 (star) amino acids. (a) Approximate starting position for each of the twelve C-terminal Piezo constructs are indicated by the blue arrows. (b) Approximate location of the residues mentioned in the chimaera study [28] is indicated by the purple arrows.

	Construct #1	Construct #2	Construct #3	Construct #4	Construct #5	Construct #6	Construct #7	Construct #8	Construct #9	Construct #10	Construct #11	Construct #12
	# aa	MW (kDa)	# aa	MW (kDa)								
Human PIEZO1	862	99	825	94	796	91	746	85	718	82	572	65
											527	60
									494	57	462	53
Mouse PIEZO1	895	103	852	98	823	95	770	88	736	84	581	66
											538	62
									507	58	475	54
Mouse PIEZO2	849	97	821	94	792	91	745	85	715	81	568	65
											528	60
									496	57	466	53
Drosophila PIEZO	846	97	814	93	785	90	744	85	713	81	620	71
											580	66
									549	63	522	60
C.elegans PIEZO	858	97	821	92	793	90	750	85	712	80	639	72
											581	66
									553	63	516	59
									477	54	450	51
									415	47		

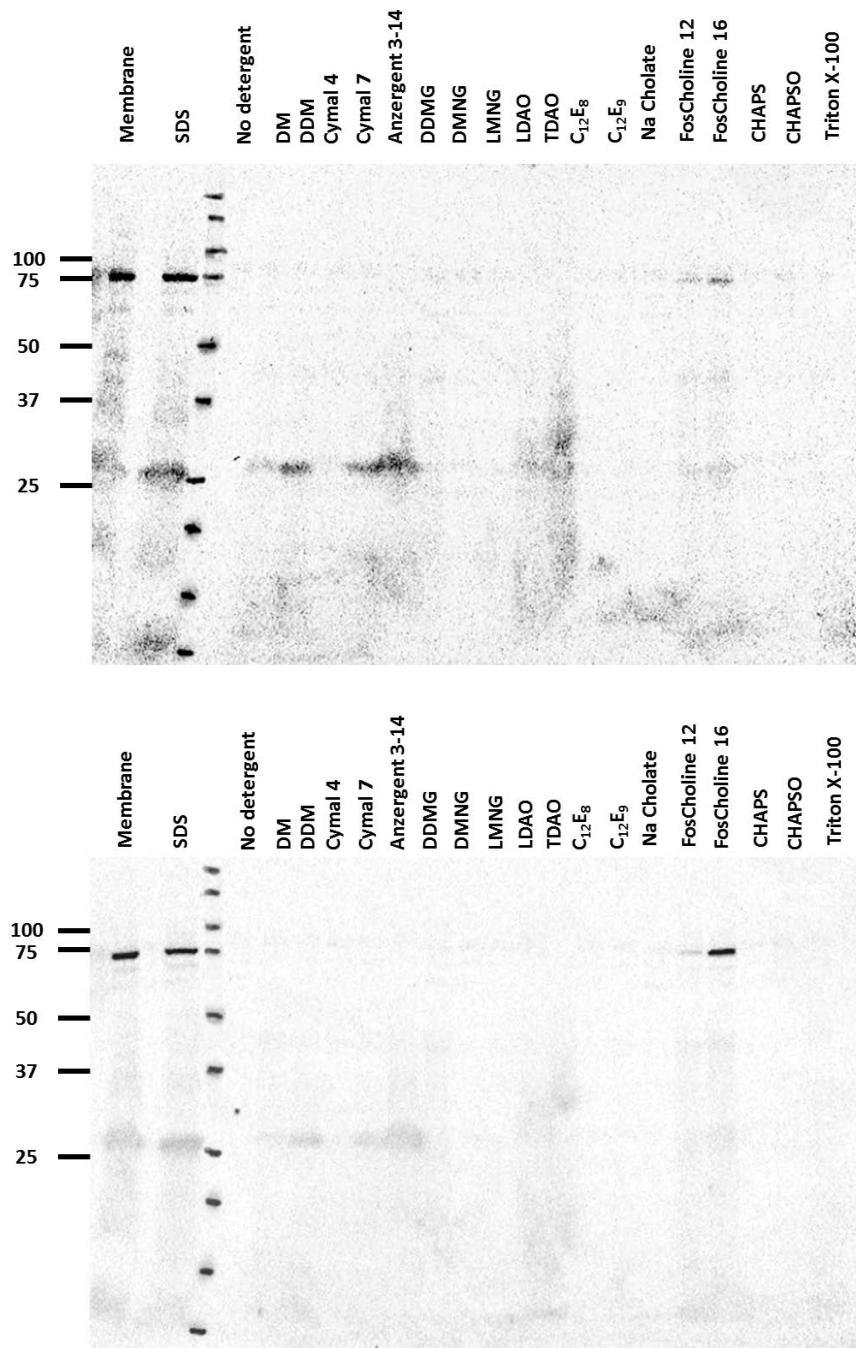
**Table 1. Summary of the number of amino acids and size for the C-terminal constructs.** Size is listed in the number of amino acids (#aa) and the molecular weight (MW) in kilo Dalton (kDa). C-terminal constructs that were chosen for initial investigation are boxed in red.



