Ion Mobility - Mass Spectrometry Methods for Studying Membrane Proteins

by

Sarah M. Fantin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2020

Doctoral Committee:

Professor Brandon T. Ruotolo, Chair
Professor Phillip Andrews
Professor Robert T. Kennedy
Professor Bing Ye

Sarah M. Fantin

fantins@umich.edu

ORCID iD: 0000-0002-2509-1143

Dedication

This thesis is dedicated to Dorothy Evans. If you had not supported me applying to University of Michigan, and told me that I had to go if I got int, this would not have happened. Good call G-Ma.

Acknowledgements

I am grateful to have so many advocates, supporters, and collaborators during my academic career. I first thank Professor Brandon Ruotolo for the countless valuable experiences he has enabled during my thesis work. From advising on research directions, to mentoring on career paths, to creating opportunities to advance our skills, it obvious that he respects his lab group and puts great effort into our training. I thank him for providing so many opportunities to present my research and develop my education skills. In my five years of graduate school, he has managed to juggle advising a full lab group, planning a conference from scratch, and the addition of three newborns at home, while still finding time to talk with me about the struggles of membrane proteins and be a reassuring "cheerleader" that my experiments are in fact useful. Thank you for being an amazing mentor.

Next, I thank my committee members and collaborating faculty who have helped shape my research and career trajectory. My committee members of Professor Robert Kennedy, Professor Bing Ye, and Professor Phil Andrews have given me important feedback on interpreting my experimental results and steered the direction of this thesis through my candidacy and data meeting. I would additionally like to thank Professor Robert Kennedy, Professor Charles Sanders, and Professor Anne McNeil for their support in advancing my professional career.

Much of the work presented in this thesis is collaborative, and I have had the pleasure of working with supportive and engaged collaborators out of my university. Professor Shelagh Ferguson-Miller and Dr. Jian Liu provided samples and valuable discussion for Chapter 2. Professor Melanie Ohi and Dr. Pramod Yadav supplied samples for Chapters 3 and 5 as well as

help inform our experiments. Professor Charles Sanders, Dr. Geoffrey Li, and Dr. Hui Huang provided samples and direction for Chapters 3, 4, and 5, and performed cellular assays for Chapter 5. I would also like to thank my collaborators within the Ruotolo group, particularly Kristine Parson (Chapters 2, 3, 5), Brock Juliano (Chapter 3), and Iliana Hampton (Chapter 5).

Being a part of the Ruotolo lab group has been an extraordinary time and part of that is fueled by the extraordinary scientists I have been able to call colleagues. First, my original mentor Shuai Niu. I am not sure you realized that "to TSPO" would lead to at least three students worth of membrane protein related thesis projects, but I thank you for it. All five years I kept your advice about rage quitting experiments at heart, but unfortunately, I have never caught up to your skill level in Dota 2. To past lab members Joseph Eschweiler, Jessica Rabuck-Gibbons, Yuwei Tian, Daniel Polasky (Dan), and Sugyan Dixit (Suggie), thank you for your valuable lab leadership and setting the standard for years to come. Dan, your commitment to "making code do the work" has so greatly impacted the lab, and both you and Suggie's contributions to the CIUSuite2 project enabled all of the work in this thesis. Suggie, thank you for letting me talk to you for long enough that I answered my own questions, being there for me during candidacy, and letting me sing the cool part of Creep so loudly. Now to sum up one of the most amazing experiences of grad school in a few sentences, mentoring, and in turn being mentored by, Kristine Parson and Iliana Hampton. Kristine, working with you on the TSPO project and learning more about nanodiscs through your project was inspiring to me and really kept me motivated to push onwards. Iliana, your contributions to Chapter 5 so far and willingness to take on so much as a rotator are impressive, and I am excited to see how your thesis takes shape! Since I don't think I'm allowed to write paragraphs about everyone, I would also like to thank these friends and colleagues: Varun Gadkari (for your dedication to the lab and leadership as a post-doc), Chunyi Zhao (for all of the fun times

we have had over the past five years), Cara D'Amico (for always teaching us something outside of native MS at group meeting and being a freaking awesome bridesmaid), Daniel Vallejo (for answering all of my illustrator questions and pushing for great causes even outside of lab), Yilin Han (for never allowing there to be a dull moment and being my ASMS buddy), Carolina Rojas-Ramirez (for sharing memes, caring about MS so much, and being my number one conference roommate), Brock Juliano (for blowing up micelles with lasers and being such a good friend), Chae Kyung Jeon (for being ambitious enough to take on a much needed role in lab and for being so supportive in our sport), and Rosendo Villafuerte (for making my last year of grad school so much fun and being your wonderful self).

The professors I had in my undergraduate career led me to where I am today and continue to support me throughout my academic career. Thank you to the chemistry department at University of Indianapolis, including my undergraduate research advisor Professor Brad Neal, for preparing me well for graduate school. You are all the kind of mentors I aspire to be! Additionally, I thank Professor Michael Johnson at University of Kansas for his support during my NSF-REU which sold me on graduate school.

Lastly, I thank my family and friends for their support. Thank you to my parents, John and Renee Fantin, who have supported me during grad school and are always ready to celebrate achievements. I also thank my sisters Leah and Heidi for listening to my problems and asking me science questions that make me feel like I'm useful. Thank you to my extended family in Michigan, living near you has been a lot of fun. A huge thank you to my honorary sister, Harleen Athwal. You have had my back since undergrad and you are still my motivator when times get rough. Finally, to my husband Nathan, thank you for everything including moving to Michigan with me,

chauffeuring me so I wouldn’t have to ride the bus, and listening to enough science talk that I think you could probably pass as a chemist.

Table of Contents

List of Figures

List of Appendices

Abstract

Membrane proteins are important therapeutic targets which play vital roles in cellular function. As protein structure is closely linked to its function, the characterization of membrane proteins is vital to elucidating their involvement in disease and potential druggability. Despite this, membrane proteins are underrepresented in structural databases due to challenges associated with their hydrophobicity and difficulty in obtaining high purity samples. There is an imminent need for methods to study the dynamic structure, folding, and ligand interactions of membrane proteins. Native mass spectrometry (MS) has emerged as a method capable of handling the polydispersity of membrane proteins samples to yield details of their structure and function. Particularly, ion mobility MS (IM-MS) has shown promise in assessing the organization of membrane protein complexes and, through collision induced unfolding (CIU) experiments, the relative stability of gas phase membrane proteins. However, challenges remain in developing robust, high-throughput methods for the study of membrane proteins through IM-MS.

