PURIFICATION AND CRYSTALLIZATION OF THE PANNEXIN 1 CHANNEL

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Julia Marie Kumpf

January 2015

ABSTRACT

Autocrine and paracrine signaling via extracellular nucleotides are one of the most ubiquitous signaling mechanisms in the mammalian organism. Cells release nucleotides both by exocytosis and through large pores in the plasma membrane. Pannexins are non-selective channels that open a 10nm pore, allowing for ATP release in many signaling events. These channels have been shown to play essential roles in vasodilation/vasoconstriction, activation of the inflammasome, the taste sensation, glia-neuron communication, and apoptosis. This wide variety of physiological roles suggests multiple channel mechanisms or protein activators, the details of which are difficult to understand in the absence of a high-resolution structure. In this work, purification of pannexin is established and protein crystals obtained. Weakly diffracting crystals are improved through a combination of biochemistry and protein engineering. These important steps on the path towards structure determination have revealed some insights into the function of this unique channel protein.

BIOGRAPHICAL SKETCH

Julia Kumpf grew up in Medfield, Massachusetts, a suburb of Boston. Inspired by her teachers and excited by science, she graduated high school in 2004 and went on to further her studies at the University of Massachusetts Amherst. Julia worked with Dr. Scott Auerbach in theoretical physical chemistry studying the energy landscape of proton transfer reactions for optimizing proton electrolyte membranes in hydrogen fuels. She graduated *summa cum laude* in 2008, earning a B.S. in Chemistry and a B.A. in Physics with a concentration in secondary education. She went on to teach in a high-needs public charter school establishing the school's first chemistry and physics curricula. Julia now graduates from Cornell University in the field of biophysics with an M.S. in biophysics in 2015.

ACKNOWLEDGMENTS

No accomplishment, in science nor in life, is possible without the incredible support of a community of people. I have been fortunate in my time at Cornell to have such a community of scholars, teachers, scientists, and friends.

First and foremost, I would like to thank my advisor and mentor, Dr. Toshimitsu Kawate, for his guidance, encouragement, and support throughout this project. By his example, I have learned to be independent in my pursuit of knowledge and have the utmost respect for this quest. By encouraging me to follow my curiosity, I had a much richer experience in a wider realm of science than if I had only pursued the work discussed in this thesis. I was able to explore a number of techniques like single-molecule biophysics and chemical engineering, and expand my experience and passion for education through teaching science to undergraduates and incarcerated men. No amount of thanks can repay Toshi for the independence he afforded me in driving my own path of graduate education.

Many thanks to Kevin Michalski for being a partner in the study of pannexin. Through our discussions, more ideas were formed than possible alone. Kevin acquired some of the whole-cell patch clamp recordings discussed in this work. Thanks to other past and present Kawate lab members Dr. Akira Karasawa, Dan Levenstein, and Diego Arenas for scientific discussions and non-scientific pursuits.

Thanks to those who taught me new techniques and allowed me to explore different avenues of science. Thanks to the Cerione lab for their T4 lysozyme construct, help with lipidic cubic phase crystallography, and their know-how about plasma membrane vesicles. Thanks to Dr. Laura Byrnes and the Sondermann lab for their support with crystallography and multi-angle light

scattering, and to the Gruner lab for help with high-pressure cryocooling. Though our work is not discussed here, Mark Richards has been the collaborator I always wanted. Thanks to him and the rest of the Daniel lab for teaching me how to engineer lipidic bilayers. Thanks to Avtar Singh and the Zipfel lab for teaching me about fluorescent proteins and the tricks of single-molecule microscopy.

Most humble appreciation to those who supported me personally through this journey. Thanks to my family who know my heart and encourage me to follow it. Thanks to Lucy Brennan for constantly challenging me to be better: everything from scientific discussions to co-teaching at the prison to board games. Eternal gratitude to Sean P. Mangan – you are the calm to my storm.

Finally, thanks for financial support to Cornell University, the NIH Molecular Biophysics Training Grant, and the NSF Graduate Research Fellowship Program.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAs	Amino acids
Amp	Ampicillin
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BzATP	2'(3')-O-(4-Benzoylbenzoyl)adenosine-5'-triphosphate
$C_{12}E_8$	Octaethylene glycol monododecyl ether
cAMP	Cyclic adenosine monophosphate
CBX	Carbenoxolone
CHESS	Cornell High-Energy Synchrotron Source
CMC	Critical micelle concentration
CNS	Central nervous system
CTD	C-terminal domain
Cx43	Connexin-43
DDM	n-Dodecyl β-D-maltoside
DMEM	Dulbecco's modified eagle medium
DMNG	Decyl maltose neopentyl glycol
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DPBS	Dulbecco's phosphate buffered saline
DNA	Deoxyribosnucleic acid'
E. Coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
FBS	Fetal bovine serum
frPanx1	Pannexin-1 from *Xenopus tropicalis*
FPLC	Fast protein liquid chromatography
FSEC	Fluorescence size-exclusion chromatography

GFP	Green-fluorescent protein
Glu	Glutamate
Gln	Glutamine
HEK	Human embryonic kidney cells
His	Histidine
hPanx1	Pannexin-1 from *Homo sapien*
HPLC	High performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
LMNG	Laurel maltose neopentyl glycol
Lys	Lysine
Kan	Kanamycin
mRNA	Messenger RNA
NCBI	National Center for Biotechnology Information
NDP	Nucleoside diphosphate
NG	Nonyl glucoside
NMP	Nucleoside monophosphate
NTD	N-terminal domain
NTP	Nucleoside triphosphate
oligo	Oligonucleotide
P1	Passage 1 virus
P2	Passage 2 virus
P3	Passage 3 virus
Panx1	Pannexin-1
Panx2	Pannexin-2
Panx3	Pannexin-3
PBS	Phosphate buffered saline

PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEI	Polyethylenimine
Phe	Phenylalanine
PIs	Protease inhibitors
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
Pro	Proline
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
Ser	Serine
siRNA	Small interfering RNA
SOC	Super optimal broth
TBS	Tris buffered saline
TM	Transmembrane helix
Trp	Tryptophan

LIST OF SYMBOLS

~	Approximately
°C	Degrees Celsius
Å	Angstrom (10^{-10} m)
λ	Wavelength
µm	Micrometers
µL	Microliters
θ	Angle
bp	Nucleotide basepairs
Cu^{2+}	Copper(II) ion
d	Lattice spacing
g	grams
g	Relative centrifugal force
kDa	Kilodaltons (1.661×10^{-27} kg)
(m/m)	Ratio of two substances, as measured by mass
mg	Milligrams
mL	Milliliters
mM	Millimolar
MPa	Megapascals
mV	Millivolts
n	Integer value
Ni^{2+}	Nickel(II) ion
nm	Nanometers
nL	Nanoliters
rpm	Revolutions per minute
(v/v)	Ratio of two substance, as measured by volume
X	Times greater than

CHAPTER 1

SIGNIFICANCE

Pannexins are large pore-forming plasma membrane channels permeable to molecules up to 1kDa. Pannexins are expressed ubiquitously throughout the human body, perhaps underlining their importance, and have thus far been identified as players in maintenance of vasculature tone, central nervous system function, T-cell activation, and apoptosis. Pannexin is able to mediate such varied cellular functions by acting as an ATP-release channel in the extracellular purinergic signaling pathways (Chekeni et al., 2010).

The mechanism of this non-selective pore must be novel because its opening apparently does not affect membrane integrity during ATP release, but when rendered constitutively active by caspase cleavage of the C-terminus the cell undergoes apoptosis (MacVicar & Thompson, 2010). Because of its central role in purinergic signaling, pannexin has been identified as an important drug target for a variety of immune system diseases. A structure will greatly facilitate the development of therapeutics to modulate pannexin function, and likely lead to pharmacological treatments for diseases such as asthma, chronic obstructive pulmonary disease, ischemia-reperfusion injury, arthritis, sepsis, and inflammatory bowel diseases.

This work is an important step in the process to understanding pannexin's varied functions. Establishing purification protocols has allowed this protein to be studied *in vitro*, and modifications required to crystallize this protein have already provided functional insights into various regions of the protein. This work will provide the foundation for solving the atomic structure of pannexin, which will be essential for discovering this channel's unique gating mechanisms and role in human physiology.

CHAPTER 2

INTRODUCTION

2.1 Purinergic Signaling

Extracellular nucleotides are essential molecules for autocrine and paracrine signaling; nearly every cell in the mammalian organism releases these nucleotide mediators (Robson, Sévigny, & Zimmermann, 2006). This intercellular communication mechanism is of particular importance in both the adaptive and innate immune responses (Junger, 2011), and in the central and peripheral nervous systems, playing a role in sensory function, glia-neuron communication, and muscle contraction (Abbracchio, Burnstock, Verkhratsky, & Zimmermann, 2009).

Important molecules in purinergic signaling

Purinoreceptors, such as metabotropic P1 and P2Y and ionotropic P2X receptors, are among the most abundant receptors in mammalian tissue (Abbracchio et al., 2009). P1 receptors are G-protein coupled and regulate intracellular cAMP production upon binding extracellular adenosine. P2Y are also G-protein coupled receptors, but are involved Ca^{2+} signaling, either by cAMP/IP3-regulated release of Ca^{2+} from intracellular stores, or through interaction with ion channels. P2X receptors are ligand-gated cation channels sensitive to varying concentrations of extracellular ATP and involved in Ca^{2+} signaling within the cell (Abbracchio et al., 2009; James & Butt, 2002). These purinoreceptors are expressed in different combinations by different cell types, leading to exquisite spatio-temporal control over downstream effects. Further emphasizing the importance of these signaling pathways is the near ubiquitous expression of membrane-bound ectonucleotidases on cell surfaces. These ectoenzymes catalyze the hydrolysis of NTPs to NDPs, NMPs, and nucleosides, creating the signaling molecules necessary to

stimulate these pathways (Robson et al., 2006), and have even been shown to colocalize to sites of ATP release, suggesting a mechanism for competition with receptors for a limited supply of nucleotides (Joseph, Buchakjian, & Dubyak, 2003).

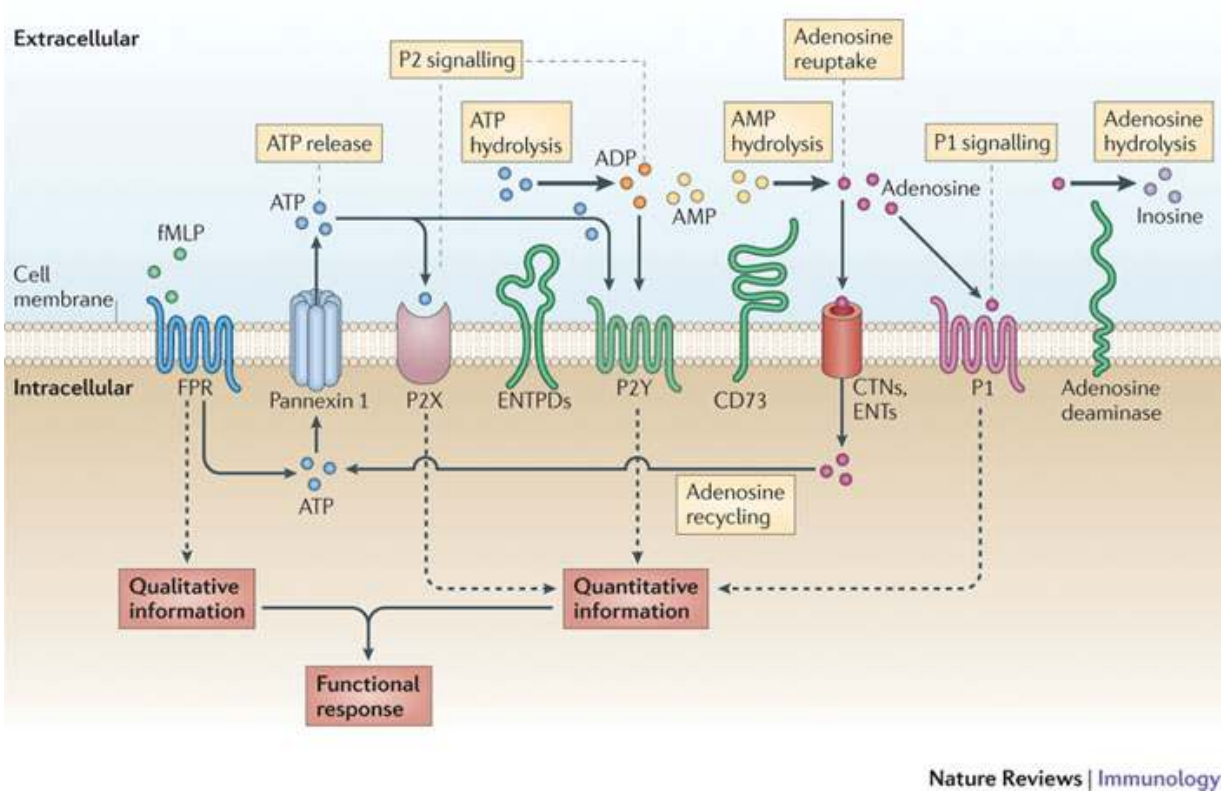

Figure 1: Key protein mediators of purinergic signaling (Junger, 2011).

Purine release from cells

Though the effects of purinergic signaling are becoming well-understood, our comprehension of an essential step – how purines are released from cells – is still in its infancy. In excitatory and secretory cells, there is much evidence to suggest that ATP is packed into vesicles (such as chromaffin granules and synaptic vesicles), and then released by exocytosis (Lazarowski, Boucher, & Harden, 2003). Only just recently the VNUT family of 12-transmembrane transporters was identified and shown to be responsible for packing ATP into vesicles using electrochemical force created by a pH gradient (Sawada et al., 2008).

In non-excitatory cells, the mechanism may be more complex. ATP may be released through exocytosis as well, however, large outward ATP gradients have been observed as a result of mechanical stress, which are unlikely to be caused by exocytosis (Praetorius & Leipziger, 2009). Further, it has been shown that ATP can be released by at least two other mechanisms from mouse erythrocytes, which are incapable of forming vesicles (Qiu, Wang, Spray, Scemes, & Dahl, 2011). In these cells, ATP has been shown to be released in a large, outwardly-directed gradient, which would be difficult to produce using exocytosis or membrane transporters. There is good reason to believe, therefore, that ATP could be released through large protein pores in the plasma membrane.

Connexin-43 (Cx43) was originally proposed to be the mediator of this process. Known to transmit large molecules (up to 1 kDa) between cell cytosols by forming gap-junction channels, this protein was proposed to open alternatively to the extracellular space as a hemichannel under some conditions (Spray, Ye, & Ransom, 2006). However, hemichannel activity from connexins can usually only be seen once extracellular calcium is greatly decreased or removed – a condition that is unlikely to have physiological significance. Further, electrophysiological activation of connexin channels requires very large depolarization steps, and as of yet have not been shown to open under physiological conditions (Praetorius & Leipziger, 2009).

2.2 Pannexin Channels

The pannexin family of channel proteins

Pannexins are a recently-discovered family of integral membrane proteins that form large-pore channels in the plasma membrane, and may be the protein responsible for non-excitatory ATP release. Pannexins release ATP during mechanical stress at resting membrane potentials (Bao,

Locovei, & Dahl, 2004), making them a better candidate than Cx43 for the ATP release channel during purinergic signaling. Further, pannexins and connexins have overlapping pharmacology, making effects of the two difficult to decouple. It is possible that early studies showing ATP release by antagonism of connexin channels may also have been antagonizing pannexin channels (Spray et al., 2006). Further, a more recent study of ATP release from astrocytes was shown to be unaffected by Cx43 knockout, but significantly decreased by siRNA knockdown of Panx1, confirming that Panx1 and not Cx43 is responsible for ATP release from these cells (Iglesias, Dahl, Qiu, Spray, & Scemes, 2009).

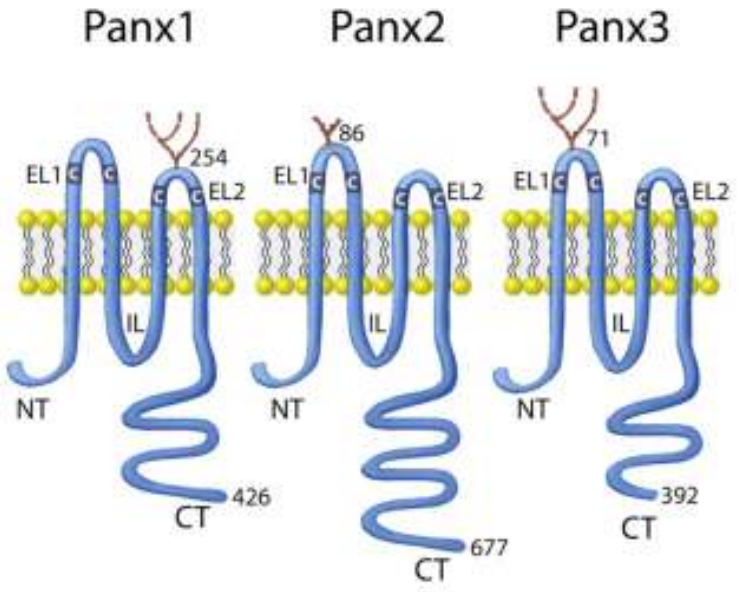

Figure 2: Schematic representation of pannexin family members. Panx1 and Panx3 show the most similarity in length as well as sequence. Panx2 has a much longer C-terminus, less sequence identity, and is predicted form a larger oligomer (Penuela, Gehi, & Laird, 2012).

Pannexins were identified in 2003 because of their sequence homology (25-36%) with innexins, which form gap junctions in invertebrates. Though neither pannexin nor innexin share homology with connexin, all three protein families have a similar membrane topology with four transmembrane domains and intracellular N-terminal domain (NTD) and C-terminal domain

(CTD). Pannexin, uniquely, has two predicted disulfide bonds on extracellular loops and one N-linked oligosaccharide (MacVicar & Thompson, 2010; Penuela et al., 2007).