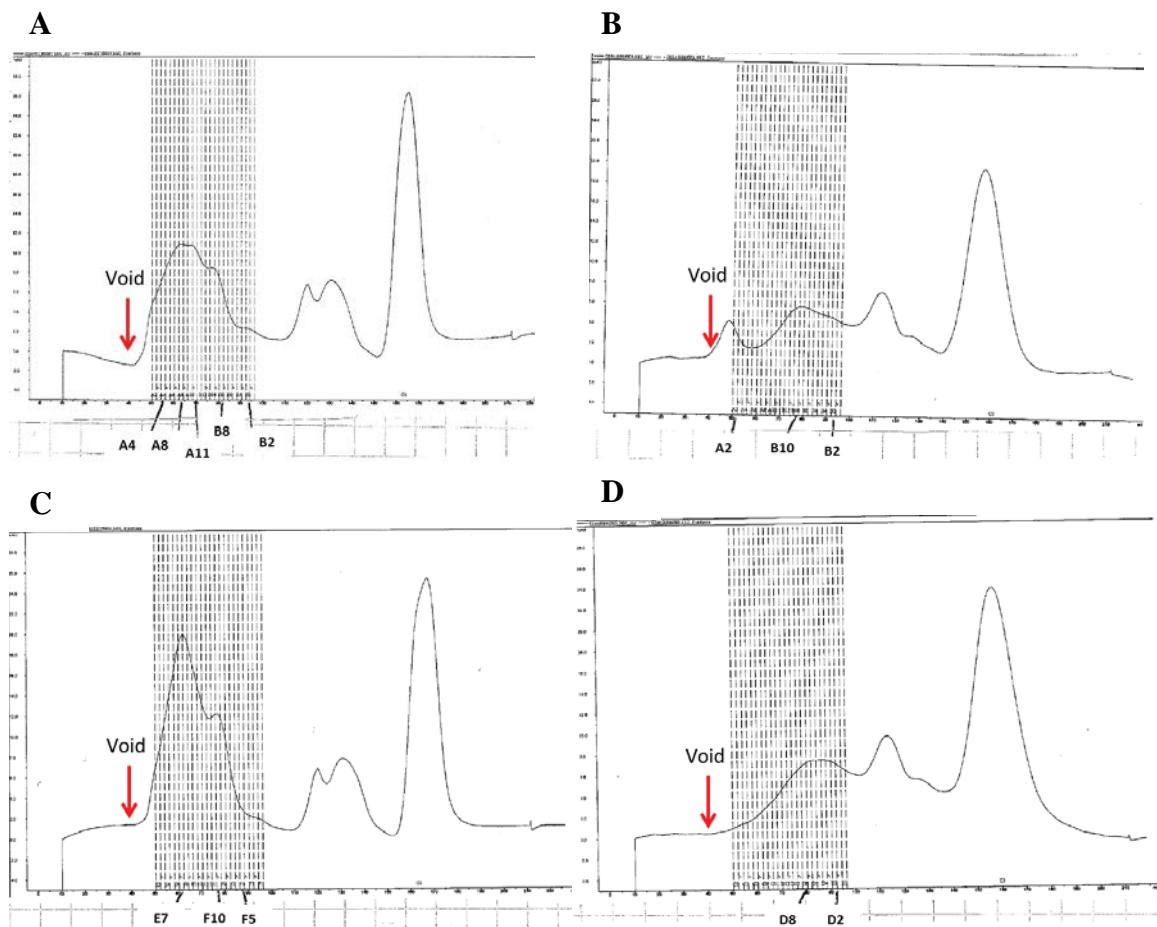
**Figure 3. Representative Western blot results for the initial test expression of several C-terminal Piezo constructs.** Test expression was done using *E. coli* expression system at small scale (4 mL culture) at two different expression conditions: 37°C for two hours or room temperature for overnight (RT ON). The alphabet represents different Piezo homologs: sample A is human PIEZO1, sample B is mouse PIEZO1, and sample C is mouse PIEZO2. The number represents different truncation sites as labelled on Table 1. Thus, lane A5 corresponds to construct #5 of human PIEZO1, whereas lane C8 corresponds to construct #8 of mouse PIEZO2. Here, C6 shows the strongest signal at the appropriate location. Due to technical error, C7 was not included on this blot. On a separate session of test expression, however, expression level of C7 is comparable to C6 (data not shown).



**Figure 4. Representative Western blot result from SDS PAGE of a detergent screen using several mild detergents.** Here, two C-terminal Piezo constructs from mouse PIEZO2 were extracted using twelve detergents that are relatively mild. Overall, none of these detergents are able to extract the C-terminal Piezo construct efficiently. Expected molecular weights for C5 (left) and C6 (right) are 81 kDa and 65 kDa, respectively. C5 expression level is much lower than C6.



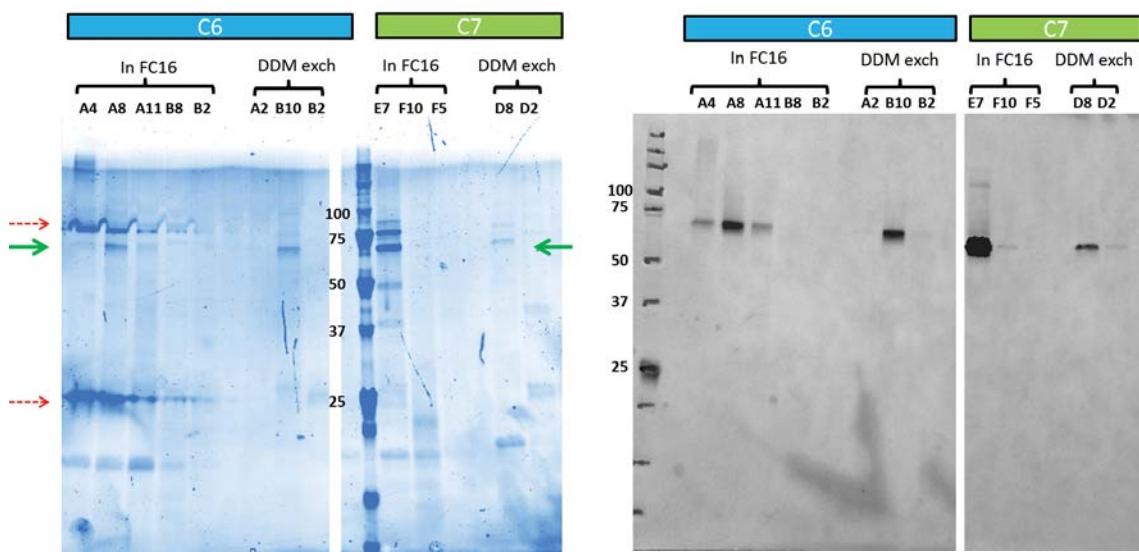
**Figure 5. Representative Western blot results of a detergent screen which include more detergents.** Results show that Fos-Choline-16 is the most optimum detergent for both C6 (top) and C7 (bottom). In fact, Fos-Choline-16 shows comparable protein extraction to SDS for C7. Expected molecular weights for C6 and C7 are 65 kDa and 60 kDa, respectively.