In Chapter 2, we develop a workflow for studying ligand binding of the integral translocator protein (TSPO) with the aim of revealing the identity of an endogenous ligand. As part of this workflow, a system for automated de-noising of CIU fingerprints is presented, allowing for the quantitative stability characterization of eight different lipids and ligands. For the first time, quantitative CIU measurements of stability and classification are used to determine that the ligands endogenously bound to TSPO are a mixture of phospholipid isoforms bearing the PG head group. In Chapter 3, we present a workflow for elucidating key differences among disease associated protein variants through IM-MS. Seven mutant forms of peripheral myelin protein (PMP22) are

shown to possess differences in dimer formation and gas phase stability that correlate with previously published data about cellular trafficking and disease state. Further, from our analysis, we construct a potential mechanism of how dysregulation of PMP22 leads to disease in which mutations cause destabilization of PMP22 monomers leading to the formation of dimeric complexes that are poorly trafficked in cells. In Chapter 4, we present methods to study three mutant forms of the voltage sensing domain of the KCNQ1 potassium channel. We find that CIU of the KCNQ1 mutants classify well according to their observed level of trafficking in cells, indicating a role for specific structural triggers in KCNQ1 trafficking. The function-based classification represents a more broadly applicable method for studying structure differences in proteins with many known mutations.

In Chapter 5, we explore the fundamental relationship between membrane protein solubilization techniques and IM-MS. The integral transmembrane proteins, PMP22 and TSPO, and a monotopic membrane protein, cytochrome P450 3A4, are prepared in multiple solubilization techniques, including micelles, bicelles and, nanodiscs. Noticeable differences in oligomerization are found for TSPO samples liberated from detergent micelles and nanodiscs, and for shifts in the observed charges states are noted. Upon CIU analysis of the complexes, we find that differences in unfolding trajectories and gas phase stability exist for both PMP22 and CYP3A4 when liberated from the varying solubilization agents, which we interpret to relate to the local protein environment. Our experiments do not show that one solubilization method is universally superior at preserving the gas phase structure of membrane proteins as each protein system had different patterns of compaction in terms of CCS, number of CIU transitions, and stability across the solubilization methods.

Chapter 1 Introduction

Membrane proteins play vital roles in many cellular functions including cell signaling, transport, and enzymatic reactions. Embedded in biological membranes, these proteins exist in complex environments, where interactions with lipids, sterols, and other proteins regulate their function.[1] Additionally, the folding and structure of membrane proteins are intimately linked to their environmental interactions.[2] Methods to study membrane proteins are essential to understanding how they perform their cellular functions and, as many membrane proteins are implicated in disease states, how pharmaceuticals may be developed to correct their dysfunction.[3] However, due to their insolubility in aqueous solutions and the complex environments in which they exist natively, membrane proteins are challenging analytes. Here I review some of the general properties, challenges, and methods for studying membrane proteins, with an emphasis on native ion mobility mass spectrometry (IM-MS), which is the focus of this dissertation.

1.1 Membrane proteins and human disease

Membrane proteins represent over 60% of therapeutic drug targets and nearly 80% of drugs approved by the FDA act on membrane proteins.[3] These therapeutic protein targets vary widely in structure, function, and the types of diseases they are implicated in. One example of a well characterized membrane protein target is the serotonin transporter, on which multiple classes of antidepressants act. The role of the serotonin transporter in depression,[4,5] its structure and mode of drug binding,[6–9] and its protein-protein interactions have all been reported.[10] Contrastingly, for

many other membrane proteins implicated in disease much less is understood about their exact biological roles and how disruption of these roles cause disease states. One example of this is the mitochondrial 18 kDa Translocator Protein (TSPO), which is implicated in a range of illnesses including neuroinflammation,[11,12] cancer,[13,14] and cardiovascular disease,[15] based on changes in its expression levels and the observation of disease-correlated mutations. While TSPO is proposed to be involved in the transport of cholesterol and porphyrin, researchers are actively working to elucidate TSPO-related disease eitologies.[16] Peripheral myelin protein 22 (PMP22) is another example of a membrane protein where mutations have been correlated to a range of disease states, including Charcot-Marie-Tooth 1,[17–19] but the mechanism by which these mutations lead to various forms of neuropathy is not wholly understood.[20] For both TSPO and PMP22, no FDA approved drugs currently exist.

1.2 Structure of membrane proteins

Membrane proteins can be classified broadly into two categories, integral and peripheral, based on the way they interact with the membrane.[21] Peripheral proteins do not embed in the membrane and instead interact with the polar lipid head groups on the surface of the membrane in a reversible fashion, as seen in Figure 1-1. Integral proteins have one or more regions of the protein embedded in the membrane and interact with the hydrophobic tails of the lipids in the bilayer in addition to the polar lipid head

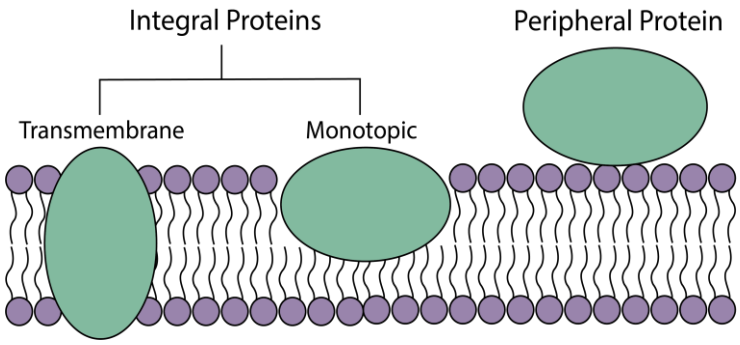

Figure 1-1. Types of membrane proteins. Integral proteins embed in the lipid bilayer while peripheral proteins interact with only the surface of the membrane. Transmembrane integral proteins span the entire membrane and monotopic integral proteins only penetrate part of the bilayer.

groups. Integral proteins can be further divided into classes of transmembrane, which span the entire length of the bilayer, and monotopic proteins, which permanently associate only with one leaflet of the bilayer. For transmembrane proteins, the structure of the regions within the membrane contain either α-helical or β-strand motifs, often rich in hydrophobic residues.[22–24] These motifs commonly manifest as α-helical bundles and β-barrels when multiple membrane spanning domains are present.