Connexin channels have monomers of many different sizes and form homomeric hexamers (Maeda et al., 2009), but the oligomeric state for pannexin channels is reported to be more complex. Pannexin family members Panx1, Panx2, and Panx3 have been suggested to have different oligomeric states: Panx1 may form hexamers, while Panx2 may form a larger state (perhaps 7- or 8-mers), and Panx3 has not yet been studied (Ambrosi et al., 2010). These differences in oligomerization are somewhat surprising as the three human channels show over 50% sequence conservation (Baranova et al., 2004). Because oligomeric state can have huge implications for channel gating mechanisms, the question of pannexin assembly is, therefore, an area that must to be studied further.

Pannexin channel function

Little is currently known about pannexin channel activity, and activation mechanisms are debated among experts. Panx1 channels show large, voltage-activated outward currents when held at membrane potentials above +20mV (Bruzzone, Hormuzdi, Barbe, Herb, & Monyer, 2003), a condition that is not likely to be physiologically relevant. Bao, and colleagues have found that single-channel recordings of hPanx1 have at least 5 distinct open states, and are mechanosensitive (Bao et al., 2004). The Panx1 C-terminus appears to be pore-associated and removal of this region is required to activate the channel, suggesting a ball-and-chain-like mechanism (Sandilos et al., 2012). Locovei, and colleagues, have shown that Panx1 can be activated by intracellular Ca^{2+}, G-protein coupled P2Y receptors, and $P2X_7$ channels (Locovei, Scemes, Qiu, Spray, & Dahl, 2007; Locovei, Wang, & Dahl, 2006). This has implicated a role

for pannexin in a long contemplated large-pore-forming mode of trimeric $P2X_7$ channels (Pelegrin & Surprenant, 2006). The individual currents of Panx1 versus $P2X_7$ channels in co-transfected cells are seen by some researchers but not others (Dahl & Keane, 2012). Pannexin has a variety of known pharmacological inhibitors (Dahl & Keane, 2012), and has recently been shown to be inhibited by ATP itself (Qiu & Dahl, 2009). This seemingly counterintuitive closing mechanism may represent a negative feedback loop important in maintaining cell integrity.

Pannexin subtypes

Panx1, so far the best-studied subtype, has been found to be ubiquitously expressed throughout human tissue, perhaps signifying its importance. Especially high concentrations of mRNAs were detected by Northern blot in the heart, skeletal muscle, testis, and ovary (Baranova et al., 2004). Panx1 channels have been shown to play important ATP-release roles in many physiological processes including vasodilation/vasoconstriction, activation of the inflammasome, the taste sensation, ischemic cell death, and apoptosis, as well as in HIV-1 viral infection of $CD4^+$ cells (Penuela et al., 2012). These channels appear to form with a mixture of glycosylated species: Gly0, the non-glycosylated form, Gly1, the ER-resident high-mannose species, and Gly2, a higher-molecular weight complex oligosaccharide. Further, a glycosylation-deficient mutation in hPanx1 (N254Q) showed lesser cell surface localization, but those channels that did successfully reach the plasma membrane were functional as assayed by dye-uptake (Penuela, Bhalla, Nag, & Laird, 2009).

Panx2 appears to be expressed only in the brain and not in any other tissues. Interestingly, Panx2 is co-expressed with Panx1 in many parts of the adult rat brain, such as the hippocampus, olfactory bulb, pyramidal cells and Purkinje cells of the cerebellum (Baranova et al., 2004), but

unlike Panx1 was found in prenatal brains and further upregulated in the postnatal stage (Penuela et al., 2012). Panx2 may not be able to form homomeric channels, but may form heteromeric channels with Panx1 (Bruzzone et al., 2003), and may do so in a glycosylation-dependent manner (Penuela et al., 2009). Interestingly, when Panx2 was overexpressed in HEK293 cells, its localization was found to be mainly within intracellular compartments, but co-expression with Panx1 but not Panx3 more than doubled its cell-surface expression (Penuela et al., 2009).

Panx3 is the least-studied homologue, and was found to be non-functional when expressed in *Xenopus* oocytes (Bruzzone et al., 2003). Panx3 is expressed mainly in mouse osteoblasts, synovial fibroblasts, whole joints, cartilage and cochlear bone, as well as low levels in human hippocampus, and has been implicated as being important for stimulating chondrocytes and osteoblasts to differentiate (Iwamoto et al., 2010).

2.3 Physiological Roles of Pannexin

In apoptosis

Apoptotic cell death occurs in all multi-cellular organisms – humans turn over cells at the extraordinary rate of nearly 1 million cells per second. This process is extremely efficient and well-regulated, and is an essential part of development, immunity, and homeostasis of many cell types. However, it is extremely difficult to observe apoptotic cells due to the speed and proficiency of phagocytic removal, even in tissues undergoing constant cell turn-over, such as in the thymus (Ravichandran & Lorenz, 2007). There is significant evidence to show that apoptotic cells communicate their presence to phagocytes by releasing ATP and UTP as 'find-me' signals during the early stages of apoptosis. As these molecular signals diffuse away, phagocytes expressing the $P2Y_2$ receptor recognize and follow a nucleotide concentration

gradient back to the cell and complete phagocytosis promptly (Elliott et al., 2009). Chekeni, et al. were able to show that Panx1 was required for both release of ATP/UTP from apoptotic cells and recruitment of phagocytes. siRNA knockdown of Panx1 from apoptotic Jurkat T-cells coupled with CBX inhibition proved to nearly eliminate ATP release, while having no effect on the progression of apoptosis. Further, they showed that this nucleotide release is directly correlated with Panx1 C-terminal cleavage by caspases 3 and 7, known to be expressed during apoptosis (Chekeni et al., 2010). Without these extracellular signals necessary for the efficient clearing of dead cells, developmental abnormalities, autoimmune disorders, and non-resolving inflammation can occur.

In the central nervous system

Though originally glia were thought to serve only as structural reinforcement for neurons, it is now widely accepted that the function of these non-excitatory cells provide essential support for neuronal activity (Kimelberg, 2010). For example, purinergic signaling appears to be critical for altering brain vasculature during times of neuronal metabolic stress, such as during hypoxia (lack of oxygen) and hypercapnia (lack of glucose) (Iadecola & Nedergaard, 2007). This response is hypothesized to be stimulated by neurons via glutamate and ATP co-release, presumably through exocytosis of VNUT-packed vesicles, leading to local activation of purinoreceptors on astrocytes by Ca^{2+} influx, propagation of Ca^{2+} waves, and further spreading of ATP release. Astrocyte endfeet then release ATP, which interacts with vascular receptors that can dilate or constrict the vessels nearby the neuronal activity to bring nutrients to actively firing neurons (Erlichman, Leiter, & Gourine, 2010). In support of pannexin's role in this astrocytic ATP signaling wave, Iglesias, et al. showed that in cultured astrocytes, when Panx1 is stimulated

to open by either voltage clamp or the $P2X_7$-agonist BzATP, ATP release occurs and can be diminished by Panx1 knockdown by siRNA(Iglesias et al., 2009). Further, they were able to show the same channel activation and ATP release exist in astrocytes from Cx43-null mice, suggesting that it is indeed Panx1 channels and not Cx43 hemichannels that are responsible for ATP-mediated signaling in these glia.

In the immune system

In the immune system, ATP release has been identified both from damaged tissue that activates inflammatory responses, and from human neutrophils and T-cells during normal physiological conditions. It is well-known that T-cell activation requires binding of T-cell receptor with antigen at the immune synapse, resulting in increase in intracellular Ca^{2+} levels. Recently, Panx1, $P2X_1$, and $P2X_4$ (but not $P2X_7$) channels were seen to translocate to the immune synapse and initiate an autocrine ATP-mediated signaling event that directly led to Ca^{2+} entry (Woehrle et al., 2010). Inhibition of Panx1 was enough to stop Ca^{2+} entry and T-cell activation, pointing at pannexin's critical role in antigen-stimulated immune response. Human neutrophils have also been seen to release ATP through Panx1 channels, but not through Cx43, in response to bacterial peptides and inflammatory mediators that activate Fcγ, IL-8, C5a complement, leukotriene B_4 receptors (Chen et al., 2010). Inhibition of Panx1 in this system blocked neutrophil activation and diminished response to bacterial infection, further implicating pannexin as an excellent drug target for neutrophil-induced organ damage (Nussler, Wittel, Nussler, & Beger, 1999).

2.4 Essential Questions about Pannexin Channels

Previous studies have elucidated the central importance of the pannexin channel in a wide variety of purinergic signaling events, from T-cell activation to glia-neuron communication to

apoptosis. Blocking of pannexin activity using pharmacological inhibition can stop undesirable events, making pannexin an important drug target. Despite its significance, basic questions are still unanswered about pannexin's mechanism: How is pannexin activated *in vivo*? How does pannexin release ATP during signaling events while maintaining cell integrity? How exactly do pannexins, purinoreceptors, and ectonucleotidases interact to control signaling?

This work is an essential contribution to this emerging field by showing the first successful purification of a pannexin channel. Further, this work shows that pannexin is capable of being crystallized, an important step on this path to solving an atomic-resolution structure. I expect a crystal structure will open up the field to understanding of the molecular mechanisms behind pannexin's unique function, and facilitate the development of therapeutics against a wide variety of diseases.

2.5 Principles of Membrane Protein Crystallography

Since the first protein structure was determined through crystallization 150 years ago, the technique has become essential to the fields of biology and biochemistry. The three-dimensional structure of a protein is extremely useful for uncovering the molecular mechanisms by which proteins function and interact. The key reason function is deduced from structure is because it is protein structure that is well-conserved through evolution, while primary sequence is not (Brenner, 2001). Though solving a protein's structure is essential, the process is quite involved and often requires considerable effort.

Membrane proteins are a particularly difficult challenge for structural biology, though they regulate some of the most important cellular functions. Approximately 25% of all proteins are membrane proteins, yet fewer than 150 unique protein structures have been determined, and

very few of these are eukaryotic membrane proteins (Carpenter, Beis, Cameron, & Iwata, 2008). Purification and crystal formation are the main bottlenecks in structure determination of these proteins crystal formation.

Overview of structure determination

Figure 3 demonstrates each step in the process of structure determination, from choosing a protein target to depositing a structure in the Protein Data Bank. While crystallization is often considered to be the bottleneck in this process, indeed each step is a separate challenge that must be overcome (Brenner, 2001).

In this work, we will show many successful steps in this process: selection of a target pannexin protein, cloning of coding sequences, expression in Sf9 insect cells, solubilization optimization through detergent screening, establishment of purification procedures, quality assessment using multi-angle light scattering, and successful protein crystallization. When crystals were not diffraction-quality, the protein target was redesigned, recrystallized in new conditions, and rescreened for high resolution diffraction. In the end, we were able to push the diffraction resolution from initial crystal hits that diffracted to 25Å to optimized proteins and crystals that diffracted 9Å.

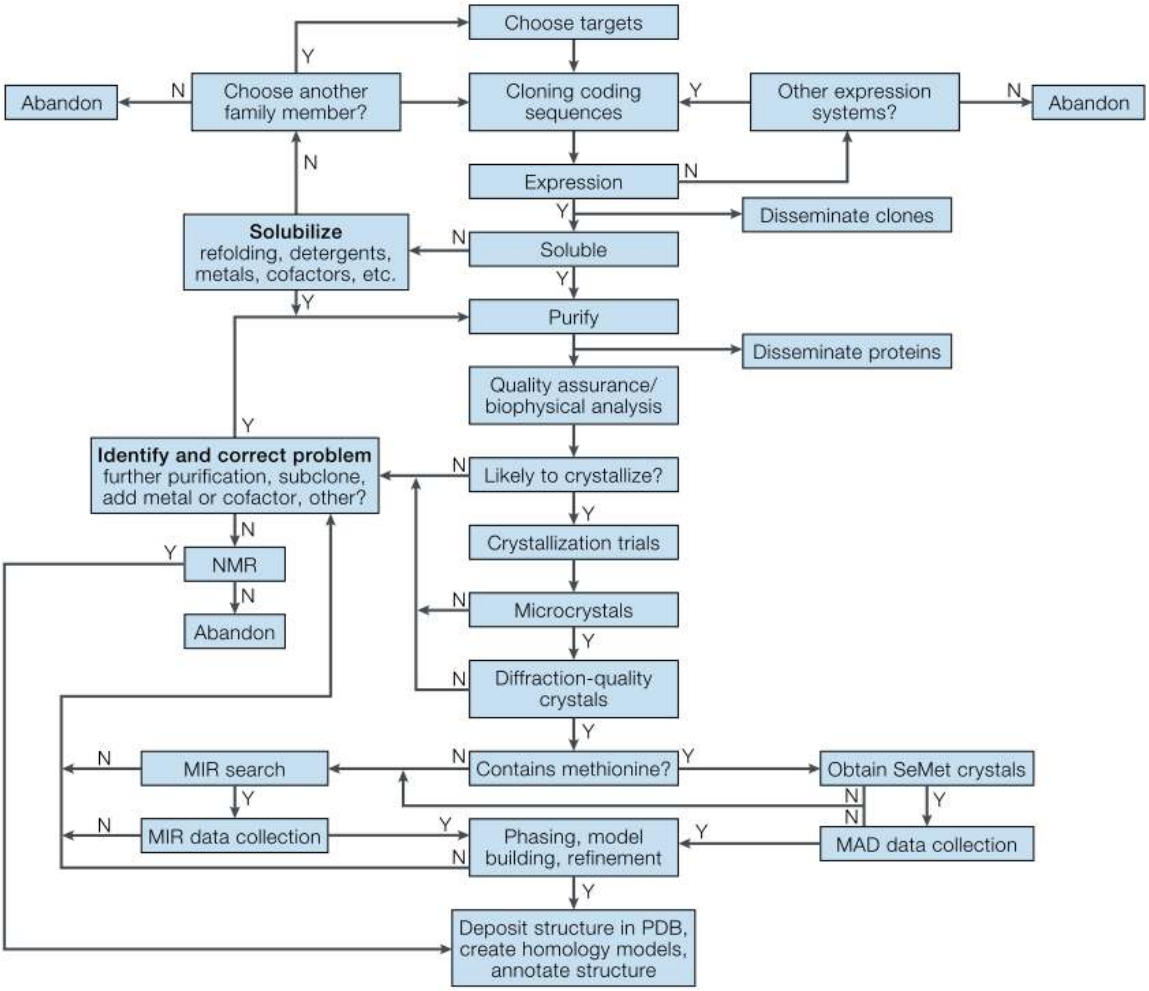

Figure 3: Process of protein structure determination, suggesting the troubleshooting and path change necessary when steps of the process fail. N indicates that a process failed and Y that it has succeeded (Brenner, 2001).

Purification

Integral membrane proteins, such as pannexin, contain both a hydrophobic core, which embeds the protein in the phospholipid bilayer, and hydrophilic regions which mediate intercellular and extracellular functions. This chemical duality is the primary reason membrane protein structures are more challenging to solve. The key is in extracting membrane proteins from the membrane while maintaining native quaternary structure. Detergent molecules, which also have hydrophobic and hydrophilic regions, must be used. Detergents create micelles around the

proteins; the hydrophobic end stabilizes the core of membrane proteins, while the hydrophilic end faces out towards the solution maintaining solubility. The choice of detergent is an essential part of both extraction and crystal formation: the detergent must be mild enough to solubilize the protein in its native conformation over long periods of time in changing chemical conditions, while simultaneously creating a small, compact micelle that allows for good protein-protein contacts to form a crystal (Carpenter et al., 2008).

Thus, purification starts by disrupting cells and isolating membranes using differential centrifugation. Proteins are extracted from membranes using detergents and then must be isolated from all other membrane-bound macromolecules. Proteins are expressed using an affinity tag: a short 8 amino acid sequence that binds tightly and specifically to a certain chemical moiety. In this work, a Strep II tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) is expressed on one end of the protein of interest. This tag is a high affinity ligand for streptavidin, a bacterial protein which is fused onto StrepTactin resin. When the membrane extract is incubated with this resin, tagged proteins of interest will bind to the resin, while all other proteins will be washed away. Tagged proteins are then eluted from the resin by adding D-desthiobiotin, an analog of streptavidin's native ligand, D-biotin, which competitively releases the tag and protein of interest with it (Schmidt & Skerra, 2007).

The tag is then cleaved from the protein using the enzyme thrombin. Finally, the protein is isolated using size-exclusion chromatography (SEC), or gel filtration. The protein is run across a long column of resin made of cross-linked polymer beads. These beads are engineered to contain pores of a particular size. Thus, as proteins run through this column the smaller proteins get "stuck" in these cavities while larger proteins run straight through, ultimately separating by

size (Nelson & Cox, 2008). Since pannexin (~600kDa), EGFP (33kDa), and thrombin (37kDa) are vastly different in size, the SEC column efficiently separates the pannexin protein from others. This protein is collected and its purity determined by SDS-PAGE gel. If found to be pure, crystallization can be attempted.

Crystallization Theory

Protein crystallization, like the crystallization of small molecules, involves a slow dehydration until enough water leaves that the proteins begin making direct contact. These weak protein-protein interactions can either form in a regular way, causing the proteins to self-assemble into a crystal lattice, or the interactions can form in an irregular way, causing the proteins to become insoluble aggregates. The huge variability between protein's intrinsic properties (such as flexibility, surface charge distribution, and conformational homogeneity) make it impossible to predict chemical conditions that will be right for crystallization (Rupp, 2010). Thus, many different chemical conditions, as well as many different protein constructs, must be tested in order to find a single crystal hit.