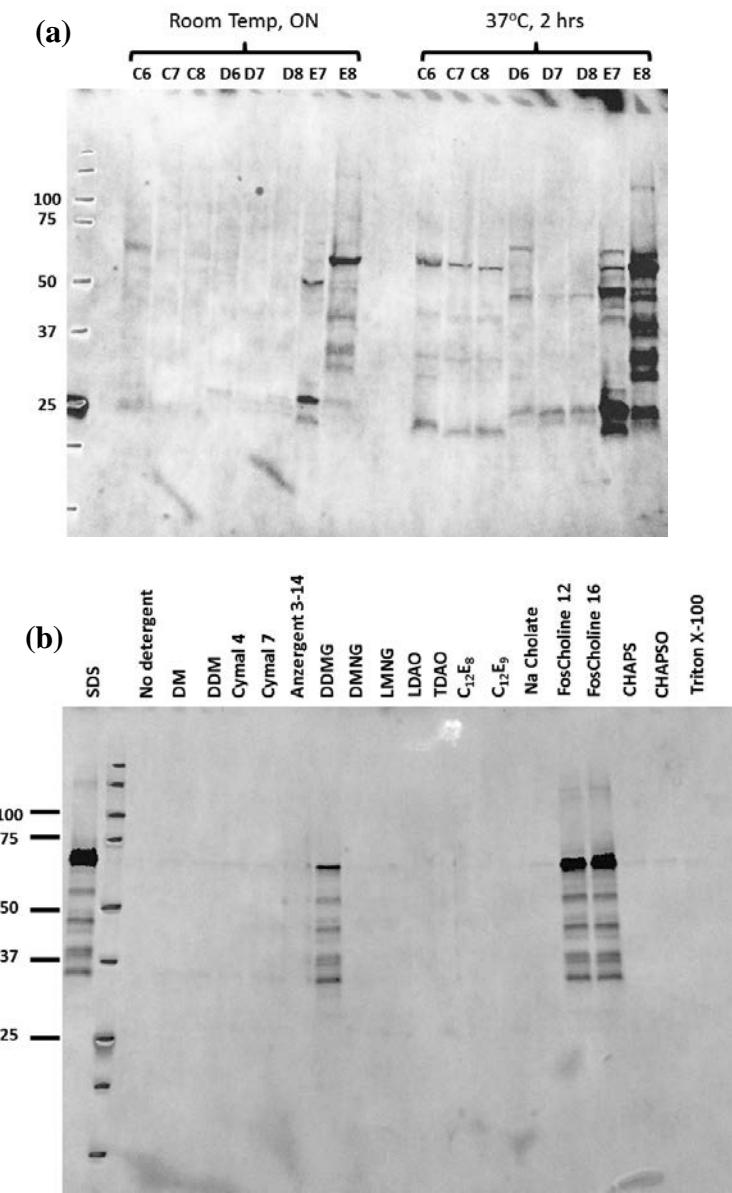
	Construct	Extraction	Purification
A	C6	Fos Choline 16	Fos Choline 16
B	C6	Fos Choline 16	DDM
C	C7	Fos Choline 16	Fos Choline 16
D	C7	Fos Choline 16	DDM

**Figure 6. Gel filtration profile of two C-terminal Piezo construct: C6 and C7.**

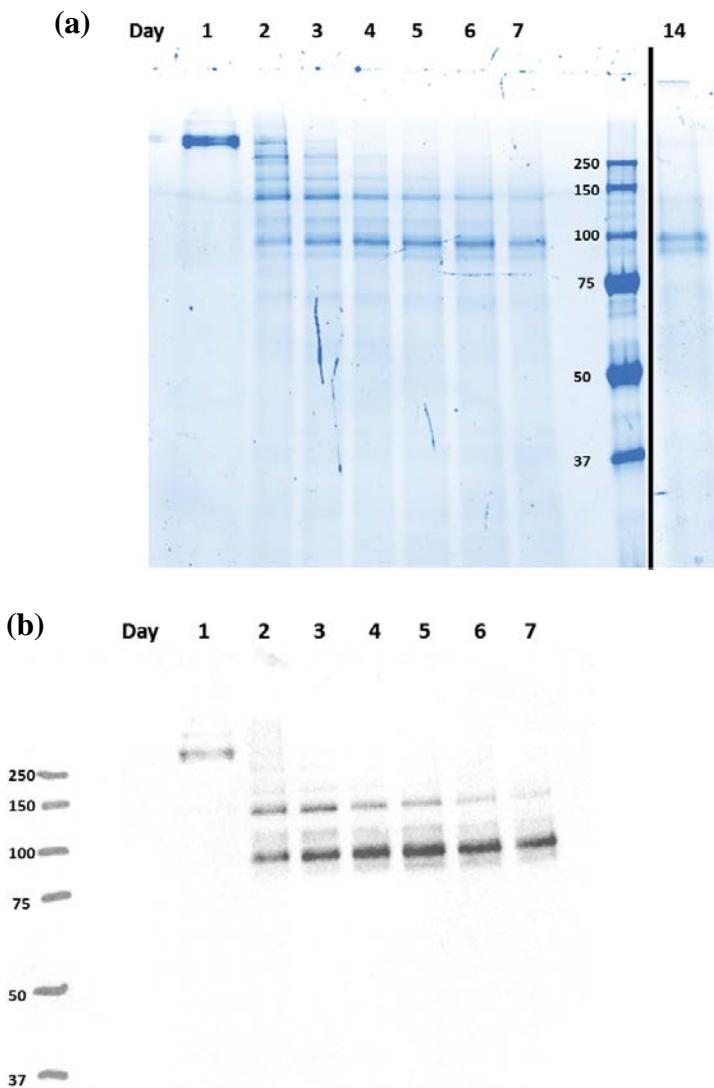
Tandem purification was carried out such that the elution from the NiNTA column is directly injected into Superdex S200 16/600 gel filtration column; imidazole peak is apparent as a large peak at the very end of the chromatogram.