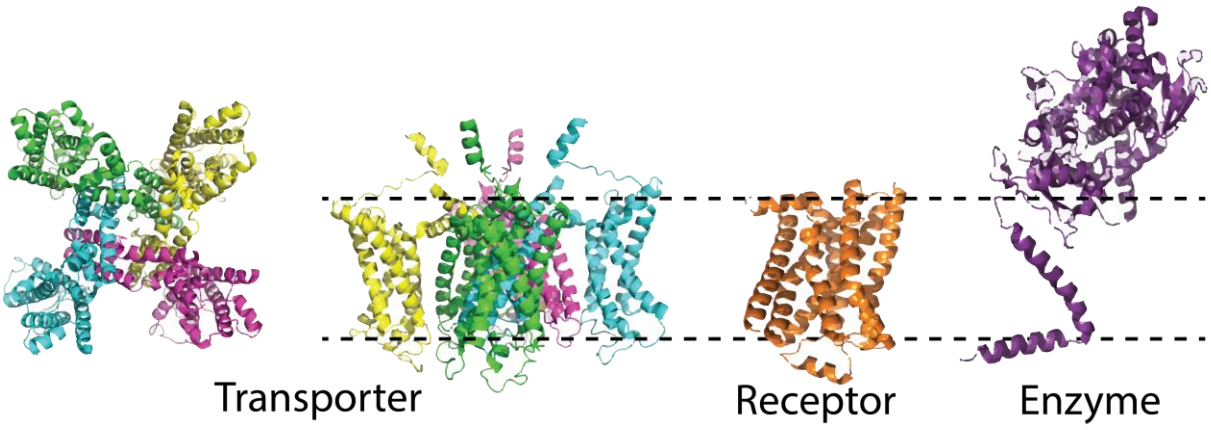

Figure 1-2. Transmembrane protein functional classes. From left to right: top-down and membrane view of a tetrameric voltage gated potassium channel (PDB 6EBM), membrane view of a β-2 adrenergic receptor GPCR (PDB 5D5A), membrane view of a bitopic cytochrome p450 (PDB 4LXJ).

Transmembrane proteins are uniquely suited to perform several functions in the cell including transporting ions and conveying signals across the lipid bilayer. Accordingly, most transmembrane proteins fall into one of three functional categories: transporters, receptors, and enzymes (Figure 1-2).[25] As biological membranes are generally impermeable to most polar molecules, transporter proteins are key for regulating the flow of molecules in and out of cells and organelles.[26] Families of transporters are sensitive to different environmental conditions, such as voltage gated ion channels which respond to charge differences across the membrane.[27] All transporters characterized thus far pass through the membrane multiple times and are thus multi-pass transmembrane proteins. The receptor category of transmembrane proteins is responsible for initiating cellular responses upon specific ligand binding.[28] G-protein coupled receptors (GPCRs)

are one family of receptor transmembrane proteins of great importance to the pharmaceutical industry, as 33% of all small molecule drugs target this protein class.[3]. While GPCRs are multi-pass proteins with seven transmembrane domains, some transmembrane receptor proteins pass through the membrane only once, referred to as bitopic or single-pass proteins.[25] Membrane enzymes are often bitopic proteins, including many of the cytochrome P450 monooxygenases that are responsible for metabolizing many small molecules.[29,30] Note that cytochrome P450s can also be monotopic (e.g. CYP 3A4).[31]

1.3 Membrane proteins are challenging analytes

Despite their importance, membrane proteins are understudied in comparison to their soluble counterparts.[32] This is due to a combination of factors that make membrane proteins challenging analytes. Membrane proteins are difficult to express and purify, which can lead to low yields and impure protein samples which are not amenable to traditional biophysical techniques for structure determination.[33,34] Additionally, the intrinsic hydrophobicity of membrane proteins makes them poorly water soluble and unable to retain their native structure in aqueous solution. As such, a solubilization method must typically be employed to mimic their native hydrophobic environment in order to carry out detailed structure and function characterization efforts.[35,36] Further complicating the study of membrane proteins is the complexity of their native environment, where they may interact with other proteins,[37] endogenous ligands,[38,39] or lipids.[40,41] These challenges have all contributed to the dramatic underrepresentation of membrane proteins in structural databases.[42]

1.3.1 Solubilization methods

To conserve the native structure of membrane proteins in the absence of a cell membrane, multiple solubilization techniques have been developed. The three methods used in this thesis: detergent micelles, bicelles, and nanodiscs, are shown in Figure 1-3. Historically, detergent micelles have been the most popular technique for solubilizing purified membrane proteins and they remain widely used today.[34,42] Detergents can have varying structures and charge, but all possess the ability to form a micelle above certain critical concentration.[33] This concentration is referred to as the critical micelle concentration or CMC in membrane protein purification and analysis protocols. With the optimization of the detergent choice, concentration, and solution conditions, a membrane protein can be encapsulated within detergent micelles, wherein hydrophobic protein residues interact with hydrophobic parts of the detergent and the polar head groups of the detergent solubilize the complex. [42,43]

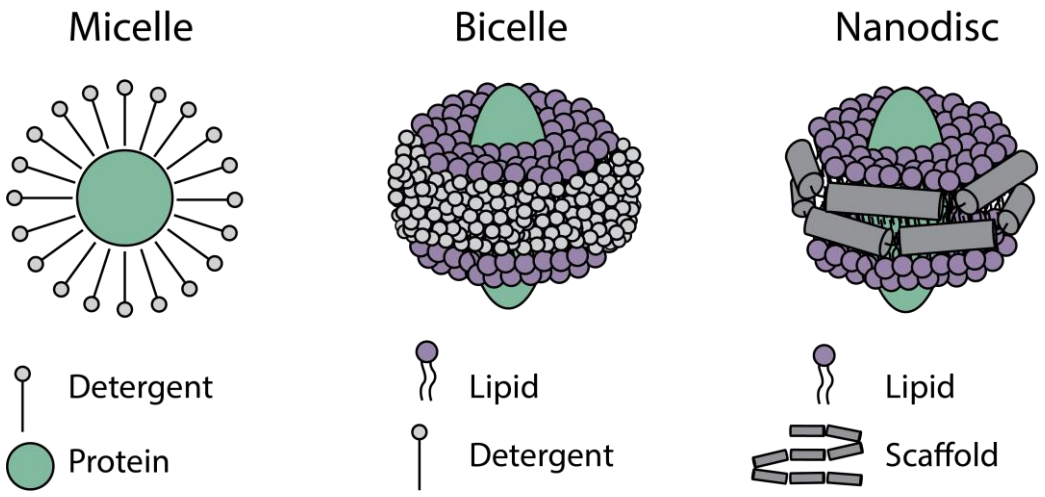

Figure 1-3. Solubilization techniques for membrane proteins employed in this work.

More complex solubilization methods aim to better mimic the native environment of membrane proteins through creating a lipid bilayer. There are many approaches to create this bilayer, each with advantages and disadvantages,[44–46] but only bicelles and nanodiscs will be described. In both bicelles and nanodiscs, the membrane protein is encapsulated by a lipid bilayer,

which consists of one or more lipid type. In bicelles, detergents are used to surround the edges of the bilayer to create a fluid, yet discrete, bilayer structure.[47] In nanodiscs, a scaffold protein or peptide is used in lieu of detergents.[48] For both methods, the thickness of the bilayer,[49] overall sizes of the particles,[48,50] and composition of lipids in the bilayer[51,52] can be optimized for protein incorporation.