The most common technique for growing crystals is the hanging-drop vapor diffusion technique. In this technique, a protein solution is mixed with a precipitation reagent cocktail on a siliconized glass coverslip. This cocktail includes a precipitant, additives, and buffer to optimize pH. When mixed with the protein in 1:1, 1:2, or 2:1 ratios (v/v), the protein will ideally remain soluble. This coverslip is then sealed with the drop hanging above a well full of the precipitation cocktail, called "mother liquor," and incubated at a constant temperature.

In this sealed system, over time, a vapor equilibrium is established. In search of this chemical equilibrium, water slowly diffuses from the drop to the mother liquor, dehydrating the proteins

and precipitants in the drop. As the concentration increases the solution becomes supersaturated and, depending on the dehydration kinetics needed to reach thermodynamic equilibrium and the level of supersaturation, the drop mixture will fall into one of three zones – metastable, nucleation, or precipitation – as diagrammed in Figure 4.

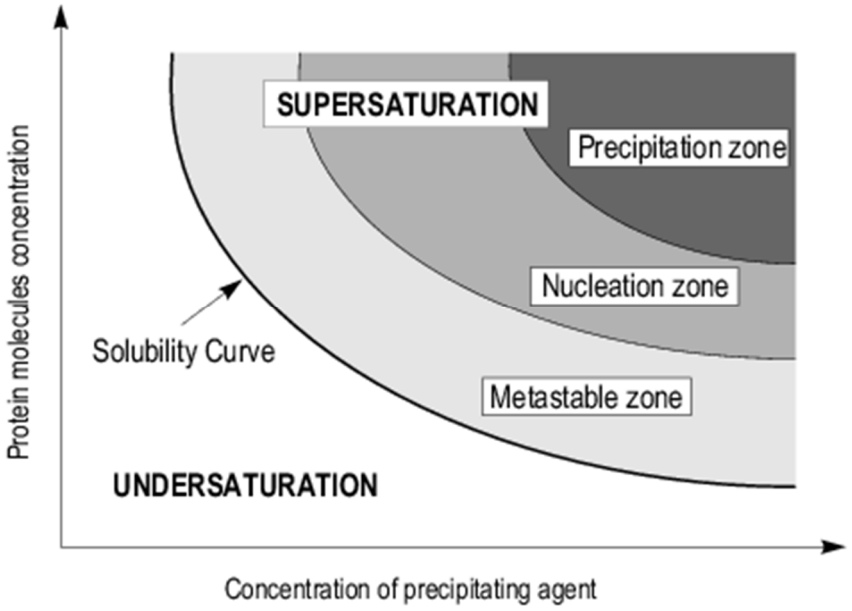

Figure 4: Crystallization diagram that schematically represents the solubility phase space of a vapor diffusion drop as a function of protein and precipitant concentration (Ducruix & Giege, 1992).

It is desirable to initially form an undersaturated solution of protein where the soluble protein is stable and in thermodynamic equilibrium. To form a crystal the protein must become supersaturated; a metastable state forms where the kinetics are hindered and do not allow equilibrium to be reached. In the metastable zone, the supersaturated protein will remain soluble, never able to attain the activation energy needed to reach equilibrium (Rupp, 2010).

It is only within the nucleation zone that the initial stages of crystal formation can occur. Proteins in the nucleation zone may overcome the activation barrier by spontaneous or external creation of homogeneous nucleation sites. This kinetic barrier stems from favorable collisions between

protein molecules. These collisions must produce contacts whose gain in enthalpic energy (due to charge stabilization or hydrophobic packing) offsets the loss in entropic energy (due to the increase in order). Once the kinetic barrier is overcome, thermodynamic equilibrium can be reached as excess proteins come out of solution and the protein crystal grows. In this case the protein concentration, but not precipitant concentration, would decrease until a saturated solution is reached. Further crystal growth will continue in this cyclical pattern of dehydration, supersaturation, activation energy reached, crystal growth, and a return to saturation until the equal concentrations between drop and well solutions is reached and diffusion of water reaches chemical equilibrium (Rupp, 2010).

If spontaneous homogeneous nucleation does not occur, the precipitation zone may be reached. This represents the point at which the concentrations of protein and precipitant are so high that kinetics dictate immediate precipitation must occur. Thus, amorphous aggregates form and crystallization is no longer possible (Rupp, 2010).

Crystallization, as discussed previously, is the primary bottleneck in structure determination. In practice, it is impossible to predict what this solubility diagram will look like for various proteins, as the kinetics and thermodynamics are affected by the crystallization conditions (buffers, salts, precipitant, pH, temperature, etc.) and even by the protein itself (detergent choice, flexible regions, pI, surface charges, conformational differences, etc.). Thus, conditions must be screened and hit conditions optimized. When crystallization fails, the protein itself, the detergent it is stabilized in, and the precipitation conditions, must all be modified.

X-Ray diffraction

X-ray diffraction is a powerful technique that can be used to determine the atomic level structure of any molecule. X-rays are a high-energy form of electromagnetic radiation with extremely small wavelengths of 0.01-10nm (1-100Å). Because the length of a typical bond is in the 0.15nm (15Å) range, X-rays can resolve atomic-level structures, even molecular structures as large as a protein.

When X-ray radiation encounters the electron density of a molecule (in atoms and within covalent bonds), the X-rays scatter. This scattering from a single molecule produces an irregular modular function and is too small to be measured. But when molecules are crystallized into a well-packed lattice with high degrees of symmetry, scattered X-rays can give rise to constructive interference and signal amplification, which we call crystal diffraction (Rupp, 2010). When a high-resolution diffraction pattern for an entire 360° rotation of a crystal can be obtained in an X-ray diffraction experiment, the protein's molecular structure calculated.

The diffraction pattern produced is a complex pattern of spots at specific intensities that are specific distances apart from each other. The intensities of the diffraction pattern come from the electron density of the molecule, while the spacing of the spots (the sampling rate of that molecule) come from the crystal lattice itself (Rupp, 2010).

The electronic structure of the molecule dictates how the X-ray electromagnetic field will scatter. Specifically, the molecular scattering function is a superposition of all atoms in the molecule, where heavier atoms scatter more, lighter atoms scatter less, and space between atoms do not scatter at all. Each atom *j* contributes to the scattering intensity by an atomic scattering

factor, $f_{s,j}^{0}$. The molecular scattering function, seen in the intensities of spots in the diffraction pattern, is a Fourier transform of the electron density of the molecule.

This scattering function can be intensified by arranging many molecules in a periodic lattice. However, when X-rays diffract off a crystal, a discrete sampling of the scattering function occurs as a result of the lattice planes of the crystal. This sampling pattern can be measured and the geometry of the lattice itself calculated. This relationship is known as Bragg's law and is diagrammed in **Error! Reference source not found.**.

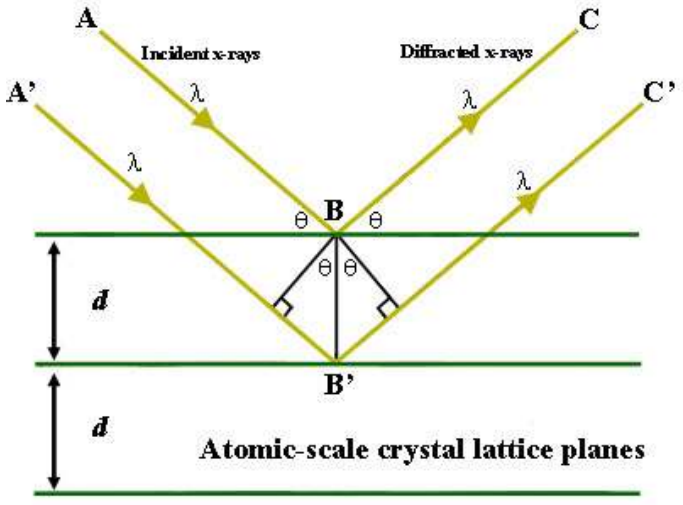

Figure 5: Bragg's Law. This diagram represents how X-rays can produce a diffraction pattern. This result occurs when incident X-ray beams scatter from electron density in symmetric crystal lattice planes and interfere constructively (Henry, Eby, Goodge, & Mogk, 2012).

An incident plane wave A approaches the crystal lattice plane, with wavelength λ and at angle θ, will reflect off the plane at the same angle. Because this scattering can occur at multiple planes within the crystal, another wave A' might scatter in a similar fashion with a similar point such that the two waves A and A' interfere constructively at the detector (points C/C'). For this interference to occur, the second wave must be in-phase with the first, meaning it's phase is some integer number of wavelengths, nλ, different) and have travelled a distance of 2*dsinθ

farther than the first wave, where d is the spacing between lattice planes. Thus, Bragg's equation becomes:

$$n\,\lambda = 2\,d\,\sin\theta$$

Putting it all together, the crystal lattice dimensions, d (or d_{hkl}) represent the real lattice of the crystal, while the diffraction spacing are shown in reciprocal space, $1/d_{hkl}$ where h, k, and l represent the indices of each reflection. The lattice spacing and atomic scattering factor are combined with the exponential summation formula for wave superposition, and the complex structure factor F_h results:

$$F_h = \sum_{j=1}^{all\ atoms} f_{s,j}^{0} \cdot e^{2\pi i \cdot hx_j}$$

Thus, the scattering function of the unit cell is expressed in fractional coordinates x_j and reciprocal lattice index h. Thus, for each diffracted X-ray the atomic scattering function (and thereby the electron density of the molecule) can be calculated (Rupp, 2010).

One major factor in determining the quality of a protein crystal is its resolution, which is defined as the limit in diffraction angle at which X-ray reflections can be detected. To obtain enough reflections to solve a protein's structure, the resolution of diffraction must be close to the atomic scale (2-4Å), resolving the electron densities of individual amino acid side chains. Crystals that diffract to high resolution must consist of rigidly packed molecules with very small lattice spacing. Indeed it is this challenge that this work will partially, and future work must successfully, overcome.

CHAPTER 3

MATERIALS AND METHODS

3.1 DNA Construct Creation

Pannexin and innexin gene constructs were synthesized (GenScript) and codon optimized for expression in Sf9 insect cells. Genes were created with BamHI and XhoI restriction sites at the N-terminus and C-terminus, respectively, for use with Kawate Lab vectors. All vectors for mammalian cell and insect cell expression were previously established (Kawate, unpublished) and outlined in Appendix A.

General molecular biology

Genes were cut and pasted between vectors as new expression constructs became necessary. By containing features such as EGFP, a thrombin cleavage site, and various affinity tags within the vector, genes were able to be simply dropped into a new vector to modify expression conditions. Double digestions using restriction enzymes BamHI and XhoI were used to extract gene sequences from synthesized vectors for 3-6 hours at 37°C using standard conditions (New England BioLabs). DNA was run on a 1% agarose gel and extracted using a Gel Extraction Kit (Qiagen or Omega). DNA was ligated into vectors for 30min on ice using T4 Ligase and standard conditions (New England BioLabs).

Ligated DNA was transformed into DH5α *E. Coli* for 30min on ice, heat shocked for exactly 45 seconds at 42°C, and returned to ice. Bacteria were then grown on LB plates with ampicillin (LB+Amp) overnight at 37°C to select for bacteria that took up the plasmid and corrected the nicks in the ligated DNA. These bacteria were picked to be amplified overnight in 8mL

LB+Amp cultures and incubated with agitation for 12-16 hours. Bacteria were then broken and plasmids harvested using a DNA Miniprep Kit (Qiagen or Omega).

Plasmid DNA was sent to the Cornell Institute of Biotechnology Genomics Facility for DNA sequencing analysis. Sequences for the gene in question and flanking sequences were returned and verified against their expected sequence.

PCR

The pannexin protein needed to be modified in various ways throughout this work to make it more easily purified or crystallized; one common mechanism was through sequence deletion. PCR was utilized to extract out only the sections of the gene that were to be recovered. Short DNA oligonucleotides (18-21 bp) were designed be complementary to the beginning and end of the sequences. Oligos, template DNA, dNTPs, reaction buffer, and Herculase II Fusion DNA Polymerase were mixed in standard reaction conditions and submitted to thermal cycling as recommended (Agilent Technologies). For more complicated sequences multiple rounds of PCR were necessary; fragments created in the first round were used as both oligo and template in the second/third rounds. Successful PCR fragments were gel purified, digested, ligated into standard lab plasmids, and transformed into DH5α *E. Coli* as described on page 34. A single bacterial colony was selected, amplified, and the plasmid extracted. Each gene sequence that underwent PCR was verified by DNA sequencing.

Mutagenesis

Point mutations or short (2-4bp) sequence deletions were achieved using Quikchange mutagenesis. Quikchange primers were typically longer than PCR primers (20-30bp) and included the mutated base in the middle of the oligo. In the quikchange reaction, both primers

annealed to the same sequence of the gene, and proceeded to copy the entire plasmid. Standard quikchange thermal cycling was used (Stratagene). Template plasmids were then digested with DpnI which would cut only methylated DNA (from bacteria) and not the newly mutated DNA. Finally, the plasmids were transformed into DH5α *E. Coli*, a single colony isolated and amplified, the plasmids extracted and sequenced as described on page 34.

3.2 Protein Expression and Characterization

HEK cell expression

Proteins were produced for screening purposes to characterize the expression level and monodispersity of the protein construct. These proteins were first screened in HEK cells due to their ease of transfection. HEK cells were grown in sterile 6-well plates in DMEM +10% FBS until they reached 60-90% confluence, at least 24 hours. Transfection mixtures were prepared in 150µL OPTI-MEM (Life Technologies) by mixing 4ug DNA and either 6µL PEI (1mg/mL) or 6µL FuGENE-6 (Promega) and allowed to incubate at room temperature for 20 minutes. Transfection mixtures were dropped carefully onto cells and incubated for 48 hours before use without changing the media.

Protein screening using fluorescence size exclusion chromatography

Proteins tagged with EGFP on the N- or C-terminus were screened for correct size, monodispersity, and expression level in whole-cell lysates of HEK cells using fluorescence size-exclusion chromatography (FSEC). Two days after transfection, HEK cells were washed with 1mL DPBS, then removed from the plate and suspended in 1 mL DPBS. After centrifuging to collect cells at 3,000rpm for 5min, cells were broken by solubilization into 150µL of S-Buffer. This buffer, containing 1% DDM (a mild detergent) and a protease inhibitor tablet (Roche) in

DPBS, was mixed thoroughly by pipetting up and down and then rotating at 4°C for 20 minutes. Insoluble debris was removed by ultracentrifugation at 40,000rpm for 30 minutes.

The whole-cell lysate, including the solubilized EGFP-tagged proteins, were run over a Superose6 gel filtration column (General Electric) to separate by size, and run with FSEC Buffer (containing 1mM DDM in PBS) to maintain a DDM concentration at least two times higher than the critical micelle concentration. Finally, proteins were run through a fluorescence detector. They were excited by 488nm light and the emission detected at 520nm, corresponding with the excitation/emission spectrum of EGFP. Thus, only the proteins of interest tagged with EGFP were visible in the whole-cell lysate, allowing for quick, cheap, efficient characterization. Proteins that showed higher levels of expression, low levels of aggregation, and a monodisperse oligomeric peak were earmarked for purification.

Sf9 cell expression

Once characterized, proteins needed for purification and crystallography were expressed in Sf9 insect cells using the Bac-to-Bac Baculovirus infection system (Life Technologies). The pFastBac plasmids were previously modified (see Appendix A) for tagging with EGFP and the Strep affinity tag in both the N-terminal and C-terminal positions.

Plasmids were transformed into DH10Bac *E. Coli* cells on ice for 30 minutes, heat shocked for exactly 45 seconds at 42°C, and recovered in SOC medium for 4 hours in order to allow time for the donor gene from pFastBac to become incorporated into the DH10Bac "bacmid" DNA. Bacteria were then plated on LB with kanamycin, gentamicin, tetracycline, IPTG, and Bluo-gal and grown overnight at 37°C. White colonies (indicating the gene had been incorporated) were selected.

DH10Bac was then amplified in 6mL cultures and the high molecular weight bacmid DNA extracted using a phenol chloroform extraction. Cells were solubilized and RNA degraded using Buffers 1, 2, and 3 from the DNA Miniprep Kit (Qiagen or Omega), as recommended, to extract the soluble cell macromolecules. The soluble extract was then vortexed well in a 1:1 (v/v) ratio with a Phenol:Chloroform:Isoamyl Alcohol reagent (25:24:1, v/v) to remove lipids and other hydrophobic molecules. The resulting suspension was centrifuged at maximum speed to separate layers, and the aqueous phase retained. The soluble extract was next vortexed 1:1 (v/v) with chloroform, separated by centrifugation, and again the aqueous phase retained. Finally, a 2:1 (v/v) ratio of 100% ethanol was added and incubated for 15min at -20°C to precipitate the bacmid DNA. Once DNA is observed, it was spun at top speed for 15min in a 4°C centrifuge. The DNA pellet was rinsed with 70% ethanol, dried in a vacuum, and then resuspended in 50uL water.

To make the Passage 1 (P1) virus, Sf9 cells were plated with $8x10^5$ cells/well in a 6-well plate in 2mL Sf-900 III media (Life Technologies), then transfected with 10uL bacmid DNA and 8uL Cellfectin (Invitrogen) in 200uL media. Cells were incubated for 5 hours, then the media replaced with a fresh 2mL and placed in a 27°C incubator. After 5-7 days, the P1 virus was dense within in the media above cells, as judged by the brightness of EGFP fluorescence. The virus was harvested and stored by extracting the 2mL media, filtering, 1:50 Pen-Strep added, and incubated at 4°C.

Passage 2 (P2) virus, created from P1 virus, and was found to be more cost-effective for producing protein than P1 or P3. This virus was made by infecting a 100mL culture (at ~$1x10^6$

cells/mL) with 50uL P1 virus, growing at 27°C, and then harvesting by filtration on Day 3. P2 virus was stored at 4°C and never retained for more than 4 weeks.