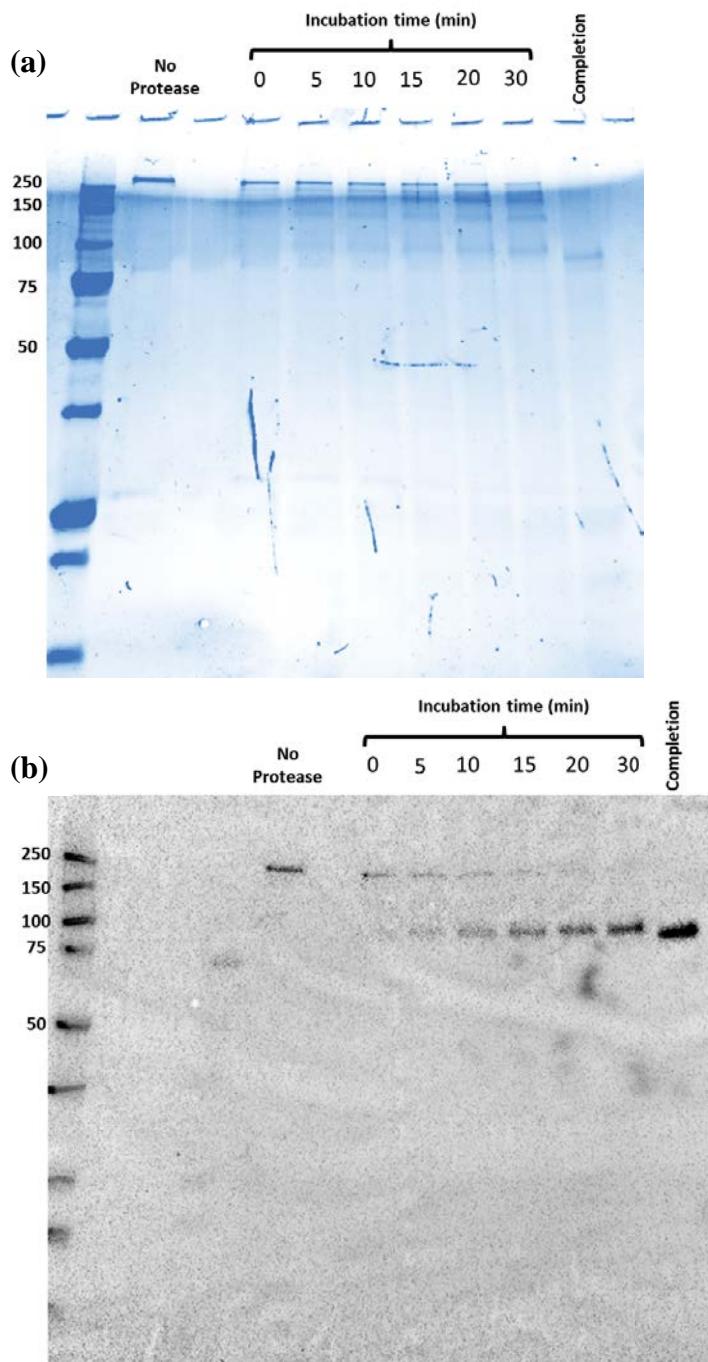
**Figure 7 Purification results of C6 and C7 visualized in a ‘stain-free’ gel image (a) and a Western blot (b).** Overall, purification was not optimized as apparent from the presence of low molecular weight contaminant bands. C-terminal Piezo construct band is indicated by green solid arrow. The 25 kDa and 75 kDa molecular weight markers for the Western blot ladder (b) showed strong signals when visualized using ‘stain-free’ setting on the BioRad gel imager. Given that the stain-free gel image (a) is the same gel used for the Western blot (b), it is very likely that the two bands (red dotted arrows) represent the overflow signal from the 25 kDa and 75 kDa molecular weight markers.



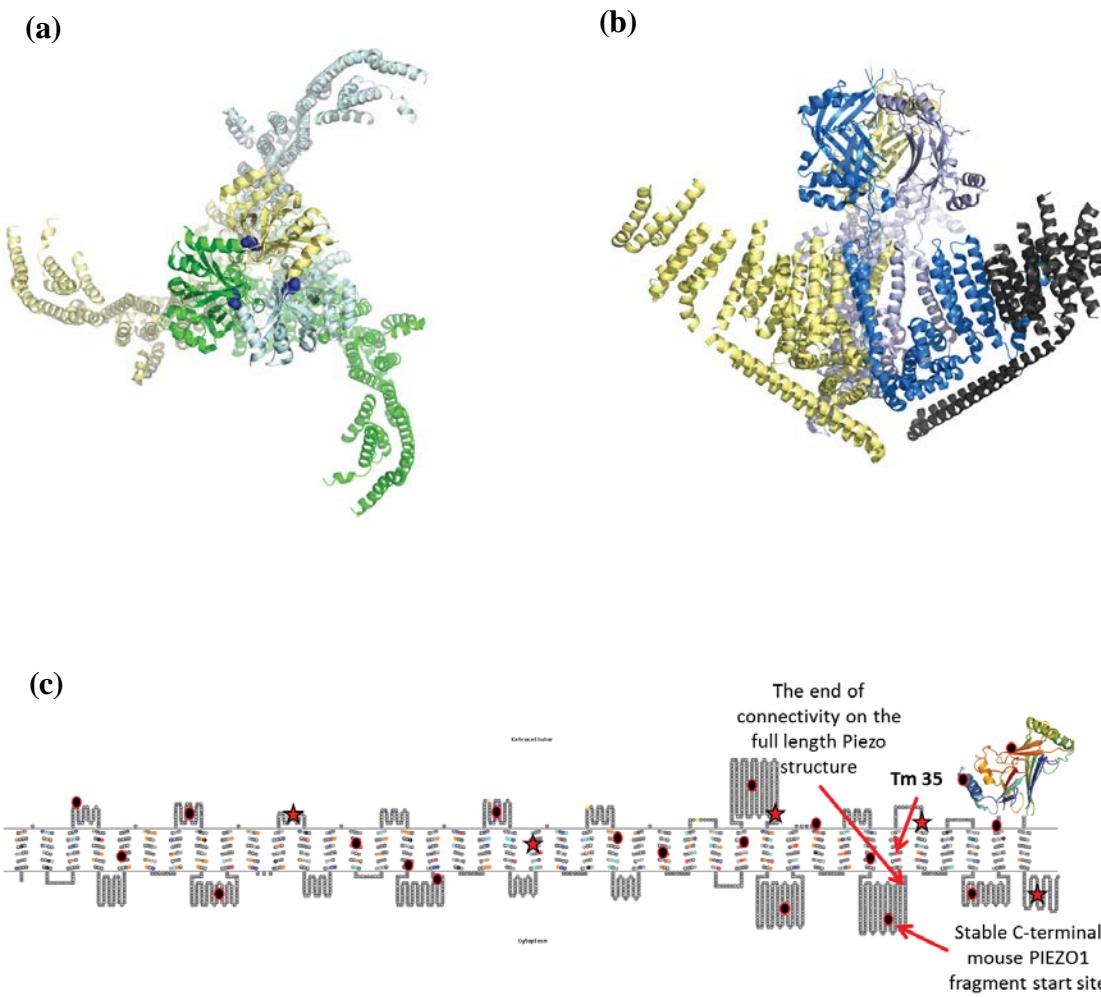
**Figure 8. Western blot result of the test expression and detergent screen.** (a) The level of protein expression of the C-terminal constructs from *Drosophila* PIEZO (D6, D7, and D8) and *C.elegans* PIEZO (E7 and E8) are compared to the expression level of C-terminal constructs from mouse PIEZO2 (C6, C7, and C8). Here, it appears that C-terminal construct #8 from the *C. elegans* PIEZO (E8) showed the highest protein expression level. (b) Detergent screen analysis shows that Fos-Choline-16 and Fos-Choline-12 (as well as DDMG to a lesser extend) can extract this construct from *E. coli* membrane efficiently.