1.3.2 Membrane proteins are underrepresented in structural databases

It is estimated that up to 30% of proteins in eukaryotic cells are membrane proteins, but only a small fraction of these proteins have solved structures in databases.[32,33,53] For example, as of 2015, only 3% of unique structures in the Protein Data Bank (PDB) were membrane proteins.[54] Protein structure is intimately linked to function,[55] and thus correcting this disparity has been a major area of research over the last few decades. Towards this end, membrane protein specific structural databases have been created in addition to the PDB, including mpstruc,[56] Orientations of Proteins in Membranes (OPM),[57] and PDBTM.[58] These databases contain almost 12,000 unique membrane protein structures. Advances in structure determination techniques, specifically through advances in electron microscopy (EM), have allowed the proportion of transmembrane (TM) proteins in structural databases to eclipse that of monotopic and peripheral proteins combined, as shown in Figure 1-5.[32] Transmembrane α-helical proteins comprise 57% of total reported structures across all databases, while β-barrel structures comprise only 11%, which loosely translates to their expected representation in the mammalian proteomes.[32,53,59] It is important to

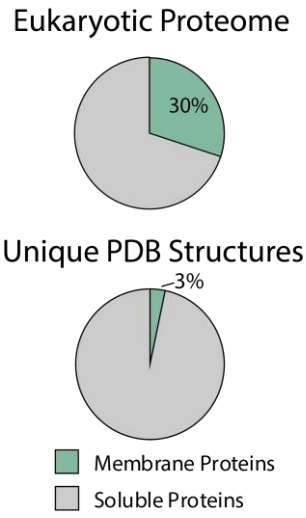

Figure 1-4. Percentage of membrane proteins in the eukaryotic proteome and percentage of membrane proteins in the PDB.

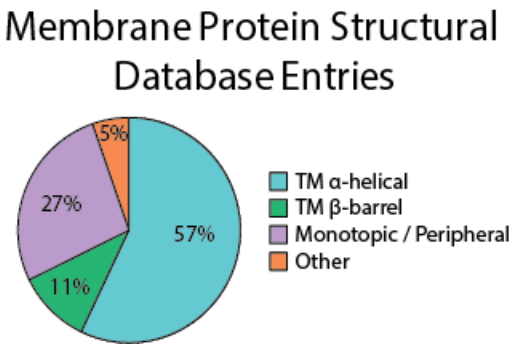

Figure 1-5. Percentage of total membrane protein structure entries across four databases.

note that for all of these reported structures, the full transmembrane region or protein sequence is not necessarily present. Some may also include engineered additions to the protein that aim to increase the proteins stability or amenability to the structure determination technique, which may or may not alter protein structure.[60]

1.4 Biophysical techniques for studying membrane protein structure

The majority of membrane protein structures entered across the PDB and membrane specific databases were obtained through X-ray crystallography.[32] In X-ray crystallography, a purified protein is first crystallized and then subjected to X-ray diffraction analysis. The pattern formed during diffraction is then used to map the electron density of the crystal, allowing for production of atomic resolution 3D structures. For membrane proteins, this process is complicated by both the difficulty of forming crystals from membrane proteins and the typically low purity and concentration of membrane proteins.[33] A major advancement enabling the success of membrane protein crystallography was the development of lipid cubic phase (LCP) methodologies,[61] which has enabled the structural determination of membrane protein complexes,

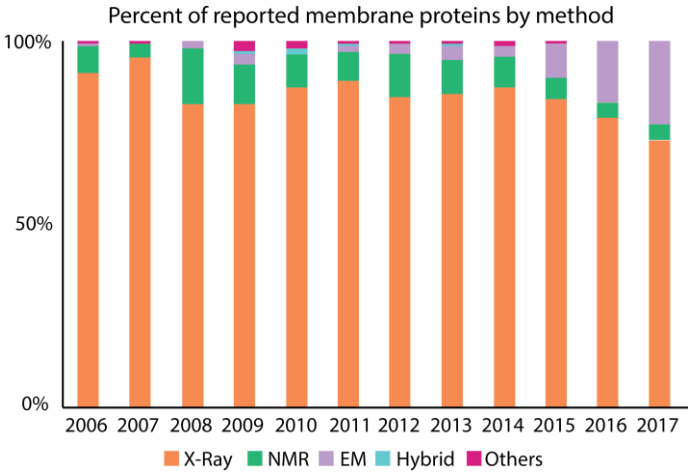

Figure 1-6. Percent of membrane protein structure reported in databases by method per year. *Adapted with permission from Shimizu et al. 2018.*

such as the beta-2-adrenegic G protein complex.[62] Ligands or lipids interacting with membrane proteins have also been observed through this method, however, the electron density alone is not always sufficient to identify these small molecules.[14]

As mentioned above, the number of membrane protein structures derived from EM methods has significantly grown over the past 5 years (Figure 1-6).[32] In this method, a crystal is not required and individual particles are placed in a grid, or within vitrified ice (cryoEM), irradiated with electrons. The structures produced by cryoEM can possess comparable resolution to those produced by X-ray crystallography, however, the success of this approach is limited for proteins under 100 kDa.[63–65] Conversely, nuclear magnetic resonance spectroscopy (NMR), where radio frequencies are used to excite nuclear spin states for atoms within a protein sample and the observed shifts in observed spin frequency are used to measure atomic proximity, can be used to study the dynamics of membrane proteins up to approximately 40 kDa.[66] In NMR, a crystal is not required and multiple protein conformations can be observed leading to measurements capable of capturing dynamic protein motions, but pure and homogenous samples are still required. Both EM and NMR are also compatible with membrane proteins solubilized in lipid bilayer mimetics such as bicelles and nanodiscs.[67–69] In light of the challenges associated with high resolution structural techniques, a plethora of spectroscopic and biophysical assays have been developed to study the structure membrane proteins,[42] including surface plasmon resonance,[70] circular dichroism,[71] and fluorescence assays.[72]

1.5 Mass spectrometry of membrane proteins

Mass spectrometry has recently emerged as a method especially suited for the analysis of native membrane proteins, due in part to its ability to handle complex mixtures and lower sample concentration requirements.[52,54,73–76] In MS, analytes are ionized and transported to a mass

analyzer and detector. The ionization of membrane proteins has been mainly performed through