Finally, purification-quality proteins were expressed from Sf9 cells infected with P2 virus. Between 1-6 liters of Sf9 cells were grown from a density of $0.5x10^6$ cells/mL in 2L flasks at 27°C, and infected with 16mL of P2 virus at a density of $3.5\text{-}4x10^6$ cells/mL. Some protein constructs were found to produce better protein when Sf9 cells were grown with shaking at 27°C for 48 hours, while others needed to be grown with shaking at 27°C for 24 hours, followed by 48 hours at 18°C.

3.3 Protein Purification

Two days after infection, Sf9 cells were collected by centrifugation (1000*g*, 10min), washed with TBS, and then resuspended in TBS with protease inhibitors (aprotinin, pepstatin A, phenylmethylsulfonyl fluoride, and leupeptin). Cells were pressurized to 750Psi with a N_2 gas in a cell disruptor, and incubated at 4°C for 20-60min. Cells lysed as they were released, and the lysate was collected. Large cellular debris was removed by centrifugation at 10,000*g* for 10 minutes, then the membrane fraction collected by ultracentrifugation at 100,000*g* for 1 hour.

The membrane was weighed and solubilized with 360mL S-Buffer: 1X TBS containing 0.5g $C_{12}E_8$ detergent per 1.0g of membrane. Membrane was homogenized using a Daunce homogenizer, and mixed at 4°C for 1hr. The solubilized membrane proteins were isolated from insoluble aggregates by centrifugation at 100,000*g* for 1 hour. TALON (Clontech) or Strep (GE Healthcare) affinity resin was added to the supernatant (1mL per 1g membrane) and rotated at 4°C overnight. The resin was collected by 1-3 gravity filtration columns, washed with 10-20 column volumes of Strep-Tactin Wash Buffer, and eluted with 6 column volumes of Strep-

Tactin Elution Buffer (containing 2.5mM desthiobiotin). Elute was concentrated to 1mL using 100K concentrators (Amicon), and injected 0.5mL at a time over a Superdex200 gel filtration column (General Electric) and run in crystallization buffer (20mM HEPES, 50mM NaCl, 15% Glycerol, and a crystallization detergent at least 2X its CMC). Finally, peak fractions were collected and concentrated, and aggregate was separated by ultracentrifuge before final use.

3.4 Protein Crystallization

Nanoscale crystallization screening

When a new protein construct was isolated and deemed to be monodisperse and pure, crystallization conditions were screened using the vapor diffusion method. First attempts were typically used 96-well broad sparse matrix screens: PEG Ion I/II and Index (Hampton), MemGold (Molecular Dimensions), and the lab's own FK screens (Furukawa and Kawate, unpublished).

Initial screens were set up by the Mosquito crystallization robot (TTP Labtech) using nanoliter volumes of protein and precipitation mixture, typically 100nL:100nL and sometimes 100nL:200nL or 200nL:100nL, respectively. The drops were formed using the stock mixtures from the crystal screen, but were incubated with well solutions that contained a 15% glycerol additive. This was important because when mixed in a 1:1 ratio, the drops contained around 7.5% glycerol (the protein buffer contained 15%). Thus, when incubated over stock solutions (0% glycerol), vapor diffusion would essentially run in reverse from the desired direction: diffusing water from the well into the drop, causing the drops to swell. Adding 15% glycerol to the well solutions reversed this effect and drop dehydration was again found to occur slowly enough for crystal formation. Initial protein-detergent crystal screenings were incubated at 4°C,

while protein-bicelle crystal screens were incubated at 20°C. Drops were monitored using Rock Imager (Formulatrix) regularly for 4 weeks; any crystal hits were noted and optimized using microscale optimization.

Microscale crystallization optimization

Crystal hits were optimized for PEG length and concentration, salt concentration, and molecular additives including agonists, antagonists, and other crystallization additives. The drops were mixed on glass coverslips in a 24-well format using 500μL:500μL of protein and crystallization condition, respectively. Crystals from these conditions often grew bigger than those from nanoscale screening and so these were primarily scooped and frozen in liquid nitrogen for X-ray diffraction experiments.

3.5 Crystal Analysis by X-Ray Diffraction

Cryoprotection

Cryoprotecting the crystals was important for a few reasons: to prevent dehydration, for ease of transport to the synchrotron, and to slow radiation damage. Given that the crystallization drops contained 7.5% glycerol and the mother liquor contained 15% glycerol (in addition to the exact same salt, precipitant, and additive concentrations as the drop), the glycerol concentration was high enough to use the mother liquor for cryoprotection. Proteins were scooped with X-ray invisible loops (Hampton), washed and protected in mother liquor, and finally frozen in liquid nitrogen.

X-Ray diffraction

Frozen crystals were brought to the A1 and F1 lines of the Cornell High Energy Synchrotron Source (CHESS). Loops were mounted on the beam lines and the crystal moved to the center of the beam. We took two snapshots of the crystal, using X-ray radiation with wavelength of 0.976Å, to assess diffraction resolution limit at both 0° and 90° crystal rotations for 30 second exposures. For crystals that diffracted to at least 15Å, we captured diffraction patterns for ten 1° wedges. Using these slices and HKL2000 software, the unit cell dimensions and space group of the crystal were estimated.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Pannexin Crystallization Target Identified

A library of pannexin & innexin genes is obtained

As species evolve, their proteins continue to evolve as well: a point mutation here, a deletion there. Over evolutionary time, proteins from one species end up with varied primary structures, while their secondary, tertiary, and quaternary structures and most importantly their functions tend to be extremely well-conserved.

While understanding the human pannexin protein is the ultimate goal of this field of research, this high conservation makes it possible to infer structure and function from a number of pannexins from other species, and the homologous innexin proteins as well. Because of these small differences in sequence, each homologous protein will have slightly different chemical behavior *in vitro*, which can lead to huge differences in their crystallizability.

Thus, a library of genes homologous to human pannexin-1 (hPanx1) were obtained through homology searches of the known genomic databases using a BLAST search (listed in Appendix B). These genes were codon-optimized for insect cell expression and synthesized (Kawate). Genes were then subcloned into the pNGFP-EU2 and pCGFP-EU2 vectors for FSEC screening. It was important to tag both N- and C-termini because without much structural information it is impossible to know what affect this tagging will have on the channel.

FSEC screening identifies a well-behaved target protein

Pannexin 1, 2, 3, and innexin constructs were expressed in HEK cells for medium-throughput FSEC screening. Expression within cells (brightness, plasma membrane trafficking, and cell health) was monitored.

After a 48 hour expression, cells were lysed and lysates subjected to FSEC. The FSEC profile of a well-expressed, well-behaved protein would show very little aggregate (the first peak to elute around 15min), a large oligomeric pannexin/innexin peak (around 23min), no pannexin/innexin monomer peak (around 30min), and little to no free-GFP peak (around 37min).

While many FSEC profiles showed some of these characteristics, only one protein demonstrated all of these characteristics when screened in the mild detergent n-Dodecyl-β-D-Maltoside (DDM). This protein was the N-terminally GFP-tagged pannexin-1 protein from *Xenopus tropicalis*, henceforth referred to as frPanx1. This protein was 64% identical to the human version (see pannexin-1 ortholog alignment in Appendix C). This protein was selected as the best candidate for purification and crystallization. Its FSEC trace can be seen in Figure 6, in comparison with a few selected traces, demonstrating some of the ways other pannexin/innexin proteins behaved when solubilized.

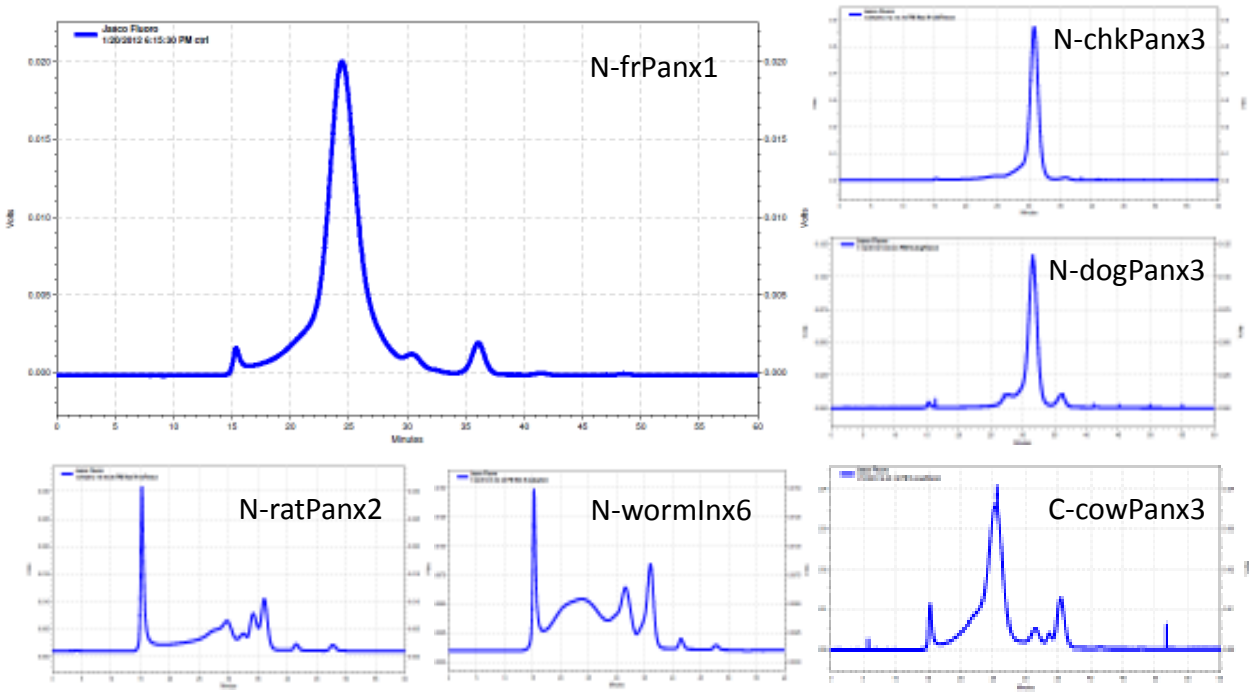

Figure 6: Representative FSEC profiles of pannexin ortholog screening. "N-" represents an N-terminal GFP tagging position, while "C-" indicates tagging at the C-terminus. Here, frPanx1 showed the best oligomeric peak (eluted around 23min) while showing minimal: aggregate peak (15min), monomer peak (30min) and free-GFP peak (37min). Other profiles show large aggregation (N-ratPanx2, N-wormInx6), monomeric protein solubilization (N-chkPanx3, N-dogPanx3), or polydispersity (C-cowPanx3, N-ratPanx2, N-wormInx6).

While only the frPanx1 protein was a candidate to move into crystallization trials, many other molecular trends were observed during screening. Overall, many of the N-terminally GFP-tagged constructs showed high expression of proteins that localized to the plasma membrane, while the same protein tagged C-terminally caused cells to be unhealthy or even die. This cell death was unable to be recovered by adding the pannexin inhibitor CBX. Thus, it is important for cell viability that the C-terminus remain unhindered. We hypothesized that this may be due to GFP interfering with some important function or folding in this region of the protein. Supporting this hypothesis, a recent paper (Sandilos et. al., 2012) showed results indicating that the C-terminal tail of hPanx1 is folded up directly into the pore region, possibly keeping Panx1 in the closed conformation. If this topology is true for all homologous Panx and Inx channels, it

would make sense that having a GFP (29 kDa) hanging from the C-terminus would interfere with this function and could hold the pore in the "open" state, causing misfolding and cell death.

The pannexin-1 and pannexin-3 constructs were overall similar in FSEC profile, perhaps because their sequences are more highly homologous. However, many showed heterogeneity in their FSEC profiles – likely mixtures of hexamers, monomers, and smaller oligomers. N-chickenPanx3 and N-dogPanx3 (Figure 6) showed extremely high expression and monodisperse profiles, but were monomeric species. C-cowPanx3 also showed a very good oligomeric peak, though the polydispersity of its left shoulder makes it not as promising a target as N-frPanx1. These constructs may be useful if frPanx1 crystallization attempts fail.

Most of the pannexin-2 constructs were localized in puncta and showed mainly aggregation and heterogeneous species (similar to N-ratPanx2 in Figure 6). These data may support a study reporting that Panx2 does not form pores on its own, but instead forms heterogeneous Panx1/2 channels (Penuela et al., 2009). Perhaps in contrast to this, a recent EM study (Ambrosi et al., 2010) showed Panx2 (already the largest pannexin) forming homogenous pores that were described to be 7/8-mers, and therefore *much* larger than Panx1's hexamers. It is possible that the Panx2 constructs screened were so large that they were unable to be solubilized effectively with DDM.

Overall, the innexin constructs were mislocalized to puncta, poorly expressed, often killed HEK cells, and showed poor FSEC profiles (Figure 6). It is unclear why the vast majority of these proteins were not trafficked to the plasma membrane. Though the innexin proteins are natively from lower invertebrate species, it is common for cells from more highly-evolved organisms, such as HEK cells, to well-express lower proteins.

Confirmation that frPanx1 is functionally homologous to hPanx1

Since each gene tested in this screening was found in a BLAST search, it was necessary to verify that the proteins coded were indeed pannexin genes and showed pannexin-like channel function. Pannexin channels have been demonstrated to produce large outward currents when held at depolarizing membrane voltages in whole-cell patch clamp (Bruzzone et al., 2003). Thus, frPanx1 and hPanx1 proteins were transfected into HEK293 cells and measured using whole-cell patch clamp; cells were held at depolarizing membrane voltages and the resulting current measured. As shown in Figure 7, frPanx1 and hPanx1 show similar large outward currents that are inhibited by the pannexin/connexin inhibitor carbenoxolone. Thus, frPanx1 has been shown to be functionally similar to hPanx1 as well as a better crystallization target.

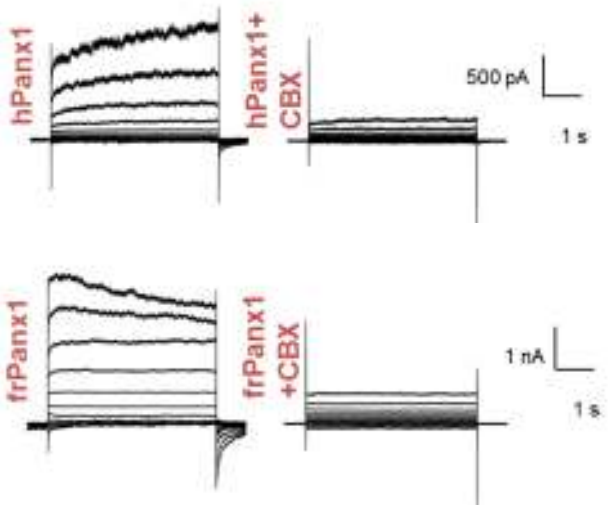

Figure 7: Whole-cell patch clamp experiments showing frPanx1 currents very similar to hPanx1 currents. Inhibition of currents by carbenoxolone (CBX) demonstrates that the identified frPanx1 channel has functional similarity to hPanx1.

4.2 Expression and Purification Protocols Established

Expression optimized for baculovirus infection of Sf9 insect cells

Membrane proteins often express in much lower quantity in recombinant systems than soluble proteins because of the plasma membrane surface area constraint. Thus, the choice of expression system is a critical one to maximize protein production. The best choices available are *E. Coli*, HEK293, and Sf9 cell expression because these cells can be grown in suspension and therefore amplified greatly to produce as much protein as possible. The *E. Coli* system was ruled out because eukaryotic membrane proteins rarely insert correctly into the prokaryotic double membrane of this gram negative bacterium. Further, an intact glycan may have been necessary to allow for correct expression and trafficking, suggesting that a eukaryotic expression system would be best. HEK293 cell expression was ruled out due to the high cost of producing and transfecting enough DNA for a large-scale culture. Thus, the baculovirus infection of Sf9 insect cells was chosen. These insect cells are known to somewhat mimic the glycosylation and trafficking pathways of mammalian cells, can be grown in culture easily, and baculovirus infection is cost efficient.

The frPanx1 gene was cloned into the pNGFP-FB2 vector, and transformed into DH10Bac *E. Coli* cells to produce the bacmid DNA. Bacmid was then transfected into Sf9 cells adhered to a 6-well plate and P1 virus was produced by secondary infection over 6 days. The P1 virus was harvested and filtered. A new suspension of Sf9 cells were grown and infected with P1 virus and grown for 1 day at 27°C then 2 days at 18°C. Finally, cells were harvested and lysed and pannexin expression evaluated by FSEC (Figure 8).

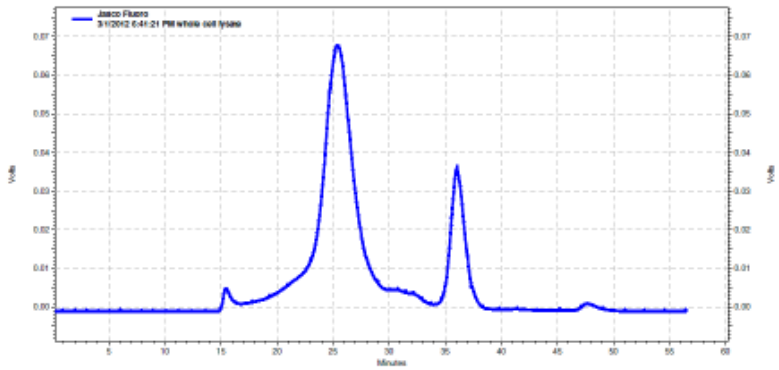

Figure 8: Whole-cell lysate of Sf9 cells infected with baculovirus expressing N-frPanx1. Appears that oligomer is well-expressed, correctly folded and assembled, and the oligomer solubilized by DDM

The FSEC trace of frPanx1 expressed in Sf9 cells was extremely similar the expression profile from HEK cells. The channel was indeed well-expressed in this cell system. The large oligomeric peak suggested that the channel was correctly-folded, well-solubilized, monodisperse, and may be able to be purified. Though there appeared to be more free-GFP cleaved, and since the affinity tag is on the GFP this would mean these proteins were unrecoverable, it is unclear whether the cleavage occurred within the cells or during lysis and therefore may be able to be protected.