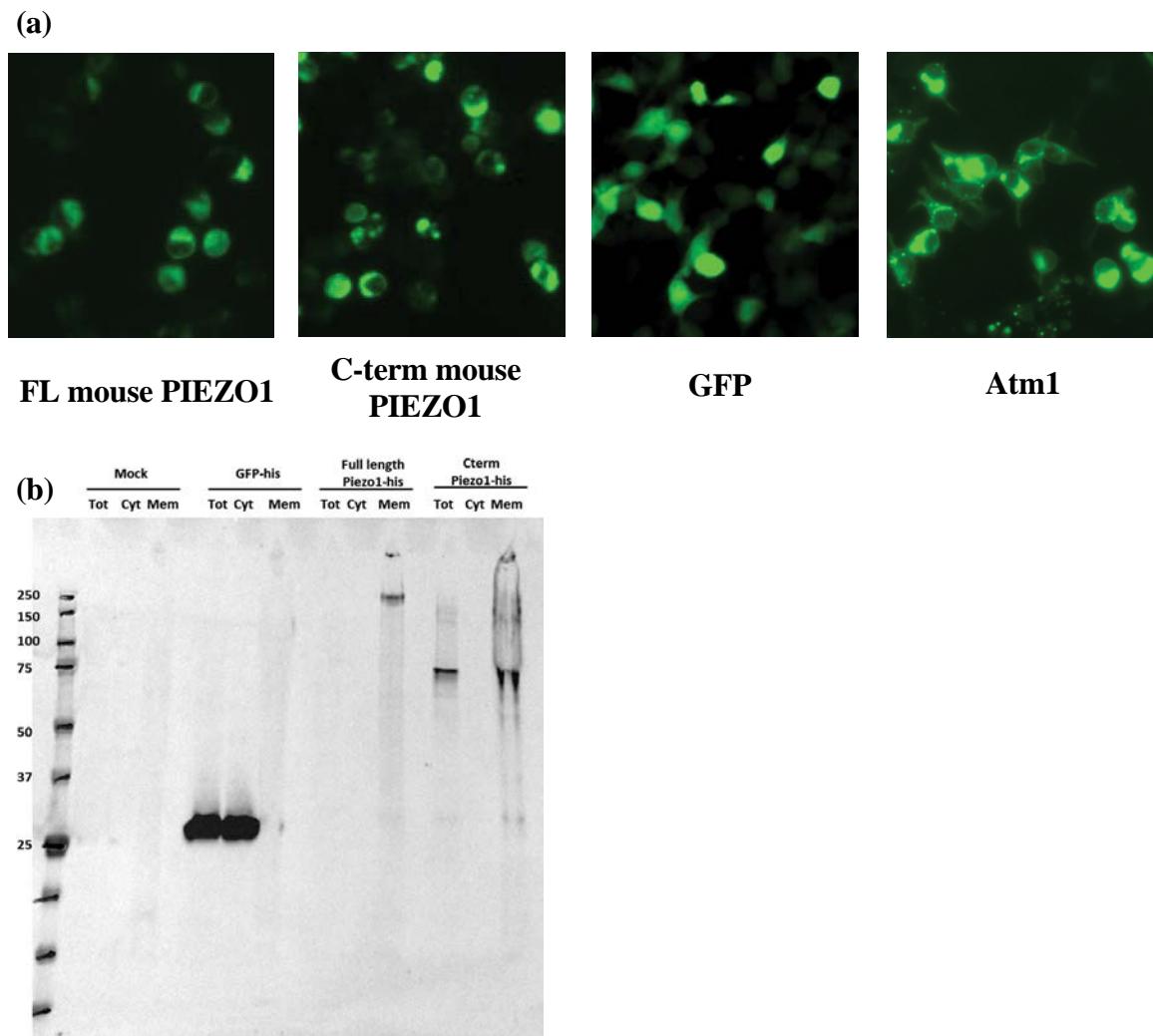
**Figure 9. Natural degradation of the full length mouse PIEZO1-GST resulted in a stable C-terminal Piezo fragment.** (a) Purified full length mouse PIEZO1-GST was degraded by day 2. The degradation was followed for seven days, and results from SDS PAGE was displayed as a Stain-free gel image (top). A separate gel was run for the same sample after 14 days (separated by black line). (b) This stable degradation product is the C-terminal fragment of Piezo as shown on the Western blot image; Western blot was done using anti-GST antibody, which recognizes the GST tag that is attached to the C-terminal end of the protein.



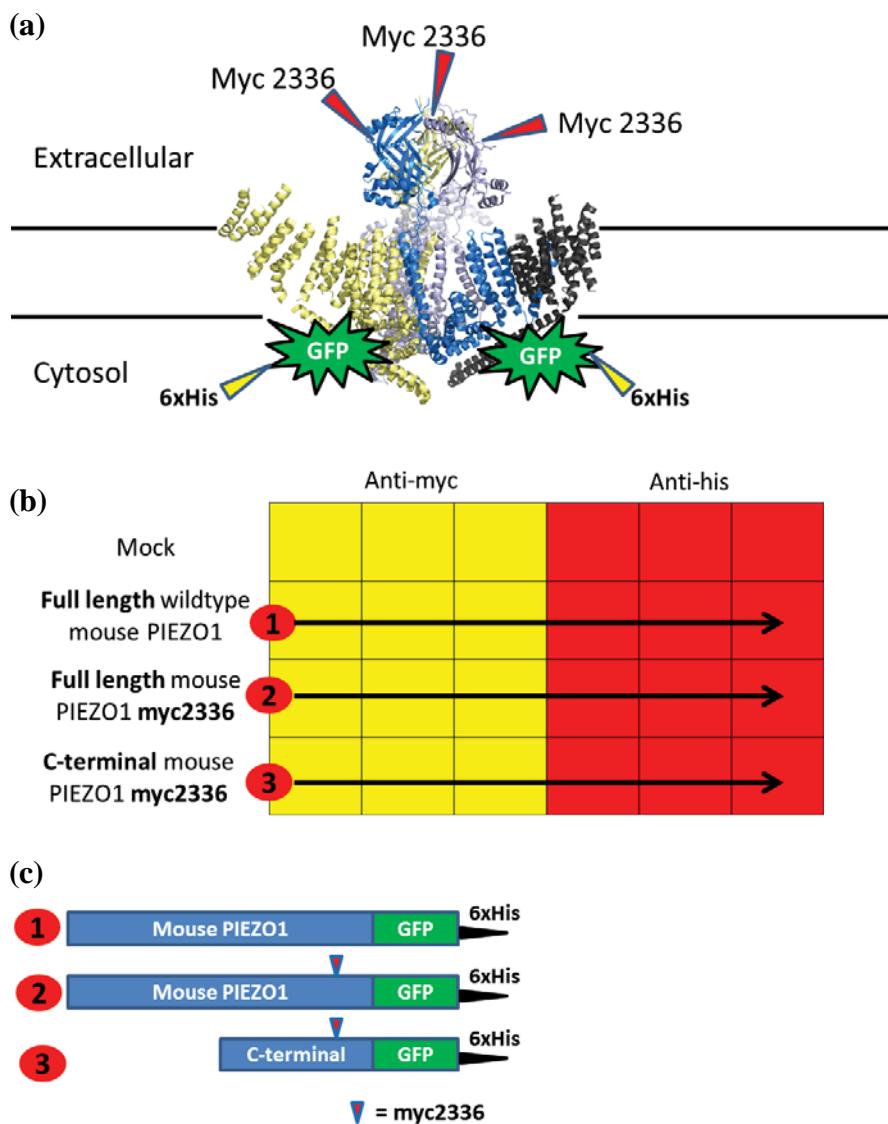
**Figure 10. Limited proteolysis of mouse PIEZO1-GST.** (a) Stain-free image of the limited tryptic digest of mouse PIEZO1 done at 4°C. Sample was taken at several time points. After 30 minutes digestion at 4°C, reaction was driven to completion by 25 minutes incubation at 37°C. (b) Western blot result shows that the resulting band is the C-terminal fragment of mouse PIEZO1.