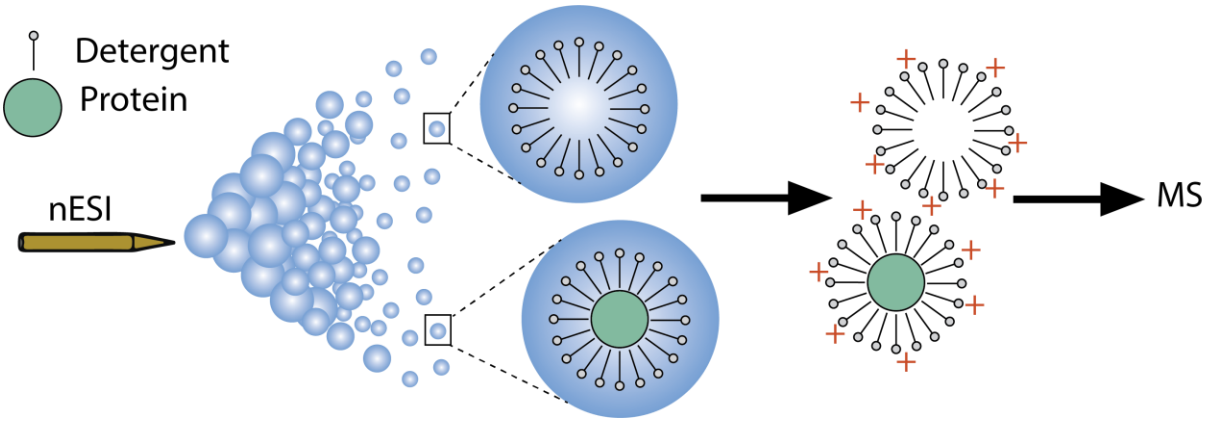

Figure 1-7. Nanoelectrospray ionization (nESI) of membrane proteins. In addition to membrane proteins encapsulated in their solubilization techniques, the solubilization agents themselves can ionize.

electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI),[77] however for the purposes of this work we will focus on ESI methodologies, which dominate the analysis of native protein samples. ESI can be used to study denatured, yet intact, membrane proteins or peptides generated through protein digests of membrane proteins. In contrast, native MS of membrane proteins, where the 3D structure of the membrane protein is left largely unchanged, can be optimally performed through nano-electrospray ionization (nESI), where smaller droplets are produced and heating of the sample is typically not needed to initiate ionization. Membrane protein ions generated through ESI or nESI for native MS analysis are most often still encapsulated in within solubilization agents, and therefore collisional activation must be applied to remove bound detergents or lipids from the target membrane protein ion.[78] Additionally, the independent solubilization agents themselves, e.g. detergents, also ionize and can thus result in an abundance of noise signals in the resulting native mass spectra. While time-of-flight (ToF) mass analyzers have shown great success in this field, higher resolution technologies, such as Orbitrap mass analyzers, can be helpful for resolving such noise signals from intended membrane protein analytes.

1.5.1 Structural proteomics, labeling, and lipidomics

Proteomics, the organism-wide study of protein expression, can be used to identify membrane proteins through their sequence,[59] elucidate post-translational modifications,[79] and reveal interactions between membrane proteins. MS-based proteomics methods typically identify peptide fragments originating from larger proteins through their measured mass and sequencing information derived from tandem MS (MS/MS) methodologies, where one m/z is selected within a first mass analyzer and then fragmented using a variety of activation techniques prior to second-stage mass analysis. Overall, proteomics methods can be divided into those that approach the problem from a bottom-up perspective, where peptides are generated by digestion of a protein in solution are the primary analytes, and those that analyze proteins in a top-down frame of reference, where denatured proteins are the primary analytes, and identification is highly dependent on fragment ions generated within the mass spectrometer. Although proteomics has been widely used to study soluble proteins, the low expression level of many membrane proteins, incompatibility with traditional urea based denaturation techniques, and a limited number of charged residues with their sequences have hampered the success of membrane proteomics.[77] Additionally, the detergents used to solubilize membrane proteins can create abundant noise signals in the mass spectra and, particularly for top-down proteomics, require additional energy within the instrument to remove from the intended analytes. Advancements in membrane enrichment techniques,[80,81] detergent removal steps,[82] and development of MS compatible detergents[83] have all aided the progress of this field in recent years. However, the current technology is not sufficient to characterize the diverse membrane proteome in its entirety,[59] and the study of native membrane proteins through proteomics is still an active area of research with relatively few published manuscripts.[84]

Deeper structural insights can be gained from membrane protein digests through labeling techniques, where the solvent accessible sites of a native protein in solution are labeled permanently or reversibly prior to digestion. Hydrogen-deuterium exchange (HDX), chemical crosslinking, and fast photochemical oxidation (FPOP),[85] have all been used to probe the tertiary structure and interactions of membrane proteins with proteins[86] and solubilization agents.[87]

As lipids are known to interact with and change the functionality of membrane proteins,[41,74,88] lipidomics is an important field adjacent to membrane proteomics. While lipids are much smaller than proteins, they are vastly complex with over 40,000 biological lipids cataloged in the database LIPID MAPS.[89] A challenge for MS-based lipidomics techniques is that some lipids are isobaric, and therefore cannot be differentiated by intact mass alone. MS/MS methods can be employed to enhance the identification of lipids, including some which are isobaric.[90] Databases and software have been curated using libraries of known MS/MS spectra to assist in the identification of lipids.[91,92] One limitation of lipidomics in respect to structural membrane protein analysis is that lipids observed to co-purify with membrane proteins may or may not represent evidence of functional protein-lipid complexes. Additionally, information on where such lipids interact with the protein, their stoichiometries, and binding strengths cannot be directly obtained through these experiments alone.

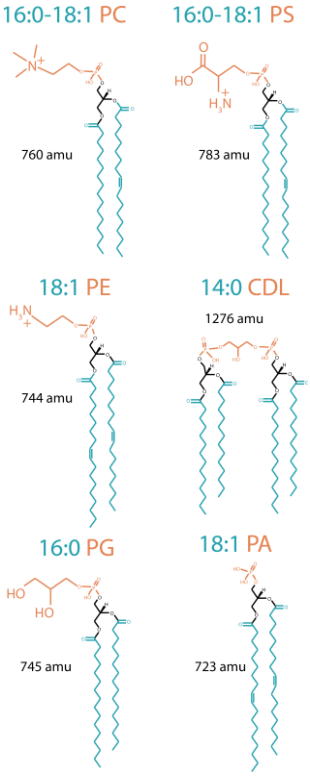

Figure 1-8. Examples phospholipids illustrating the diversity among lipids in both head group (orange) and chains (blue).