Passage 1 virus (P1) was used to infect 50mL Sf9 cells, and grown for 3 days at 27°C, creating an amplified Passage 2 virus (P2). P2 virus is known potently infect Sf9 cells. Next, protein production was optimized for P2 infection of large scale culture. Growth conditions were optimized for temperature, time, and volume of virus. Cells were infected with 3 ratios of P2 virus to 50mL of cells and grown for 4 days. After 24 hours of growth at 27°C, cells were divided into two flasks – one set continued to be incubated at 27°C and the others were moved to 18°C. Cells were removed on days 2, 3, and 4 and the protein's profile checked by FSEC.

It was clear that the largest amount of monodisperse pannexin oligomer was produced when 50mL cells were infected with 500µL P2 virus. Thus, for large-scale culture, 1L Sf9 cells would

be infected with 10mL P2 virus. Nearly the same amount of protein was found in cells that were grown at 27°C for 2 days as cells that were grown at 18°C for 3 days. Subsequent proteins may have been grown in either of these conditions.

Purification procedures established

With optimized cell expression conditions, it was next important to learn how to purify the frPanx1 protein. A 1L Sf9 cell culture was grown and pannexin expressed using the previously determined conditions. Cells were harvested, washed, lysed by Avestin, and membranes were collected by ultracentrifugation.

DETERGENT SCREENING FOR MEMBRANE EXTRACTION

Membrane extraction conditions were screened using a panel of detergents and protein stability over time was assessed by FSEC. Aliquots of the 50µL of purified membrane were combined with 0.5-1% solubilization buffer and incubated at 4°C. As expected from initial ortholog screening, frPanx1 was initially well-solubilized by DDM (Figure 9, top left), but unexpectedly these proteins fall apart overnight even when kept at 4°C (Figure 9, top right). In contrast, LMNG, sodium cholate, CHAPS, DMNG (data not shown), and $C_{12}E_8$ (Figure 9, bottom) were all better than DDM at solubilization, prevention of aggregation, prevention of GFP cleavage, and prevention of dissociation over 8 days. Therefore, any of these proteins would have been an acceptable choice for extraction. $C_{12}E_8$ was chosen for future use due to its favorable price and ease of use (comes in liquid form).

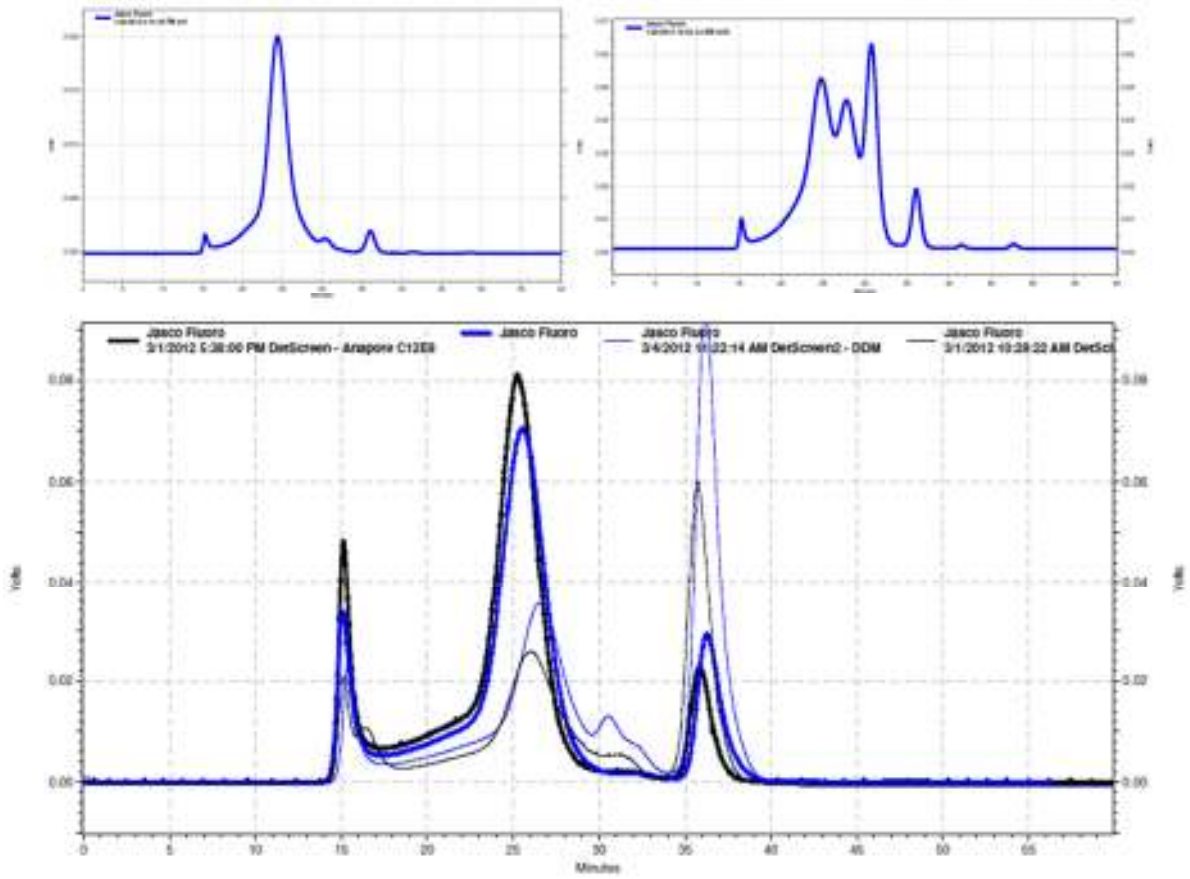

Figure 9: Detergent screen for membrane extraction. DDM solubilizes frPanx1 well initially (top left), but after 11 hours the protein dissociates (top right). Proteins solubilized in $C_{12}E_8$ are better preserved after 3 days (bold black line) and 8 days (bold blue line), as compared with DDM for 3 days (thin black line) and 8 days (thin blue line).

HIS-TAG AFFINITY PURIFICATION

After solubilization from the membrane, affinity purification was utilized to isolate only the recombinant protein. Membranes were solubilized and debris spun out (Figure 10, left). The 8xHis-tagged protein and solubilized membrane was incubated with TALON Co^{2+} resin in batch for 1 hour. Resin with bound protein was then collected and eluted with an imidazole gradient using the FPLC. Eluted protein was then checked by FSEC (Figure 10, right); a small amount of oligomer remained, but most had fallen apart into monomer and an intermediate species.

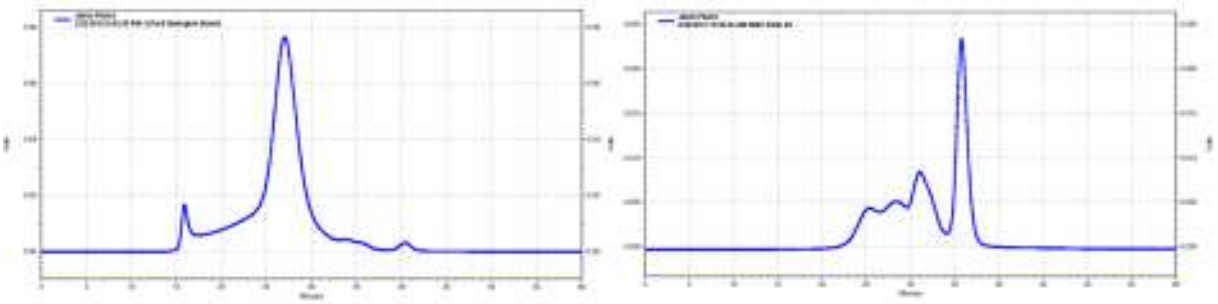

Figure 10: FSEC verification of protein stability throughout purification. The protein is intact after cell lysis, membrane isolation, and detergent solubilization (left), but dissociates after Co^{2+} resin affinity purification (right).

The oligomer dissociated during affinity purification, suggesting that the interfering molecule was either the Co^{2+}, the resin, or the imidazole. New protein was produced and membranes prepared. Solubilization buffers including relevant concentrations of $CoCl_2$ or imidazole were prepared and incubated with membrane overnight at 4°C. Figure 11 shows the resulting FSEC traces for control samples (thin black lines) and experimental samples (thick blue lines). The proteins incubated with $CoCl_2$ were not affected (Figure 11, left) while those incubated with imidazole were dissociated (Figure 11, right). It is unclear why imidazole causes pannexin to dissociate; we may speculate that perhaps histidine residues mediate the interactions between pannexin subunits, but regardless a new affinity protocol was necessary.

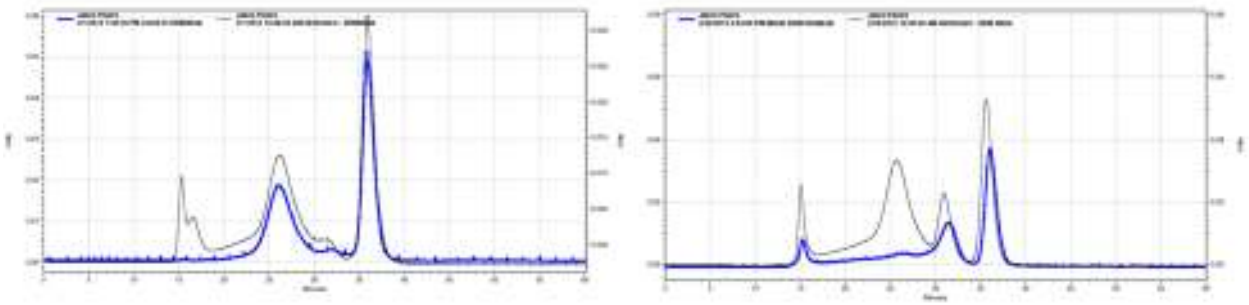

Figure 11: Imidazole caused pannexin oligomer to destabilize. FSEC traces showing overnight membrane solubilization with DDM only (black lines), DDM+$CoCl_2$ (blue line, left), and DDM+imidazole (blue line, right).

Low pH (4-5) is known to elute His-tagged proteins from a Ni^{2+} column, though it was unknown how pannexin would handle low pH. Pannexin was washed in standard PBS (pH 7.4) and then

eluted with an acetate buffer at pH 4. Because EGFP is not fluorescent in this pH range, the solution was neutralized before analysis by FSEC. Any sign of EGFP or pannexin had disappeared, suggesting that perhaps the protein had aggregated and was removed during ultracentrifugation. Though EDTA elution could have been attempted, an additional Co^{2+} removal step would be necessary. Thus, we changed the affinity purification step to use the StrepII-tag/Streptavidin system.

STREP-TAG AFFINITY PURIFICATION

The frPanx1 gene was subcloned into the pNGFP-FB3 vector which expressed the same protein construct with a StrepII-tag in place of the 8xHis tag. Bacmid, P1, and P2 viruses were made; proteins were expressed, purified, and solubilized in $C_{12}E_8$ as previously determined. The Strep-tagged protein was bound to Streptactin resin in batch for 1 hour and eluted with D-desthiobiotin. The sample was checked by FSEC and found to contain well-preserved protein oligomers (Figure 12, left).

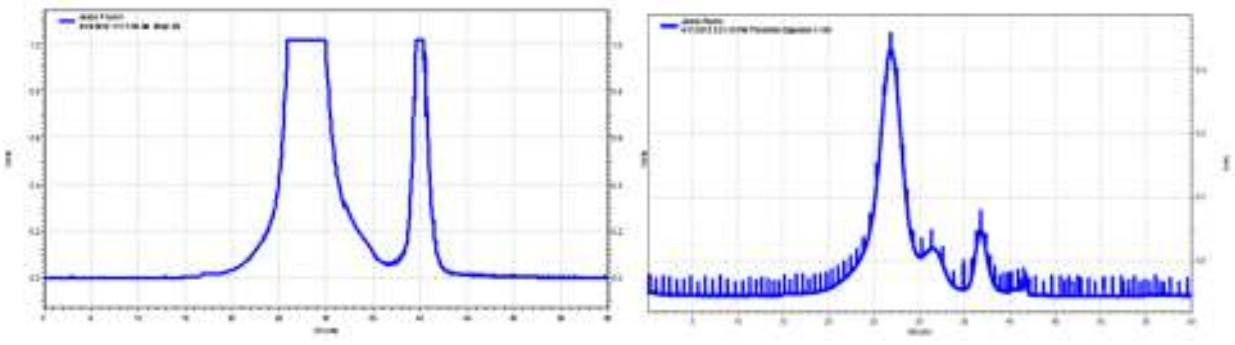

Figure 12: Pannexin survives desthiobiotin elution (left) and thrombin cleavage (right).

THROMBIN CLEAVAGE OF EGFP

Next, the EGFP and StrepII-tag were removed by thrombin cleavage. Thrombin digestion concentration was tested for overnight digestion at 4°C. 1:100 thrombin was sufficient to cleave all protein tags.

SIZE-EXCLUSION CHROMATOGRAPHY

Pannexin, thrombin, and EGFP were separated using a Superdex200 gel filtration column. This column was also used to exchange the protein into a new chemical environment that would be ideal for crystallography. A typical crystallization buffer (20mM HEPES) and low salt conditions (50mM NaCl) were chosen. Various crystallization detergents were chosen (DDM/LMNG), and the protein was stabilized in 15% glycerol. Though there was more aggregate and monomer than previously seen by FSEC, the oligomeric pannexin peak was indeed maintained through this last step of purification.

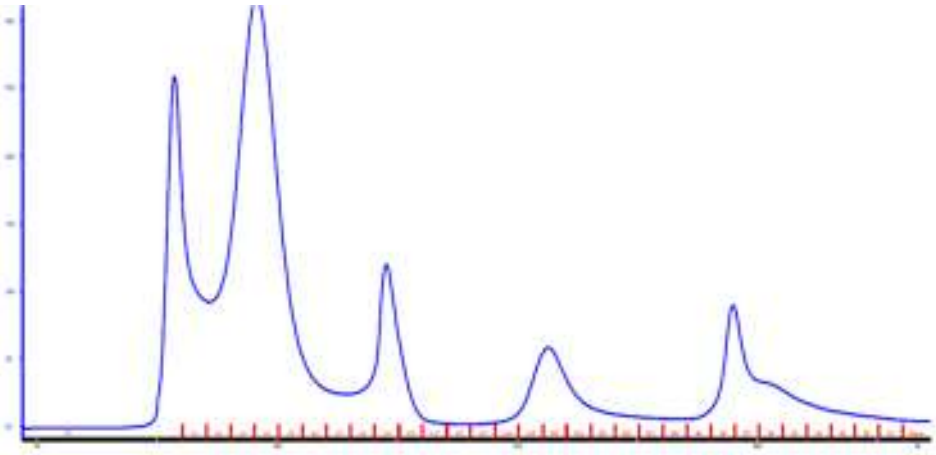

Figure 13: Size-exclusion chromatography of purified frPanx1 protein. Peaks are presumed to be (in order from left to right): aggregate, pannexin oligomer, pannexin monomer, thrombin/GFP, and degradation product.

Finally, samples were collected from the first 9 fractions (aggregate, oligomer, and monomer) and an SDS-PAGE gel was run to investigate the purity of the protein. When mixed with SDS and β-mercaptoethanol the oligomer is dissociated and the protein denatured, then resulting proteins are separated by size. In Figure 14, the pannexin monomer can be seen running as a doublet just above the 37kDa marker. Though the SEC trace suggests the oligomeric pannexin has been isolated, the SDS-PAGE gel suggests that the purified protein is extremely

heterogeneous. Next, it was necessary to address each heterogeneity to obtain a homogenous purified protein oligomer.

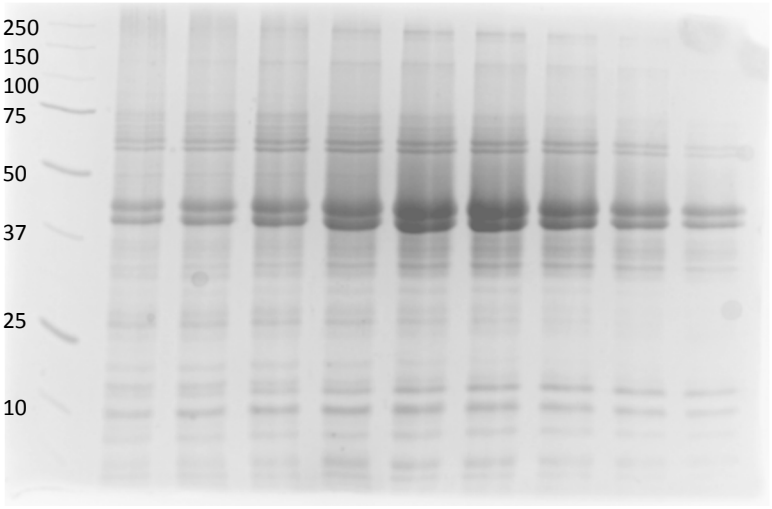

Figure 14: Purified frPanx1 protein. Each lane represents one successive fractional elution off the SEC column (shown in red in Figure 13). The protein is far from pure and therefore modifications to the protein and/or purification process must be made if there is hope for crystallization.

4.3 Pannexin Protein Modified to Obtain Protein Crystals

Through close analysis of the SDS-PAGE gel run after purification of NGFP-frPanx1 (Figure 14), hypotheses were made as to the identity of the unwanted species. After experimental verification, modified protein constructs and/or purification steps were added to increase the protein homogeneity.

A Western blot of the SEC fractions confirmed that the doublet running above the 37kDa marker was the frPanx1 monomer. An additional SDS-PAGE gel comparing the protein prior to thrombin digestion with those after SEC confirmed that the doublet running above the 50kDa marker was the GFP-tagged frPanx1 monomer (data not shown). In every digestion attempted, a small amount of this N-terminal GFP was incapable of being cleaved, suggesting its thrombin

cleavage site was inaccessible for some reason. The bands smaller than 37kDa were predicted to be degradation products.