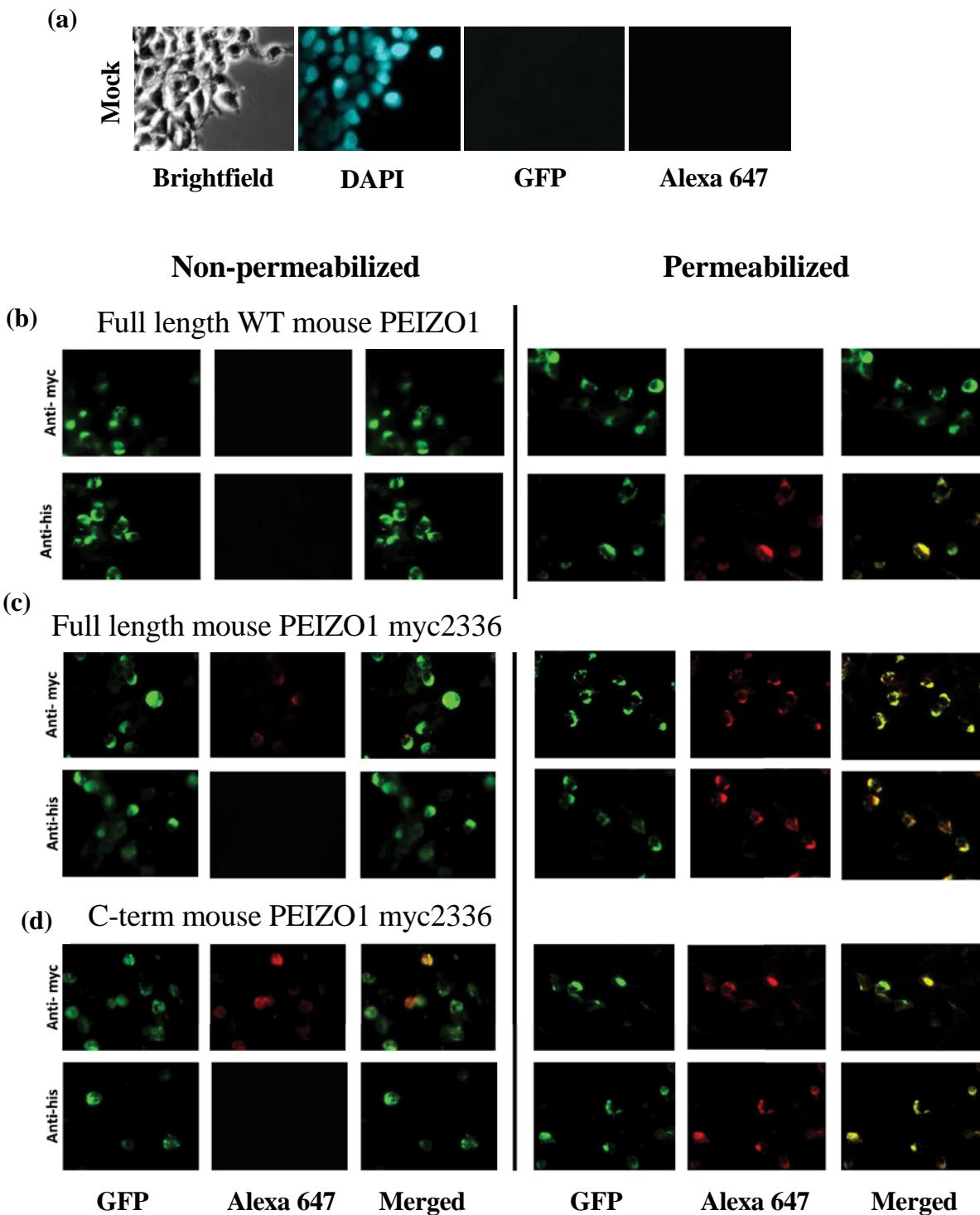
**Figure 11. Full length mouse PIEZO1 model.** (a) Top view of the trimeric mouse PIEZO1 [29], showing the three-fold symmetry centered at its ion permeation pathway. This ion permeation pathway makes up the central core of Piezo channel. (b) Side view of mouse PIEZO1, showing each monomer in a different color (yellow, light gray, and blue). Part of the blue monomer is colored black to indicate the helices in the model where connection cannot be traced back to the central core of the channel. (c) Membrane topology prediction of mouse PIEZO1 showing the approximate location of the starting residue of the stable C-terminal fragment of mouse PIEZO1 relative to the full length mouse PIEZO1 model.