1.5.2 Native ion mobility - mass spectrometry

Recently, native MS has been shown to be a powerful tool for probing membrane protein structure and function.[54,74,75,93,94] After nESI, membrane proteins can be liberated from a variety of detergent and solubilization agents in the gas phase using different techniques for activation, such as collision induced dissociation (CID). It is important to note that not all solubilization agents are equally effective in this endeavor, and screening detergents, solution conditions, and

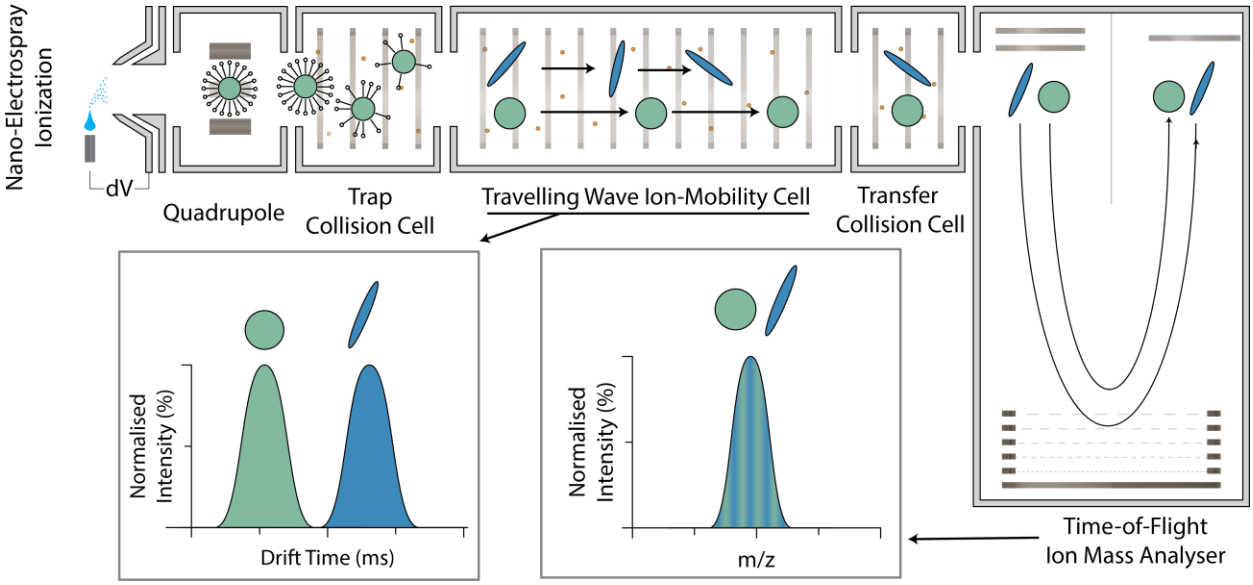

Figure 1-9. Instrument diagram of an IM-MS instrument, Synapt G2 (Waters, Milford MA). After nESI, membrane proteins are transported through the quadrupole still encapsulated in their polydisperse solubilization agents. Membrane proteins are liberated from the agents in the trap collision cell and then separated in a travelling wave IM cell. Additional energy can be added in the transfer collision cell to dissociate clusters of solubilization agents before time-of-flight mass analysis.

optimizing instrument parameters is a necessity for striking the delicate balance between the removal of solubilization agents and retention of native structure.[95] However, with successful optimization, native MS has been used to study discrete lipid binding,[74,94,96] ligand binding,[97] and specific protein-protein interactions associated with a wide range of membrane proteins.[52,75,98]

The addition of an ion mobility separation to mass spectrometry instruments (IM-MS) is beneficial for native MS of membrane proteins. IM separates protein ions according to their size and shape by recording their drift through a neutral buffer gas under the influence of a weak electric

field. The output of IM analyses is often the determination of collision cross section (CCS), or the orientationally averaged size, of ions which can be compared to cross sections calculated from

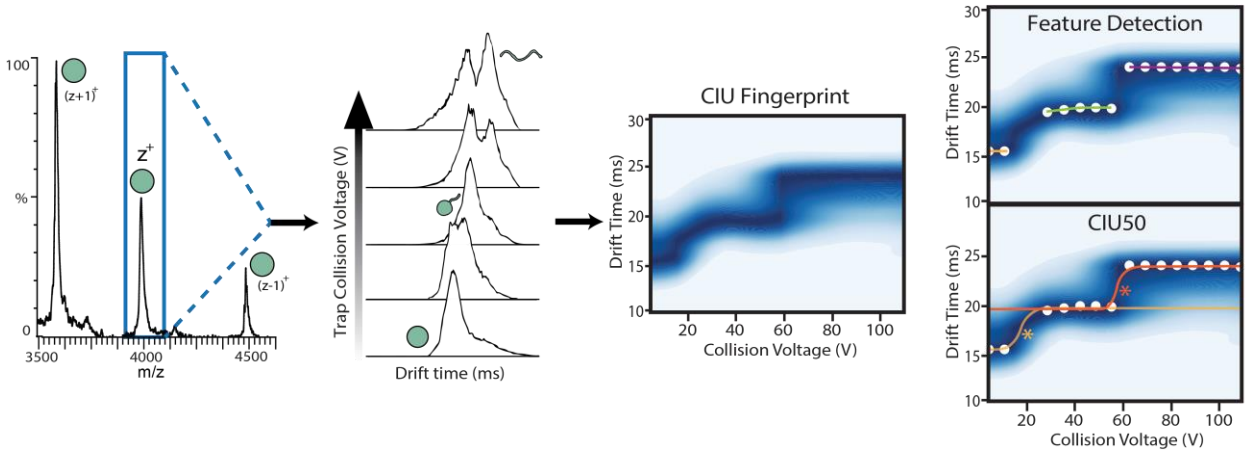

Figure 1-10. Collision induced unfolding (CIU). The arrival time distribution (ms) for a specific m/z is extracted from the IM-MS data set. As the CV is ramped in the trap of the instrument the protein unfolds, which can be visualized through CIU fingerprints. Software allows for automated analysis of CIU fingerprints including the detection of features and generation of sigmoidal fits between features yielding CIU50 values.

structures found within in databases or produced through molecular dynamics simulations.[99] Additionally, IM-MS platforms, such as the Waters Synapt G2 (Milford, MA, USA) used in this work (Figure 1-9), include supplemental trapping regions that provide opportunities to perform CID for both liberation of membrane proteins from their solubilization agents and breaking up free clusters of solubilization agents, which can greatly increase signal quality.