N/C-terminal domain deletions reduce protein degradation

The degradation products were hypothesized to be regions of the C-terminal domain – pannexin's largest cytosolic protein region. If this area were nicked by enzymes during purification, the amino acids may remain a part of the oligomer until SDS-PAGE when peptides would be denatured and run separately from the protein core.

To test this theory, and determine whether the C-terminus was necessary for proper channel folding, 11 deletions of the C-terminus and 3 deletions of the much smaller N-terminus were designed up to the transmembrane helices and around regions of high sequence conservation amongst pannexin-1 species. Genes were cloned and screened in HEK cells using FSEC. Of the constructs tested, deletion of the N-terminus up to amino acid 22 (Figure 15, left) and deletion of the C-terminus at amino acid 357 (Figure 15, right) or beyond showed a well-behaved oligomeric protein in whole-cell lysate. Larger deletions of the termini tended to cause proteins to localize to puncta and show large aggregation and degradation. This suggests that this "core" protein of frPanx1(22-357) was necessary for oligomer folding, trafficking, and/or detergent stability, since their FSEC profile was similar to wild-type frPanx1.

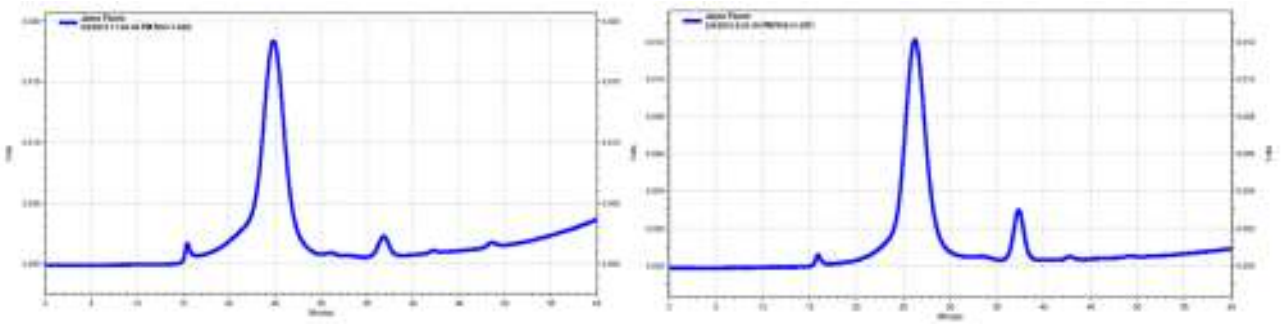

Figure 15: FSEC screening of N- and C-terminal deletion constructs. On the left, pNGFP-frPanx1(22-428) deletion of 21 AAs from the N-terminus, and on the right, pNGFP-frPanx1(1-357) deletion of 71 AAs from the C-terminus.

Since the N-terminus was already quite short and was already seen to have difficulty with GFP-cleavage, it was decided that only the C-terminus would be deleted. This construct was well-expressed in HEK cells and showed no cell death, as may have been expected if the C-terminal tail was indeed essential. Further, this construct was verified to still be functional by whole-cell patch clamp (Michalski, data not shown), making it a good crystallization candidate.

The frPanx1(1-357) gene was subcloned into pNGFP-FB3, virus produced, and protein purified as previously optimized. The SEC profile (Figure 16, left) showed a much more homogeneous separation of oligomer peak from other species. This protein was able to be purified with far less aggregate, no monomer (suggesting the oligomer is more stable), and no degradation product. This analysis was confirmed by SDS-PAGE gel (Figure 16, right) as the smaller degradation bands have disappeared.

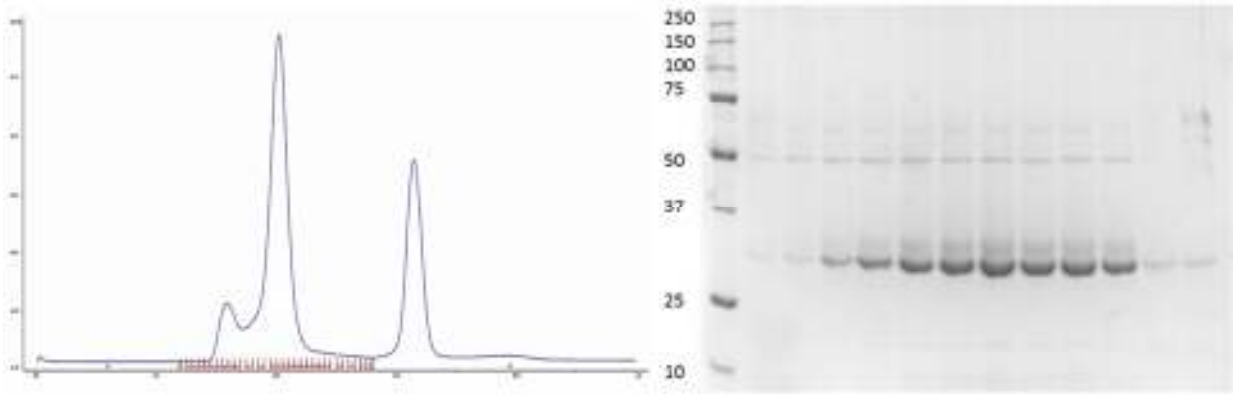

Figure 16: Purified frPanx1(1-357) SEC fractions. The oligomeric peak is well-resolved in fractions 5-8 and shows very little degradation. The pannexin monomer still runs as a much smaller doublet (under the 37kDa marker) and still shows some uncleaved GFP (around 50kDa).

Crystallization conditions were screened with this semi-pure protein, but no crystal hits were found. Thus, the protein was too inhomogeneous for crystal formation.

Glycan removal reduces heterogeneity

Pannexin-1 has four putative N-glycosylation consensus sites following the NXS/T motif, but only one of these sequences is on the predicted extracellular surface: at position 254 sequence NDS in hPanx1. This site was previously confirmed to be the only site of glycosylation and moreover was shown to be glycosylated heterogeneously (Penuela et al., 2007). The majority of pannexins show no glycan, while some others have high-mannose or more complex glycans as seen in hPanx1 whole-cell Western blots (Gehi, Shao, & Laird, 2011). Thus, it was hypothesized that heterogeneous glycosylation was causing the pannexin doublet seen in Figure 16. Though previous studies showed two different pannexin-1 glycoprotein species, and Sf9 cells are known to mimic mammalian glycosylation, overexpression or different insect cell machinery may have led to only a single glycoprotein.

Pannexin deletion frPanx1(1-357) was digested using PNGaseF in denaturing conditions and using EndoH in non-denaturing conditions (Figure 17). Both endoglycosidases were able to remove the top band in the doublet, identifying this as a glycosylated species, and collapsing the protein into the (presumably non-glycosylated) smaller band. Digestion conditions were optimized and found that an overnight digestion of 1:50 pannexin to EndoH (by mass) at 4°C was sufficient to remove the glycan (data not shown). Thus, future purifications involved glycan removal by EndoH digestion simultaneous to digestion with thrombin.

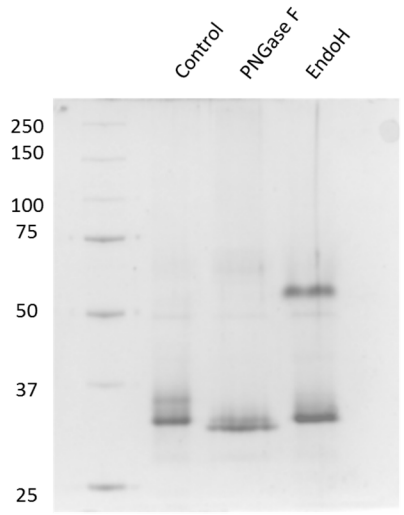

Figure 17: Heterogeneous protein glycosylation can be removed by endoglycosidases. Top band of pannexin doublet (control lane) must be a pannexin proteoglycan, as it can be removed with PNGaseF (denaturing conditions) or EndoH (non-denaturing conditions). Extra band observed in EndoH lane is probably EndoH itself, as this enzyme was not removed before running the gel.

Though digestion seemed to be effective, the most homogeneous way to produce pannexin protein would be without any glycan at all. Thus, we also tried mutating the asparagine in the single NDS motif to neutrally-, positively-, and negatively-charged serine, glutamate, and arginine residues, respectively. Small-scale Sf9 purifications were completed in tandem and

SEC fractions run on SDS-PAGE gel (Figure 18). Interestingly the charge-swapping glutamate mutation (N257E) showed the cleanest purification, though some residual NGFP was still observed. It is unclear why the N257S and N257R mutants showed degradation.

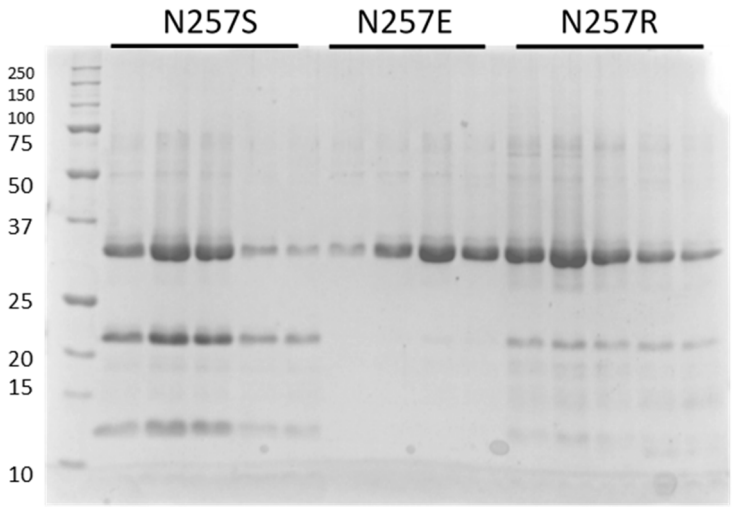

Figure 18: Glycosylation removal by genetic mutation. Asparagine in N-linked glycosylation motif NXS was mutated to serine (N257S), glutamate (N257E), and arginine (N257R). The glutamate-mutated protein shows less degradation.

Interestingly, a very slight doublet was still observed in all mutated proteins in Figure 18, and may even be visible in the endoglycosidase-treated proteins in Figure 17. Because the glycans have clearly been removed, this must be an additional form of protein heterogeneity.

These bands were further separated, extracted from the gel, and analyzed by mass spectroscopy. Both bands showed identical amino acid sequences, suggesting this heterogeneity was perhaps conformational rather than chemical. Though protein conformations are usually unable to be observed in denaturing SDS conditions, it is clear that these membrane proteins still retain some secondary or tertiary structure because they run much faster through the gel than their size would permit. Full-length frPanx1 is 49kDa, but always runs between 37-50kDa markers (Figure 14) while the 41kDa frPanx1(1-357) deletion runs between 25-37kDa. Thus, this doublet observed

in the absence of glycosylation may be indicating a conformational heterogeneity within the pannexin protein. Though some heterogeneities still existed in the modified proteins, they were indeed pure enough to form crystals.

Modified proteins were able to form crystals

Crystallization detergents were screened using dialysis and found that the protein is stable in LMNG, DMNG, C12E8, and CHAPS over a week, maybe longer. The frPanx1(1-357)(N257E) and the EndoH-treated frPanx1(1-357) proteins were switched into these detergents over the size-exclusion column, concentrated, and subjected to crystallization screening using sparse matrix screens with the hanging drop method.

Four hits, all with the EndoH-treated protein, were observed to grow into tiny crystals which were verified to be protein crystals by UV fluorescence. Each hit condition was then optimized by screening pH, PEG type and PEG concentration. Hits were reproducible (Figure 19, left) but showed huge amounts of precipitate and inconsistent growth trends in identical conditions with protein from successive preparations. Small crystals were able to be grown to 50-70μm. These crystals were broken into single rods and cryoprotected in mother liquor with 15% glycerol. The X-ray beam at the Cornell High Energy Synchrotron Source showed weak diffraction of crystals in various conditions and detergents to 20-30Å resolution. The best crystals (Figure 19, center) diffracted to 15Å resolution (Figure 19, right).

Figure 19: Crystals of frPanx1 deletion. Initial hits (left) led to optimized crystals (center) which weakly diffracted to 12Å (right).

With so few spots, it was impossible to use the diffraction patterns for very much, but HKL2000 was able to estimate a P4 crystal symmetry with unit cell geometry of 191.1Å x 191.1Å x 133.28Å. Previous EM studies (Ambrosi et al., 2010) suggest Panx1 to have a diameter of 120-160 Å, indicating one protein molecule within each unit cell with 70% solvent content.

4.4 Pannexin Crystal Quality Improvements

The ability to form protein crystals and their weak but observable diffraction was extremely promising. These observations made it clear that the pannexin-1 protein was indeed stable and well-behaved *in vitro* and relatively homogeneous. But poor crystal reproducibility, high solvent content, and weak diffraction suggested that further modification was necessary to either improve crystal formation conditions, improve the homogeneity of protein purification, or reduce flexibility within the protein itself before tight, well-packed protein crystals could be produced.

Crystallization modifications

Crystallization additives were screened to see if one could be found to improve crystal quality. Most additives showed no change, while $CaCl_2$ and 1,5-diaminopentane-DiHCl produced

smaller crystals with slightly different morphology. A small improvement in crystal formation, though not diffraction quality, was found when protein-detergent complexes were supplemented with a lipid mixture of 3:1:1 POPE:POPG:cholesterol (m/m).

Dehydration of crystals in mother liquor with increasing concentrations of PEG (13% to 25%) was able to improve diffraction. Crystals were scooped to each successive solution and incubated for 30 minutes before moved to the next. Optimum dehydration occurred at 22.5% PEG, improving crystal diffraction to 11-14 Å resolution.

With dehydration showing improved crystal packing, high-pressure cryocooling (Kim, Kapfer, & Gruner, 2005) at 200 MPa was attempted. Crystals were scooped into loops but were unable to remain hydrated long enough to manipulate and pressurize the systems. It may be worth trying this method again scooping crystals into oil to prevent dehydration.

Finally, small improvements with the lipid mixture suggested that perhaps a lipidic environment would help crystallize the protein. Neither crystallization in lipidic cubic phase (Landau & Rosenbusch, 1996), nor hanging drop crystallization with proteins in lipid/detergent bicelles (Faham & Bowie, 2002) were sufficient to grow crystals.

In any effect, though these manipulations could do some to improve diffraction, it was ultimately the protein itself that was not packing well. Further attempts to improve protein homogeneity and stability needed occur.

Tagging position switch removes GFP tag heterogeneity

For some reason, in purification after purification, the N-terminal GFP tag is unable to be completely cleaved. One attempt to solve this problem involved removing the GFP entirely,

switching to the pNNG-FB3 expression vector. However, this protein was found to be completely aggregated, suggesting the GFP may help keep the protein soluble through purification.

When the C-terminal deletion was made, pannexin channel currents were measured by whole-cell patch clamp and found to be functional. This proved to us that the C-terminus does not modulate channel function as previously suggested (Sandilos et al., 2012). Though our C-terminal GFP tagged pannexin/innexin constructs led to cell death in HEK cells, the similarly tagged frPanx1(1-357) deletion construct did not.

Thus, it made sense to try switching the GFP tag to the CTD to see if it would more be more completely cleaved during thrombin digestion (pCGFP-FB3 vector used). Indeed, subsequent purifications showed that every time the GFP was cleaved completely. Figure 20, left shows the affinity-purified protein in lane 1 and the thrombin-digested SEC-purified peak fractions in lanes 2-9.

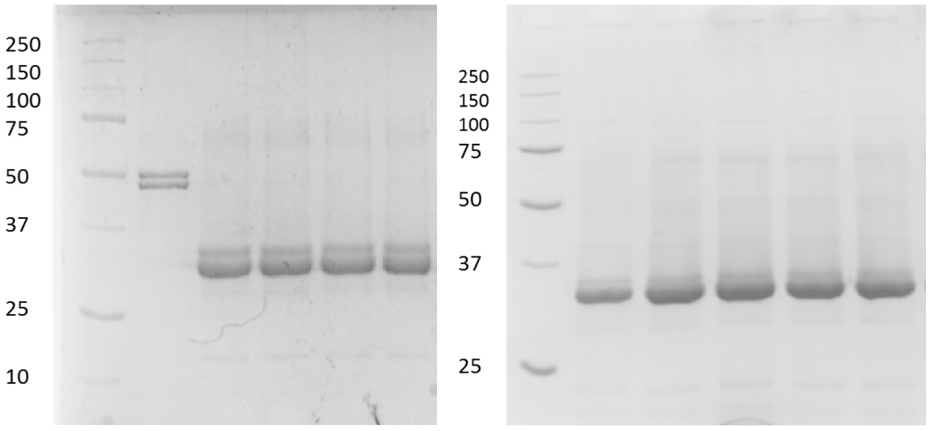

Figure 20: C-terminal tagging with GFP-Strep (left) or Strep tag only (right) leads purer protein. Left: purified pannexin-GFP is shown in lane 1. The tag was cleaved by thrombin and purified by gel filtration to produce the peak fractions shown in lanes 2-9. Right: Protein requires no tag digestion and is purified with the Strep tag intact.

Further, the pCNG-FB7 construct was designed for C-terminal expression of Strep tag (without thrombin cleavage site), presuming this short sequence might not hinder crystal growth. Pannexin was found to be stable in this construct without GFP throughout purification when 15% glycerol was included in all buffers.

Overall, purification of C-terminally tagged frPanx1(1-357) proteins increased the purity and decreased the level of aggregation. More consistency was seen in peak shape, further indicating the better monodispersity of this protein sample. Many more crystal hits were obtained than with the previous N-terminally tagged constructs, and tended to be more easily reproduced. However, the best crystals were not able to grow very large and the best diffraction observed was to 16Å resolution.