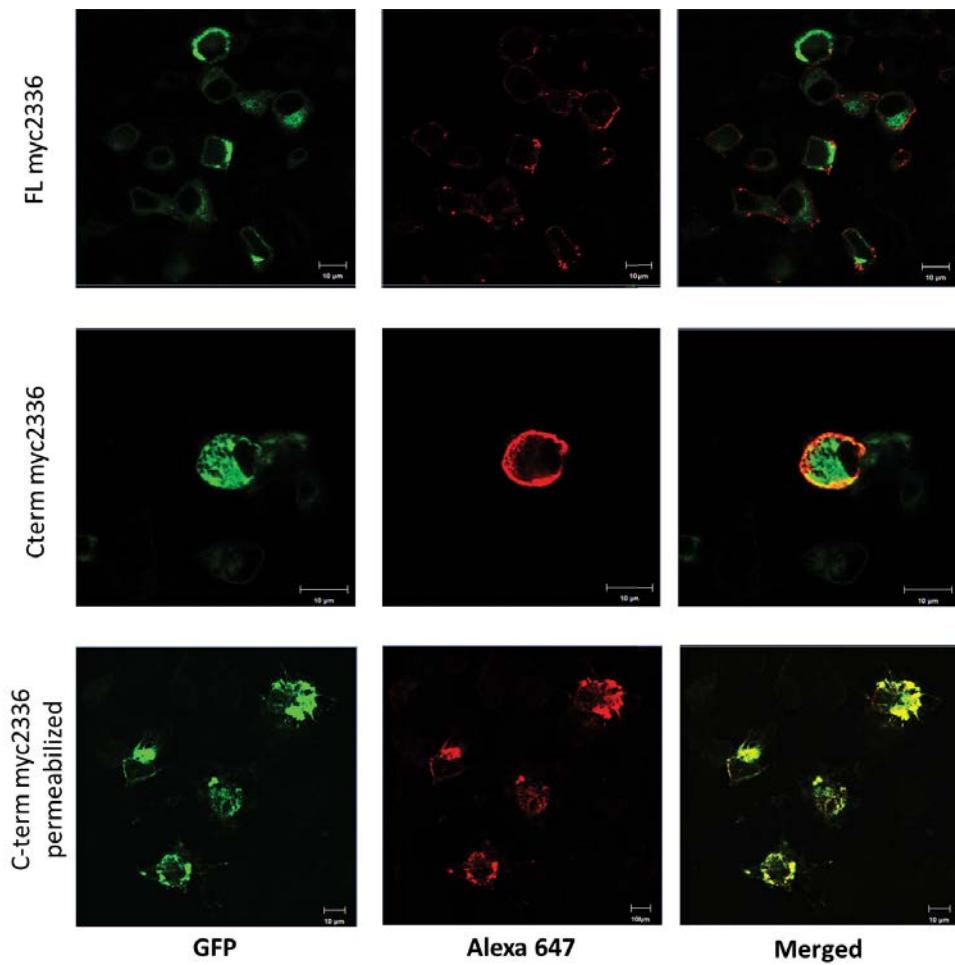
**Figure 12 C-terminal mouse PIEZO1 localization to the membrane.** (a) GFP fluorescence images of several GFP fusion constructs. Localization of the C-terminal mouse PIEZO1 resembles the localization pattern of the full length mouse PIEZO; this localization pattern is also observed in another membrane protein construct, an ABC transporter Atm1. Soluble protein, such as free GFP alone, shows uniform green fluorescence throughout cellular space. (b) Western blot result of cellular fractionation. Both full length mouse PIEZO1 and C-terminal mouse PIEZO1 can be found in the membrane (mem) fraction but not the cytosolic (cyt) fraction. Conversely, GFP can only be found in the cytosolic fraction, but not the membrane fraction.



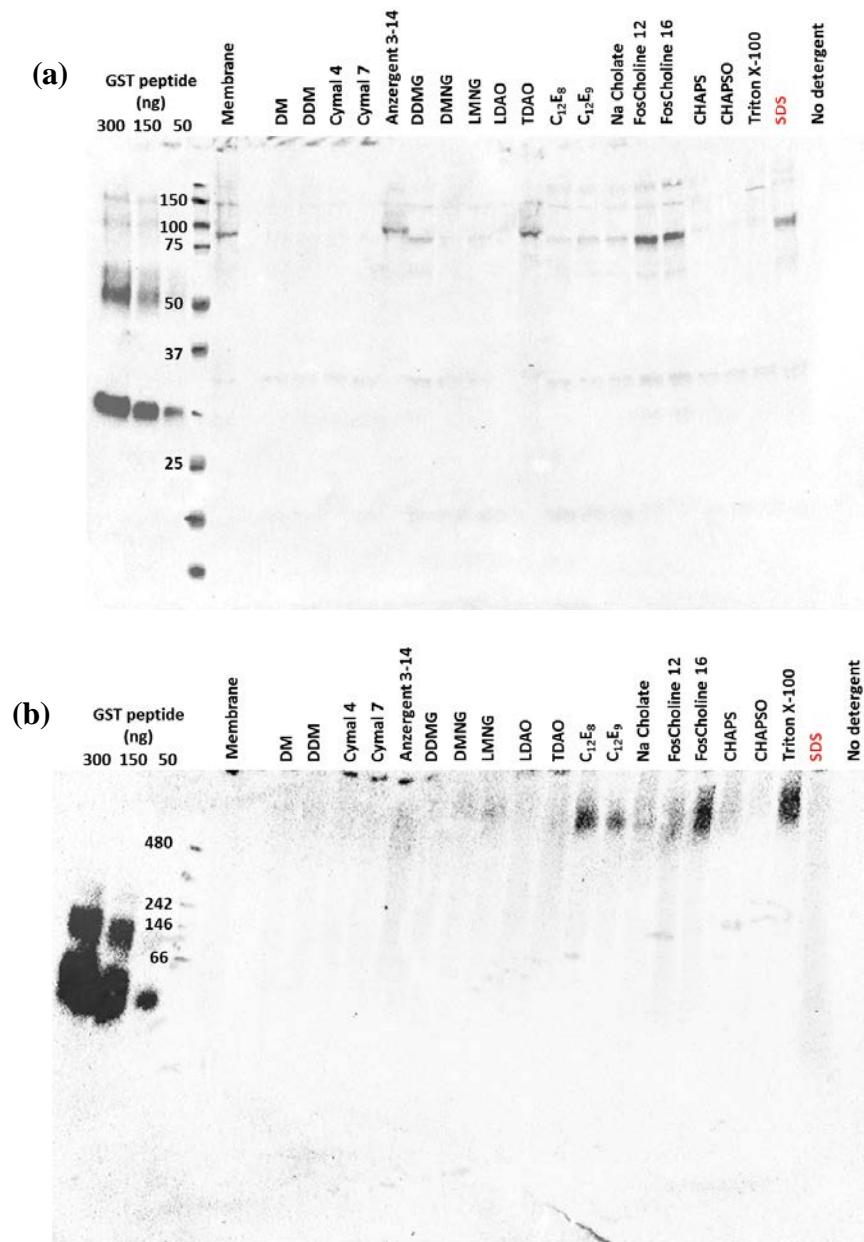
**Figure 13. Experimental design for the cellular localization of the C-terminal Piezo fragment.** (a) Schematic model showing the approximate location of the Myc 2336 tag and 6xHis tag on the full length mouse PIEZO1 model. (b) Arrangement of each construct in a 24-well plate. Each construct is transfected into 6 wells: 3 wells for anti-Myc staining and 3 wells for anti-His staining. One complete set of Immunocytochemistry (ICC) staining require two plates: one plate is for the 'live-cells' ICC staining and another plate is for ICC staining after permeabilization with 0.4% Triton X-100. (c) Diagram of the three constructs that were used for the ICC staining.