The energy experienced by ions in these trapping regions is a function of an accelerating or collision voltage (CV), and, ideally, optimized solubilization systems can be removed at relatively low CVs. At CVs higher than the threshold for removal of the solubilization agents, the membrane protein can experience collision induced unfolding (CIU), and this unfolding can be tracked through the resulting IM arrival time distribution.[100] As seen in Figure 1-10, these data can be plotted as CIU fingerprints with CV on the x-axis and arrival time on the y-axis with intensity indicated by a color scale. Recent software developments allow for the automated detection of features, or regions of conformational stability in CV space, as well as the fitting of sigmoidal

functions between features to output discrete values where transitions occur (CIU50 values).[101] CIU experiments are valuable for assessing the relative gas phase stability of membrane protein complexes[74,102] and have been used in the past to classify soluble protein systems.[73] Despite the successful characterization of a range of membrane protein systems, many IM-MS and CIU datasets reported in the literature exhibit solubilization agent related noise,[40,103] and currently, membrane protein CIU has only been demonstrated with protein liberated from detergent micelles. Methods for minimizing and handling detergent, lipid, and scaffold related noise are key to enabling CIU of membrane protein solubilized in more complex lipid mimetics.

IM-MS methods for studying membrane proteins and membrane protein complexes have sample concentration requirements in the low nanomolar range, which is significantly lower than high resolution structural techniques such as X-Ray crystallography.[61,95] IM-MS is also compatible with many different membrane protein solubilization methods over a broad range of protein sizes.[104] While IM-MS methods have been shown to yield dynamic information on membrane protein structure, atomic level protein structure cannot be solved through IM-MS alone, and thus IM-MS methods are often combined with traditional biophysical techniques.[105,106] Additionally, membrane protein IM-MS and CIU are still relatively new fields, both gaining popularity within the last decade, and their full potential is still actively being explored.[100]

1.6 Summary

This thesis focuses on developing and applying IM-MS methodologies to pharmaceutically relevant membrane proteins. It can be divided into two parts: 1) developing and applying workflows for elucidating biologically relevant information from therapeutic targets

through IM-MS and 2) exploring the fundamental effects of membrane proteins solubilization agents on gas phase protein structure and unfolding.

Chapters 2, 3, and 4 are aimed at studying the structural biology of therapeutically relevant membrane proteins. In Chapter 2, a workflow for de-noising of CIU fingerprints and the identification of an endogenously bound ligand to TSPO is revealed. This work has been published as **Fantin, S. M, Parson, K. F., Niu, S., Liu, J., Polasky, D. A., Dixit, S. M., Ferguson-Miller, S. M., and Ruotolo, B. T. Collision Induced Unfolding Classifies Ligands Bound to the Integral Membrane Translocator Protein.** ***Anal. Chem.*** **2019*, 91* (24), 15469-15476.** Chapter 3 elucidates a possible mechanism of disease for mutant forms of PMP22 through observed differences in protein oligomeric state and gas phase stability. In Chapter 4, gas phase unfolding of mutant forms of the voltage sensing domain (VSD) of KCNQ1 is demonstrated as a means to classify mutants according to their cellular trafficking behavior. Chapter 5 focuses on the effects of solubilization technique, namely micelles, bicelles and nanodiscs, on gas phase protein structure and unfolding. We find that there are significant differences in protein oligomeric state, charge state, and unfolding trajectories when such analytes are prepared using different solubilization techniques, which calls into question how such agents should be used to preserve native membrane protein structure in general. In Chapter 6 we conclude by summarizing our findings and their impact on the field of nMS, and we propose future directions related to high throughput screening and the expansion of the solubilization agent study described in Chapter 5.

Chapter 2 CIU Classifies Ligand Binding of Integral Membrane Translocator Protein (TSPO)

(*Adapted with permission from:* Sarah M. Fantin, Kristine F. Parson, Shuai Niu, Jian Liu, Daniel A. Polasky, Sugyan M. Dixit, Shelagh M. Ferguson-Miller, and Brandon T. Ruotolo. Collision Induced Unfolding Classifies Ligands Bound to the Integral Membrane Translocator Protein. *Analytical Chemistry* **2019** 91 (24), 15469-15476. Copyright 2019 American Chemical Society.)

2.1 Introduction

Membrane proteins play central roles in cellular signaling, molecular trafficking, and a wide variety of enzymatic reactions.[107] As such, membrane proteins represent a disproportionately large number of current drug targets with respect to their predicted abundance in the human genome.[108,109] Despite their importance, the poor solubility of membrane proteins in aqueous solvents often limits their expression and purification yields,[35,110] resulting in a deficit of membrane protein structural knowledge. Specifically, membrane proteins currently comprise only 3% of unique entries in the Protein Data Bank.[54] Recent advancements in NMR, X-Ray, and EM methods have moved to close this gap; however, many challenges remain in assessing membrane protein dynamics, polydispersity, and ligand binding in a high-throughput manner.[111–114]

Recently, native mass spectrometry (MS) has been deployed to address the unique challenges associated with studying the structure and function of membrane proteins. Typical experimental workflows involve the micellar solubilization of membrane proteins, the use of nano-electrospray ionization (nESI) to produce micelle-protein complex ions, and the subsequent

breakdown of detergent micelle ions using collisional activation to reveal membrane protein as well as smaller surfactant cluster ion populations. Insights into membrane protein complex organization,[84,115,116] lipid binding,[96,117,118] and membrane mimetic performance[52,118–121] have been gained by leveraging the ability of MS to manage the complex mixtures that typically result from this workflow. Ion mobility-mass spectrometry (IM-MS), which sorts ions according to both their orientiationally-averaged size and m/z ratio, can be used to capture large, global changes in membrane protein structure associated with their function.[122] IM-MS enables the use of collision induced unfolding (CIU), where collisional activation is utilized to unfold isolated protein ions in the gas phase. CIU of membrane proteins can be sensitive to ligand binding, and gas-phase stability shifts have been linked to lipid binding selectivity.[106,123,124] Despite these successes, many challenges hinder the broader implementation of CIU methodologies for membrane protein complexes. Generating tandem MS of sufficient quality and detail to identify ligands ejected from membrane proteins is currently difficult, if not impossible, to achieve. In addition, chemical noise generated by protein solubilization agents can overlap with signals associated with small membrane protein ligand complexes, creating a convoluting signal population within CIU fingerprints and obfuscating key stability information.

An example of one such small membrane protein complex is the 36 kDa dimer of translocator protein (TSPO), which plays key roles in stress response, cancer, neurodegenerative conditions, and cardiovascular diseases[13,15,125,126]. Multiple classes of ligands have been proposed to bind to specific regions of the TSPO surface, including cholesterol, protoporphyrin, lipids, and small drug-like molecules.[127,128] Despite its plurality of proposed ligands, TSPO is still an orphan receptor, in that there is currently no consensus regarding the identity of its endogenous ligands,

and many questions remain surrounding the potential allosteric impact of ligand binding on TSPO structure.[129]

Here, we develop CIU data analysis strategies and deploy these techniques to characterize TSPO-ligand complexes. We quantify stability shifts associated with lipid and porphyrin binding to TSPO, and observe a stronger stability shift upon protoporphyrin IX (PPIX) binding when compared to any of the lipid complexes we detect. Furthermore, we build a CIU classification scheme capable of differentiating two sets of TSPO:ligand complexes by their proposed binding region. Finally, we utilize CIU fingerprint classification and high resolution mass spectrometry (HRMS) to characterize an endogenous TSPO ligand, putatively identifying the bound material as a mixture of phosphotidylglycerol (PG) isomers having different levels of chain saturation. We conclude by discussing the general utility of our enhanced CIU methods for the study of membrane protein complexes.