Trypsin digestion elucidates flexible loop

This extremely pure protein (chemically, but perhaps not conformationally) was still unable to form well-ordered crystals. Thus, it must be the protein itself that is preventing good crystal formation. One possibility is that the detergent micelle is too large (DDM and LMNG both have 12-carbon chains) and very little of the protein is accessible to make crystal contacts. Another possibility is that much of the accessible protein is disordered and this flexibility makes it impossible to pack tightly.

A trypsin digestion of the purified protein was performed in non-denaturing conditions to identify regions of the protein likely to be part of a disordered loop. Trypsin is a serine protease that cleaves proteins at lysine or arginine residues. When proteins are denatured, trypsin is able to access nearly every positively-charged residue and cleave the peptide bond. But in non-

denaturing conditions, all protein structure remains intact and trypsin is only able to access unstructured residues.

Purified frPanx1(1-357) was incubated with a range of trypsin concentrations in normal crystallization buffer – at high enough concentration, the protein should be completely degraded. Interestingly, at even the lowest concentration tested (1:10,000 trypsin by mass), and then persisting into pretty high concentrations (1:500 trypsin by mass), the protein was cut into only two major species (Figure 21, left), suggesting a single flexible region of the protein.

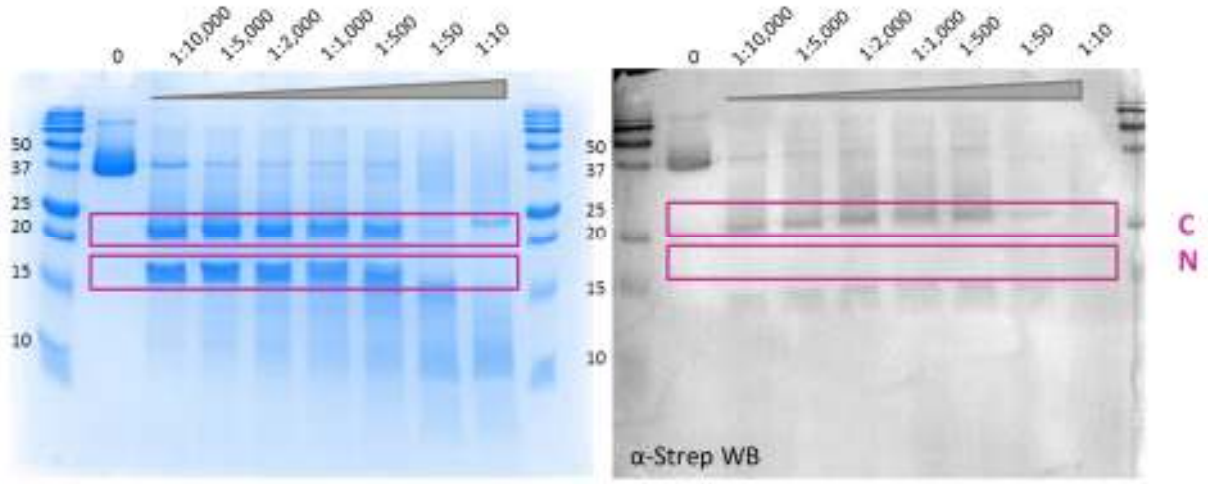

Figure 21: Trypsin digestion of frPanx1(1-357) reveals disordered region in intracellular loop. When digested proteins were run on SDS-PAGE and stained with coomassie (left), two distinct protein bands were seen, even at very low trypsin concentrations, suggesting a single trypsin-accessible protein region. A Western blot was performed with α-StrepII-tag antibody, identifying the N- and C-terminal fragments.

Western blotting for the C-terminal StrepII-tag (Figure 21, right) identified the top band as the C-terminal pannexin fragment (~22 kDa in length) and the lower band presumably as the N-terminal fragment (~16 kDa in length). Fragment sizes in comparison with predicted trypsin sites allowed me to identify the pannexin intracellular loop to be the area of cleavage, at one or many of the trypsin recognition sites.

The sum of the fragment sizes (~38kDa) fell just short of the expressed construct (41kDa), likely due to some cleavage of the C-terminus (faint degradation products observable below the N-terminal fragment, Figure 21, right). Perhaps shortening the end of the C-terminus slightly would help, but this is probably not the biggest problem.

The primary amino acid sequence of the pannexin-1 protein is extremely well-conserved between species (Appendix C), especially in the 1-357 range. However, the pannexin-1 intracellular loop is an exception to this trend. A large region of this loop shows amino acid mutations, charge-swapping, deletions, and insertions when compared across species. If this loop is indeed a disordered region that is not involved in structural integrity, this may explain why the area is poorly-conserved through evolution. This flexibility may indeed obstruct crystallization.

Modifications to the intracellular loop

PANEL OF LOOP DELETION CONSTRUCTS

First, it was necessary to identify which parts of the intracellular loop were essential and which were not. 6 deletions of varying length were made to shorten the loop. Two extreme deletions (88 and 50 AAs) were made that deleted the entire loop or the region of lowest sequence conservation, respectively. Four modest deletions (~15 AAs) to the region of poorest conservation were made using a sliding window technique to identify the essential residues.

FSEC screening showed the two proteins with drastic deletions were misassembled, but the other four deletions appeared similar to wild-type. This suggested there were many residues throughout the middle of the loop that could be deleted, but the residues flanking the transmembrane helices needed to be conserved.

A second set of loop deletions were made varying regions and lengths of deletion within the poorly-conserved region of the loop. By FSEC analysis, all proteins were assembled correctly, eluting at nearly the same time as the wild-type protein (Figure 22). Varying levels of expression were observed. The two shortest deletions (13 and 17 AAs) were more highly expressed than wild-type, but many of the constructs were the same or smaller.

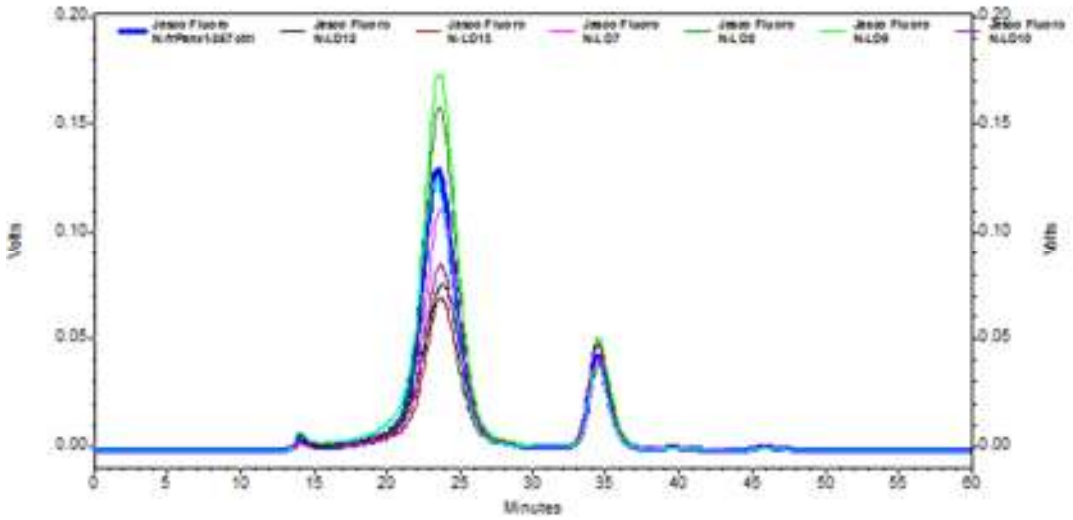

Figure 22: Pannexin intracellular loop deletions screened by FSEC. The proteins with shorter intracellular loops (thin lines) are well-folded and assembled in comparison with wild-type (thick blue line). Varied levels of cell expression are seen.

T4-LYSOZYME FUSIONS

The identification of this flexible loop suggests a location within pannexin to insert a crystallization chaperone like T4-lysozyme as was successfully shown with some G-protein coupled receptors (Cherezov et al., 2007). By replacing the disordered loop with this easily-crystallized protein, there is more protein surface area for protein-protein interactions to form a crystal.

I obtained the non-optimized T4 lysozyme DNA sequence and cloned this into the loop in wild-type pannexin and into the region deleted in 5 selected loop deletion constructs. Cells expressed very little of this protein, and much of what was present tended to aggregate, as analyzed by

FSEC. However, the protein does show a visible peak at the predicted size, indicating that pannexin is still correctly-folded and oligomerized despite expressing an entire 14 kDa protein within the intracellular loop of each subunit of the channel as well as a GFP tag on each C-terminus. The best two constructs from the screening in Figure 23, left in dark blue and green, were the constructs with the largest deletion.

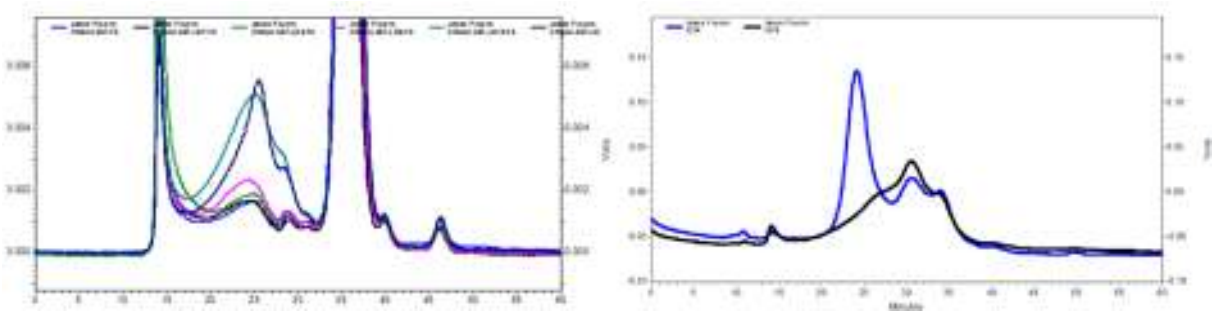

Figure 23: T4 lysozyme fusion protein screening. Fusion proteins inserting T4 lysozyme into the intracellular loop with varying lengths of flanking region were screened (left). All proteins showed correct pannexin oligomerization (peak eluted ~25min). The best two (green and dark blue on left) were purified from insect cells (blue and black on right, respectively) and just one found to be well-folded.

These two best-expressing constructs were chosen for purification; the DNAs were subcloned into the non-GFP vector pCNG-FB7 so as allow more room for lysozyme and perhaps increase expression level. Only one construct (the Δ152-190+T4 shown in blue in Figure 23, right) was indeed able to be purified, though some instability was visible due to the presence of monomers. However, the expression level of this construct was far worse than wild-type and therefore not a cost-effective crystallization target.

LOOP DELETIONS

Because there were too many successful loop deletion constructs to test, three proteins of short (Δ174-194), medium (Δ161-190), and long (Δ152-190) deletions within the same loop area were chosen. The largest deletion showed an aggregated oligomer peak upon purification, while the mid-sized deletion (identical deletion but with 9 additional AAs) eluted in the expected volume.

These proteins were screened for crystallization conditions extensively, but no hits were observed. The largest deletion construct was subjected to trypsin digestion and was indeed found to be trypsin-resistant (data not shown), verifying that it was this poorly-conserved region of the intracellular loop where trypsin was able to cut.

The shortest deletion, frPanx1(1-357)(Δ174-194) missing 20 AAs on the TM3-end of the intracellular loop, was able to be expressed and purified with zero aggregation, a feat no pannexin had thus far accomplished. Further, this protein showed nearly 1.5X higher expression than the wild-type frPanx1(1-357) in Sf9 cells and its C-terminal GFP tag was perfectly cleaved (Figure 24, top left). And surprisingly, the conformational heterogeneity seen in all previously-purified proteins no longer exists. This is the most homogeneous protein yet purified. Crystal hits were found with protein in LMNG and a mixture of NG+CHAPS, but optimization was unable to grow the crystals much larger, and they did not diffract at all.

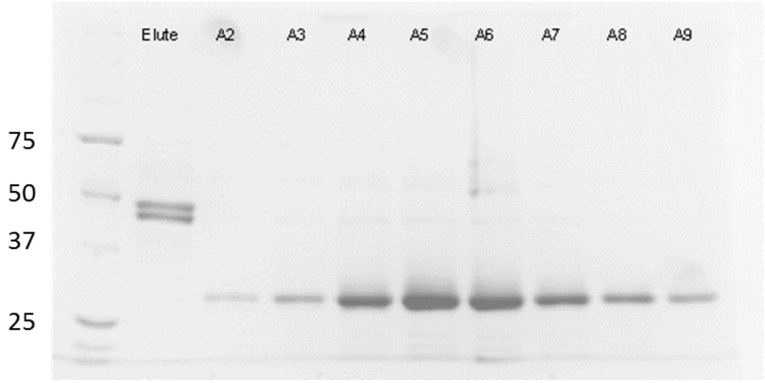

Figure 24: Purification of best loop deletion, frPanx1(1-357)(Δ174-194). The C-terminal GFP tag and glycan (two protein species in Elute lane) are cleaved to high purity and the gel filtration fractions of the oligomeric peak shown (A2-A9).

But the high purity and monodispersity of this protein led to further crystallization attempts, ultimately showing success when crystallized in bicelles. Bicelles are a lipidic medium composed of a long-chain lipid and a short-chain detergent that when mixed with membrane

proteins assemble into structures that closely resemble discs of lipid bilayer (Faham & Bowie, 2002). In contrast to detergents whose micelles can be quite large and may occlude the soluble regions of the protein, bicelles better mimic the plasma membrane making it easier for soluble portions of the protein to form crystal contacts.

Bicelles with long-lipid DMPC with short-chain detergent CHAPSO in a 3:1 molar ratio were produced as suggested by Ujwal and Bowie (2011), and mixed with protein in a 1:4 (v/v) ratio. After incubation, the protein-bicelle mixture was used in standard hanging-drop crystallization screening.

Crystal hits abounded and many were able to be reproduced easily. Tiny crystals appeared almost overnight, while larger crystals formed after 2-3 weeks. These larger crystals (Figure 25, left) were oddly-shaped and extremely brittle, but easily optimized to grow as large as 300μm. The best crystals diffracted to 9Å and showed characteristic bicelle diffraction rings (Figure 25, right).

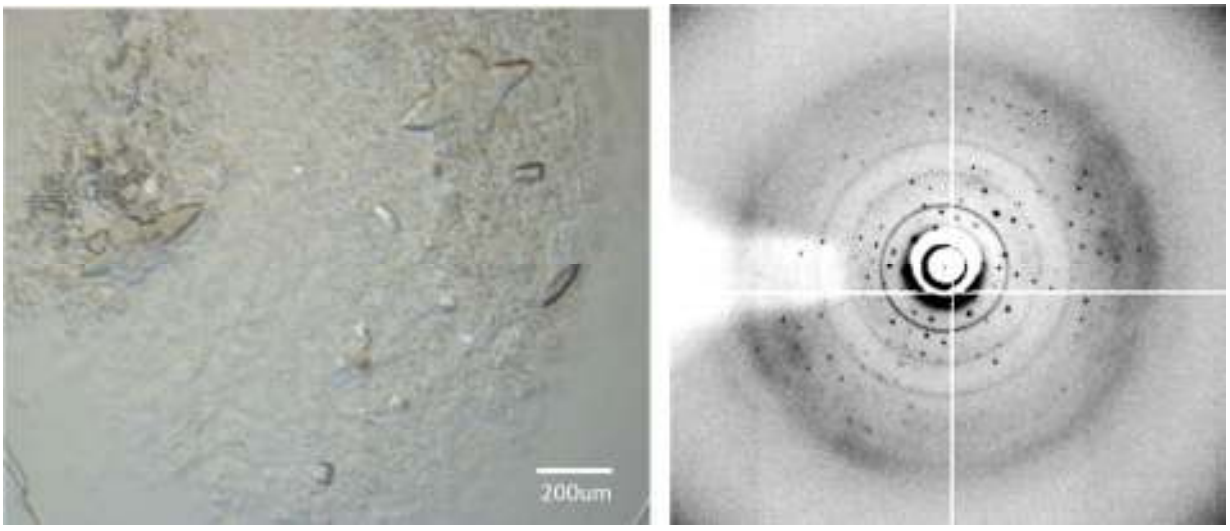

Figure 25: Protein-bicelle crystallization of frPanx1(1-357)(Δ174-194). These best crystals showed odd geometry, but were able to be grown as large as 300μm (left) and diffract to 9Å resolution (right).

CHAPTER 5

CONCLUSION

5.1 Summary of Important Findings

Pannexins are large-pore forming channels that mediate non-selective transport of ATP during purinergic signaling events important for cell function of the immune system, central nervous system, and apoptosis. Though studies have characterized pannexin malfunction as a key player in various diseases, therapeutics have been challenging to develop due to the absence of a high-resolution structure.

In this work, steps are made towards the goal of structure determination. The pannexin-1 protein from *Xenopus tropicalis* was identified as the most chemically stable pannexin or innexin protein, and thus was chosen as the crystallization candidate. Expression conditions were established to optimize protein production in Sf9 insect cells using the baculovirus infection system. Chemical conditions for purification were established. Imidazole was found to dissociate the protein, necessitating the use of the StrepII-tag for affinity purification.

Important modifications to the protein were made to enhance purity and crystallizability. The relatively unconserved regions of the protein, namely the intracellular loop and the C-terminal domain, were found to be unnecessary for pannexin channel assembly and function. Deletion of the C-terminal domain led to less degradation throughout purification and thus a purer protein. Deletions of the intracellular loop improved the likelihood of crystal formation. Finally, bicelle crystallization of this improved protein construct enhanced diffraction quality to 9Å resolution, a huge improvement over initial 20Å diffraction.

5.2 Future Directions

This work lays the foundation for future X-ray crystallography studies of the pannexin-1 channel. Certainly, the most immediate barrier towards this goal is improvement of protein and crystal quality.