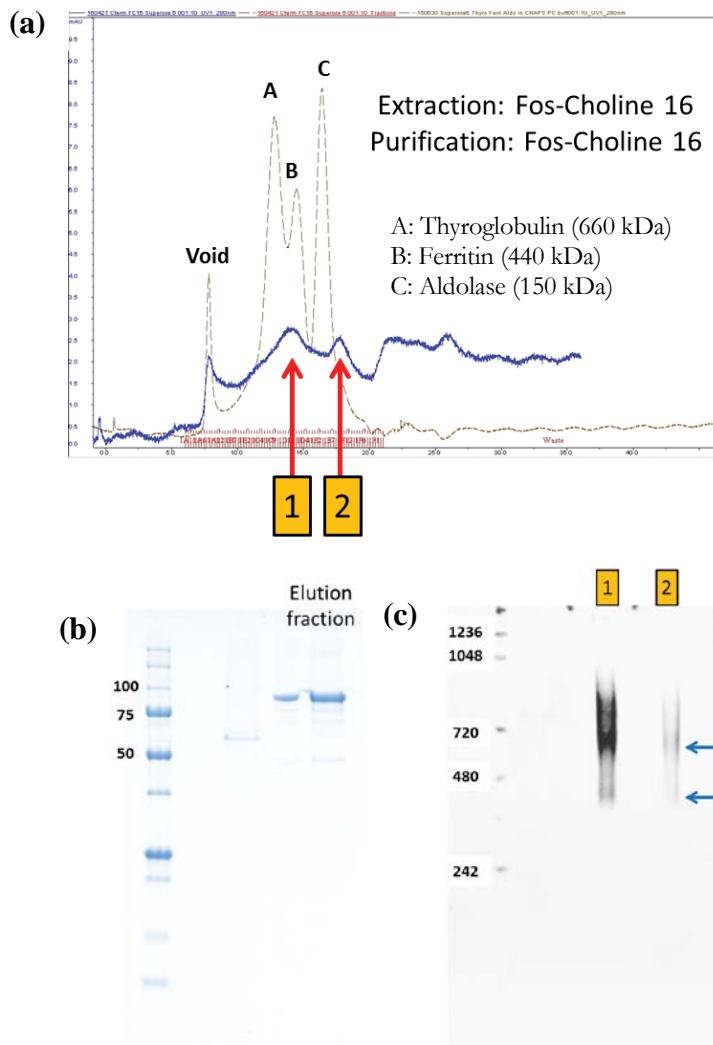
**Figure 14. Immunocytochemistry (ICC) staining results on a wide field fluorescence microscope.** (a) Mock transfected cells shows no fluorescence signal for both GFP and Alexa 647. (b,c,d) Results for wildtype full length mouse PIEZO1 (no Myc), full length mouse PIEZO1 Myc 2336, and C-terminal mouse PIEZO1 Myc 2336. Each construct is stained either with anti-Myc or anti-His antibodies in a ‘live-cells’ staining (non-permeabilized) or Triton X-100 permeabilized ICC staining set up. All constructs are expected to show GFP signals. However, Alexa 647 (red) signal can be detected only when the antibody molecules gain access to their corresponding epitope.



**Figure 15. Immunocytochemistry (ICC) staining results on confocal microscope.** Green signal indicates GFP localization, whereas red signal is the resulting signal of the ICC staining using anti-Myc antibody. Top and middle panels show samples from the live-cells staining procedure (non-permeabilized). Examples of cells which shows Alexa 647 (red) signal on the plasma membrane can be found in both full length mouse PIEZO1 Myc 2336 (FL Myc2336) and C-terminal mouse PIEZO1 Myc 2336 (Cterm Myc2336). These images shows one slice through the z-stack image sets which best illustrate the localization of the red signal surrounding cell membrane. Bottom panels show the samples from ICC staining using anti-Myc antibody after permeabilization with 0.4% Triton X-100. Here, red signal can be found at the same location where GFP signal is found in every image stack.



**Figure 16. Detergent screen of the C-terminal fragment of mouse PIEZO1.** (a) Western blot result from a SDS PAGE of detergent screen of the C-terminal mouse PIEZO1 construct. Fos-Choline-16 seems to be able to extract the C-terminal mouse PIEZO1 construct efficiently. (b) Western blot result from a BN-PAGE. C-terminal construct extracted in Fos-Choline-16 detergent-buffer condition shows a localized band, which may be an indication that the extracted protein is not completely unfolded.



**Figure 17. Preliminary results of protein purification of C-terminal mouse PIEZO1 construct in Fos-Choline-16 detergent-buffer condition.** (a) Purified protein was analyzed on a Superose 6 10/300 gel filtration column. Molecular weight standard is displayed in a dotted brown line: peak A is Thyroglobulin (660 kDa), peak B is Ferritin (440 kDa), and peak C is Aldolase (150 kDa). (b) Purified C-terminal mouse PIEZO1 on a SDS PAGE. Sample was taken prior to the gel filtration analysis. (c) Two samples which correspond to the two peaks on the gel filtration chromatogram are analyzed by western blot on a BN-PAGE (blue arrows). Both peaks show one major band between 720 kDa and 480 kDa molecular weight marker and one minor band just below the 480 kDa molecular weight marker.

## BIBLIOGRAPHY

- Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.
- Jones, Mary. *Life and Visual Perception*. City: University Press, 1998.
- Smith, Chris. *Theory and the Art of Communications Design*. State of the University Press, 1997.
- Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.
- Jones, Mary. *Life and Visual Perception*. City: University Press, 1998.
- Smith, Chris. *Theory and the Art of Communications Design*. State of the University Press, 1997.
- Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.
- Jones, Mary. *Life and Visual Perception*. City: University Press, 1998.
- Smith, Chris. *Theory and the Art of Communications Design*. State of the University Press, 1997.
- Smith, Chris. *Theory and the Art of Communications Design*. State of the University Press, 1997.
- Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.
- Jones, Mary. *Life and Visual Perception*. City: University Press, 1998.
- Smith, Chris. *Theory and the Art of Communications Design*. State of the University Press, 1997.
- Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.
- Jones, Mary. *Life and Visual Perception*. City: University Press, 1998.
- Smith, Chris. *Theory and the Art of Communications Design*. State of the University Press, 1997.
- Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.
- Jones, Mary. *Life and Visual Perception*. City: University Press, 1998.