In regards to this task, it may be worth making further modification to the intracellular loop region. Slightly shortening or lengthening around the Δ174-194 deletion may lead to slightly better crystal packing and diffraction. Other loop deletions were shown to be functional purification and crystallization was never attempted. It would likely be worth attempting to crystallize the shorter deletions (~20 AAs) of the N-terminal section of the non-conserved region to contrast the successful 9Å crystals found with the short deletion of the C-terminal section.

With the success of the lipid-like protein environment afforded by bicelle crystallization, it would be prudent to also attempt crystallization within lipidic cubic phase. This somewhat challenging technique involves reconstituting pannexin into lipids in a highly viscous bicontinuous cubic phase. In this phase, formed by monoolein or monopalmitolein and water, lipids pack into a lattice with interpenetrating aqueous channels (Landau & Rosenbusch, 1996). Because of the lateral protein mobility throughout the lipid lattice, more of the protein becomes accessible to form crystal contacts. This may be an important method for pannexin since it does not have a large soluble domain to mediate crystal packing.

Identification of this flexible loop also provides a target for antibody co-crystallization. Originally used to crystallize the β_2 adrenergic receptor (Day et al., 2007), this technique involves raising antibodies, purifying the Fab fragment, and mixing protein with Fab in the crystallization drop. This method has been shown to be effective for both stabilizing disordered

regions of membrane proteins as well as providing an improved surface for crystal contacts to form.

Improvement will be made with systematic modifications to the expression construct until the protein is coaxed into forming well-diffracting crystals. The variety of physiological roles and proposed mechanisms of action suggest the pannexin structure will reveal a novel channel architecture. A high-resolution atomic structure will open doors that currently limit the studies in this family of proteins, and perhaps facilitate the development of therapeutics against a wide variety of diseases.

REFERENCES

Abbracchio, M. P., Burnstock, G., Verkhratsky, A., & Zimmermann, H. (2009). Purinergic signalling in the nervous system: an overview. *Trends in Neurosciences*, *32*(1), 19–29. doi:10.1016/j.tins.2008.10.001

Ambrosi, C., Gassmann, O., Pranskevich, J. N., Boassa, D., Smock, A., Wang, J., … Sosinsky, G. E. (2010). Pannexin1 and Pannexin2 channels show quaternary similarities to connexons and different oligomerization numbers from each other. *The Journal of Biological Chemistry*, *285*(32), 24420–31. doi:10.1074/jbc.M110.115444

Bao, L., Locovei, S., & Dahl, G. (2004). Pannexin membrane channels are mechanosensitive conduits for ATP. *FEBS Letters*, *572*(1-3), 65–8. doi:10.1016/j.febslet.2004.07.009

Baranova, A., Ivanov, D., Petrash, N., Pestova, A., Skoblov, M., Kelmanson, I., … Panchin, Y. (2004). The mammalian pannexin family is homologous to the invertebrate innexin gap junction proteins. *Genomics*, *83*(4), 706–16. doi:10.1016/j.ygeno.2003.09.025

Brenner, S. E. (2001). A tour of structural genomics, *2*(October), 1–9.

Bruzzone, R., Hormuzdi, S. G., Barbe, M. T., Herb, A., & Monyer, H. (2003). Pannexins, a family of gap junction proteins expressed in brain. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(23), 13644–9. doi:10.1073/pnas.2233464100

Carpenter, E. P., Beis, K., Cameron, A. D., & Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*, *18*(5), 581–6. doi:10.1016/j.sbi.2008.07.001

Chekeni, F. B., Elliott, M. R., Sandilos, J. K., Walk, S. F., Kinchen, J. M., Lazarowski, E. R., … Ravichandran, K. S. (2010). Pannexin 1 channels mediate “find-me” signal release and membrane permeability during apoptosis. *Nature*, *467*(7317), 863–7. doi:10.1038/nature09413

Chen, Y., Yao, Y., Sumi, Y., Li, A., To, U. K., Elkhal, A., … Junger, W. G. (2010). Purinergic signaling: a fundamental mechanism in neutrophil activation. *Science Signaling*, *3*(125), ra45. doi:10.1126/scisignal.2000549

Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., … Stevens, R. C. (2007). High-Resolution Crystal Structure of an Engineered Human b2 -Adrenergic G Protein-Coupled Receptor. *Science*, *318*(November), 1258–1265.

Dahl, G., & Keane, R. W. (2012). Pannexin: from discovery to bedside in 11±4 years? *Brain Research*, *1487*, 150–9. doi:10.1016/j.brainres.2012.04.058

Day, P. W., Rasmussen, S. G. F., Masood, A., Parnot, C., Jose, J., Kobilka, T. S., … Kobilka, B. K. (2007). A monoclonal antibody for G protein – coupled receptor crystallography. *Nature Methods*, *4*(11), 927–929. doi:10.1038/NMETH1112

Ducruix, A., & Giege, R. (1992). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. New York: Oxford University Press.

Elliott, M. R., Chekeni, F. B., Trampont, P. C., Lazarowski, E. R., Kadl, A., Walk, S. F., … Ravichandran, K. S. (2009). Nucleotides released by apoptotic cells act as a find-me signal to promote phagocytic clearance. *Nature*, *461*(7261), 282–6. doi:10.1038/nature08296

Erlichman, J. S., Leiter, J. C., & Gourine, A. V. (2010). ATP, glia and central respiratory control. *Respiratory Physiology & Neurobiology*, *173*(3), 305–11. doi:10.1016/j.resp.2010.06.009

Faham, S., & Bowie, J. U. (2002). Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of Molecular Biology*, *316*(1), 1–6. doi:10.1006/jmbi.2001.5295

Gehi, R., Shao, Q., & Laird, D. W. (2011). Pathways regulating the trafficking and turnover of pannexin1 protein and the role of the C-terminal domain. *The Journal of Biological Chemistry*, *286*(31), 27639–53. doi:10.1074/jbc.M111.260711

Iadecola, C., & Nedergaard, M. (2007). Glial regulation of the cerebral microvasculature. *Nature Neuroscience*, *10*(11), 1369–76. doi:10.1038/nn2003

Iglesias, R., Dahl, G., Qiu, F., Spray, D. C., & Scemes, E. (2009). Pannexin 1: the molecular substrate of astrocyte "hemichannels". *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(21), 7092–7. doi:10.1523/JNEUROSCI.6062-08.2009

Iwamoto, T., Nakamura, T., Doyle, A., Ishikawa, M., de Vega, S., Fukumoto, S., & Yamada, Y. (2010). Pannexin 3 regulates intracellular ATP/cAMP levels and promotes chondrocyte differentiation. *The Journal of Biological Chemistry*, *285*(24), 18948–58. doi:10.1074/jbc.M110.127027

James, G., & Butt, A. M. (2002). P2Y and P2X purinoceptor mediated Ca2+ signalling in glial cell pathology in the central nervous system. *European Journal of Pharmacology*, *447*(2-3), 247–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12151016

Joseph, S. M., Buchakjian, M. R., & Dubyak, G. R. (2003). Colocalization of ATP release sites and ecto-ATPase activity at the extracellular surface of human astrocytes. *The Journal of Biological Chemistry*, *278*(26), 23331–42. doi:10.1074/jbc.M302680200

Junger, W. G. (2011). Immune cell regulation by autocrine purinergic signalling. *Nature Reviews. Immunology*, *11*(3), 201–12. doi:10.1038/nri2938

Kim, C. U., Kapfer, R., & Gruner, S. M. (2005). High-pressure cooling of protein crystals without cryoprotectants. *Acta Crystallographica. Section D, Biological Crystallography*, *61*(Pt 7), 881–90. doi:10.1107/S090744490500836X

Kimelberg, H. K. (2010). Functions of mature mammalian astrocytes: a current view. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *16*(1), 79–106. doi:10.1177/1073858409342593

Landau, E. M., & Rosenbusch, J. P. (1996). Lipidic cubic phases : A novel concept for the crystallization. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(December), 14532–14535.

Lazarowski, E. R., Boucher, R. C., & Harden, T. K. (2003). Mechanisms of release of nucleotides and integration of their action as P2X- and P2Y-receptor activating molecules. *Molecular Pharmacology*, *64*(4), 785–95. doi:10.1124/mol.64.4.785

Locovei, S., Scemes, E., Qiu, F., Spray, D. C., & Dahl, G. (2007). Pannexin1 is part of the pore forming unit of the P2X(7) receptor death complex. *FEBS Letters*, *581*(3), 483–8. doi:10.1016/j.febslet.2006.12.056

Locovei, S., Wang, J., & Dahl, G. (2006). Activation of pannexin 1 channels by ATP through P2Y receptors and by cytoplasmic calcium. *FEBS Letters*, *580*(1), 239–44. doi:10.1016/j.febslet.2005.12.004

MacVicar, B. A., & Thompson, R. J. (2010). Non-junction functions of pannexin-1 channels. *Trends in Neurosciences*, *33*(2), 93–102. doi:10.1016/j.tins.2009.11.007

Maeda, S., Nakagawa, S., Suga, M., Yamashita, E., Oshima, A., Fujiyoshi, Y., & Tsukihara, T. (2009). Structure of the connexin 26 gap junction channel at 3.5 A resolution. *Nature*, *458*(7238), 597–602. doi:10.1038/nature07869

Nelson, D. L., & Cox, M. M. (2008). *Lehninger Principles of Biochemistry* (Fifth.). New York, NY: W. H. Freeman and Company.

Nussler, A. K., Wittel, U. A., Nussler, N. C., & Beger, H. G. (1999). Leukocytes, the Janus cells in inflammatory disease. *Langenbeck's Archives of Surgery / Deutsche Gesellschaft Für Chirurgie*, *384*(2), 222–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10328179

Pelegrin, P., & Surprenant, A. (2006). Pannexin-1 mediates large pore formation and interleukin-1beta release by the ATP-gated P2X7 receptor. *The EMBO Journal*, *25*(21), 5071–82. doi:10.1038/sj.emboj.7601378

Penuela, S., Bhalla, R., Gong, X.-Q., Cowan, K. N., Celetti, S. J., Cowan, B. J., … Laird, D. W. (2007). Pannexin 1 and pannexin 3 are glycoproteins that exhibit many distinct characteristics from the connexin family of gap junction proteins. *Journal of Cell Science*, *120*(Pt 21), 3772–83. doi:10.1242/jcs.009514

Penuela, S., Bhalla, R., Nag, K., & Laird, D. W. (2009). Glycosylation Regulates Pannexin Intermixing and Cellular Localization, *20*, 4313–4323. doi:10.1091/mbc.E09

Penuela, S., Gehi, R., & Laird, D. W. (2012). The biochemistry and function of pannexin channels. *Biochimica et Biophysica Acta*, *1828*(1), 15–22. doi:10.1016/j.bbamem.2012.01.017

Praetorius, H. a, & Leipziger, J. (2009). ATP release from non-excitable cells. *Purinergic Signalling*, *5*(4), 433–46. doi:10.1007/s11302-009-9146-2

Qiu, F., & Dahl, G. (2009). A permeant regulating its permeation pore: inhibition of pannexin 1 channels by ATP. *American Journal of Physiology. Cell Physiology*, *296*(2), C250–5. doi:10.1152/ajpcell.00433.2008

Qiu, F., Wang, J., Spray, D. C., Scemes, E., & Dahl, G. (2011). Two non-vesicular ATP release pathways in the mouse erythrocyte membrane. *FEBS Letters*, *585*(21), 3430–5. doi:10.1016/j.febslet.2011.09.033

Ravichandran, K. S., & Lorenz, U. (2007). Engulfment of apoptotic cells: signals for a good meal. *Nature Reviews. Immunology*, *7*(12), 964–74. doi:10.1038/nri2214

Robson, S. C., Sévigny, J., & Zimmermann, H. (2006). The E-NTPDase family of ectonucleotidases: Structure function relationships and pathophysiological significance. *Purinergic Signalling*, *2*(2), 409–30. doi:10.1007/s11302-006-9003-5

Rupp, B. (2010). *Biomolecular Crystallography*. New York: Garland Science.

Sandilos, J. K., Chiu, Y.-H., Chekeni, F. B., Armstrong, A. J., Walk, S. F., Ravichandran, K. S., & Bayliss, D. a. (2012). Pannexin 1, an ATP release channel, is activated by caspase cleavage of its pore-associated C-terminal autoinhibitory region. *The Journal of Biological Chemistry*, *287*(14), 11303–11. doi:10.1074/jbc.M111.323378

Sawada, K., Echigo, N., Juge, N., Miyaji, T., Otsuka, M., Omote, H., … Moriyama, Y. (2008). Identification of a vesicular nucleotide transporter. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(15), 5683–6. doi:10.1073/pnas.0800141105

Schmidt, T. G. M., & Skerra, A. (2007). The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature Protocols*, *2*(6), 1528–35. doi:10.1038/nprot.2007.209

Spray, D. C., Ye, Z., & Ransom, B. R. (2006). Functional connexin "'hemichannels'": A critical appraisal. *Glia*, *54*(August), 758–773. doi:10.1002/glia

Woehrle, T., Yip, L., Elkhal, A., Sumi, Y., Chen, Y., Yao, Y., … Junger, W. G. (2010). Pannexin-1 hemichannel-mediated ATP release together with P2X1 and P2X4 receptors regulate T-cell activation at the immune synapse. *Blood*, *116*(18), 3475–84. doi:10.1182/blood-2010-04-277707

APPENDIX A

DNA VECTORS FOR PROTEIN EXPRESSION

The following chart represents various protein constructs flanking the gene of interest used in this study. Mammalian expression vectors contained a strong mammalian promoter and were used in this study in the HEK293 cell expression system. Insect expression vectors contained two promoters for bacterial and insect cells and were modified from pFastBac (Life Technologies), designed for the Bac-to-Bac Baculovirus Expression and used in this study in the Sf9 cell expression system.

Plasmid Name	**Species Expression**	**Tag Position**	**Affinity Tag**	**EGFP**	**Expression Profile**
pNGFP-EU2	Mammalian	N	8xHis	Y	Met-8xHis-EGFP-ThrombinClvg-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)
pNGFP-FB2	Insect	N	Strep	Y	Met-8xHis-EGFP-ThrombinClvg-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)
pNGFP-FB3	Insect	N	Strep	Y	Met-StrepTag-EGFP-ThrombinClvg-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)
pNNG-FB3	Insect	N	Strep	N	Met-StrepTag-ThrombinClvg-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)
pCGFP-EU2	Mammalian	C	8xHis	Y	Met-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)-ThrombinClvg-EGFP-8xHis
pCGFP-FB3	Insect	C	Strep	Y	Met-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)-ThrombinClvg-EGFP-StrepTag
pCNG-FB7	Insect	C	Strep	N	Met-BamHI(Gly-Ser)-**GENE**-XhoI(Ala-Ser-Ser)-StrepTag

APPENDIX B

GENES SCREENED FOR CRYSTALLIZABILITY

The following genes were screened in this study for expression level, expression profile, and crystallizablity. Sequences were obtained from the NCBI genomic databases and codon-optimized for insect cell expression.

Gene	Species	Abbreviation
pannexin-1	*Homo sapien*	hPanx1
pannexin-1	*Rattus norvegicus*	ratPanx1
pannexin-1	*Bos Taurus*	cowPanx1
pannexin-1	*Canis lupus familiaris*	dogPanx1
pannexin-1	*Xenopus tropicalis*	frPanx1
pannexin-1	*Danio rerio*	zfPanx1
pannexin-1	*Tetraodon nigroviridis*	fuguPanx1
pannexin-1	*Aplysia californica*	aplPanx1
pannexin-2	*Homo sapien*	hPanx2
pannexin-2	*Rattus norvegicus*	ratPanx2
pannexin-2	*Gallus gallus*	chkPanx2
pannexin-2	*Danio rerio*	zfPanx2
pannexin-2	*Tetraodon nigroviridis*	fuguPanx2
pannexin-2	*Aplysia californica*	aplPanx2
pannexin-3	*Homo sapien*	hPanx3
pannexin-3	*Rattus norvegicus*	ratPanx3
pannexin-3	*Bos Taurus*	cowPanx3
pannexin-3	*Canis lupus familiaris*	dogPanx3
pannexin-3	*Gallus gallus*	chkPanx3
pannexin-3	*Danio rerio*	zfPanx3
pannexin-3	*Tetraodon nigroviridis*	fuguPanx3
pannexin-3	*Aplysia californica*	aplPanx3
pannexin-4	*Aplysia californica*	aplPanx4
pannexin-5	*Aplysia californica*	aplPanx5
innexin-1	*Drosophila melanogaster*	flyInx1
innexin-1	*Hydra vulgaris*	hydraInx1
innexin-1	*Spodoptera frugiperda*	sfInx1
innexin-2	*Schmidtea mediterranea*	planInx2
innexin-3	*Caenorhabditis elegans*	wormInx3
innexin-3	*Schmidtea mediterranea*	planInx3
innexin-4	*Caenorhabditis elegans*	wormInx4
innexin-4	*Schmidtea mediterranea*	planInx4
innexin-6	*Caenorhabditis elegans*	wormInx6
innexin-8	*Caenorhabditis elegans*	wormInx8
innexin-9	*Caenorhabditis elegans*	wormInx9
innexin-10	*Schmidtea mediterranea*	planInx10
innexin-11	*Schmidtea mediterranea*	planInx11
innexin-13	*Caenorhabditis elegans*	wormInx13
innexin-18	*Caenorhabditis elegans*	wormInx18

APPENDIX C

PANNEXIN-1 PRIMARY SEQUENCE AND ALIGNMENT

Primary sequence alignment of frPanx1 with other pannexin-1 orthologs. Yellow highlighting indicates sequence identity, blue indicates near sequence identity, and green indicates a similar amino acid substitution (polar, same charge, aliphatic, aromatic).