2.2 Methods

2.2.1 Membrane Protein Sample Preparation

TSPO bearing the A138F point mutation was purified and expressed using established protocols from Rhodobacter sphaeroides.[128] All lipids (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine [PE], 1',3'-bis[1,2-dimyristoyl-sn-glycero-3-phospho]-glycerol [CDL], 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine [PS], 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine [PC], 1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-rac-glycerol) [PG], 1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-myo-inositol)[PI], and 1,2-dioleoyl-sn-glycero-3-phosphate[PA]) were purchased from Avanti Polar Lipids (Alabaster, AL). Ammonium acetate,

Protoporphyrin IX, Octyl β-D-glucopyranoside (OG), and chloroform were purchased from Sigma Aldrich (St. Louis, MO). Isopropanol and methanol were purchased from Fisher Scientific (Hampton, NH). For all native MS experiments, 30 μM TSPO was buffer and detergent exchanged simultaneously from 50 mM Tris, 150 mM NaCl , 0.20% DM (N-dodecyl β-D-maltoside), pH 8.0 to 40 mM OG , 200 mM ammonium acetate, pH 8.0 using 100 kDa Amicon Ultra-0.5 centrifugal filter units (MilliporeSigma, Burlington, MA). Detergent screening and lipid binding studies were performed following established protocols,[95] where the protein:lipid molar ratio was optimized for each lipid to be between 1:1 and 1:3, using an incubation period of 10 minutes per binding experiment. Protoporphyrin IX (PPIX) was dissolved in ammonium acetate buffer and added to TSPO in a light-free environment at an optimized protein:ligand ratio of 1:10. An endogenous ligand was noted to be present in some samples provided, and from these samples, replicates of the endogenous ligand (UNK) binding were generated without further preparation.

2.2.2 CIU Experiments

All CIU data was collected using a Synapt G2 HDMS IM-Q-ToF mass spectrometer (Waters, Milford, MA), with a direct infusion nESI source set to positive ion mode. Our instrument settings were tuned to generate intact protein ions while completely dissociating detergent adducts prior to the IM separator, including appropriately tuned settings for the source temperature (40° C), helium cell gas flow (100 mL/min), and the sampling cone (120 V). Trapping cell wave velocity and height were 116 m/s and 0.1 V. IMS wave velocity and height were 700 m/s and 32.5 V. Transfer cell wave velocity and height were 300 m/s and 10 V, with an accelerating potential of 70 V used to dissociate empty micelles. Collision cross section analysis was performed by established methods with the protein standards concavalin A, avidin, and β-lactoglobulin.

Theoretical collision cross sections of dimeric TSPO were calculatedusing IMPACT.[99,130,131] All CIU analyses were performed by increasing the trap collision voltage in a stepwise manner 50 – 140 V in 5 V increments. CIU data from the 9+ charge state of the TSPO dimer was extracted into a text-based format using TWIMExtract[132], then processed and analyzed using CIUSuite 2.[133] Data processing included two rounds of 2D Savitzky-Golay smoothing with a window of five bins and interpolation of the collision voltage axis by a factor of four.

Gaussian de-noising and the subsequent classification of Gaussian processed CIU data was performed using a modified version of the classification module of CIUSuite 2, which we term CIUSuite 2.1, in conjunction with manual feature selection. Information regarding the algorithms that underpin CIUSuite 2.1 software is available elsewhere.[134] Briefly, we modified the standard classification algorithm within CIUSuite by both standardizing and normalizing input CIU data as well as updating our approach to Gaussian fitting of IM ATDs. Our data normalization procedure is designed so that the most intense signals across different classes receive equal weight within resulting classifiers. We then standardize this data to a normal distribution possessing both a zero mean and unit variance.[135] This approach prevents bias in the support vector machine (SVM) approach that we utilize to generate downstream classifiers and dramatically improves the robustness of the classifiers produced. In CIUSuite 2.1, the normalization and standardization steps described above have been implemented for both raw CIU data and the Gaussian features that result from our de-noising procedures.[133] For Gaussian data, fitted peak centroids, widths, and amplitudes were standardized separately at each collision voltage. If multiple peaks were present at a given collision voltage, their attributes were averaged for standardization. The Gaussian fitting module of CIUSuite 2 was also updated as part of this work to utilize the optimized fitting results from the analyses of previous collision voltages as the initial guesses for the next collision voltage

data processed. Since CIU features are typically stable over several collision voltage steps, our CIUSuite 2.1 approach results in higher quality initial guesses in all cases, except for those IM datasets where new features first appear. In general, this new Gaussian fitting approach results in more consistent and higher quality fits in the noisy membrane protein CIU datasets presented here.

2.2.3 High Resolution MS Lipid Extraction Experiments

30 μL of 1:2 chloroform:methanol solution was added to 2.5 μ L of 3.3 mM TSPO in 50 mM Tris, 150 mM NaCl , 0.20% DM, pH 8.0. This solution was then vortexed and allowed to incubate for 10 minutes on ice. The resulting solution was dried under a stream of nitrogen gas and resuspended in 40 μL1:1 isopropanol:methanol, sonicated for 30 minutes, and then directly infused into a nESI source of an Orbitrap Fusion Lumos Tribrid MS (Thermo Scientific Incorporated, Waltham, MA). MS spectra were collected in positive and negative mode, with mass selection and fragmentation performed for any peak of interest at HCD energy values between 0 and 40. MS/MS data was analyzed using LipidMaps.[89]

2.3 Results and Discussion

2.3.1 Analyzing CIU Data for TSPO:Ligand Complexes

Previous reports have demonstrated the capabilities of CIU to capture stability shifts within membrane proteins upon ligand binding.[122,123,136] However, generating robust, quantitative descriptions of membrane protein complexes using CIU remains a significant challenge, prompting the creation of custom CIU data analysis software.[133,136] In this report, we present a new workflow for the automated quantitive analysis of membrane protein CIU data, which we utilize to extract stability values for a range of TSPO:ligand complexes, wherein we first detect complex ions using native MS, then perform CIU and assign each feature observed according to their voltage ranges and centroid IM drift times, allowing us to fit sigmoid functions to the transitions between the CIU features

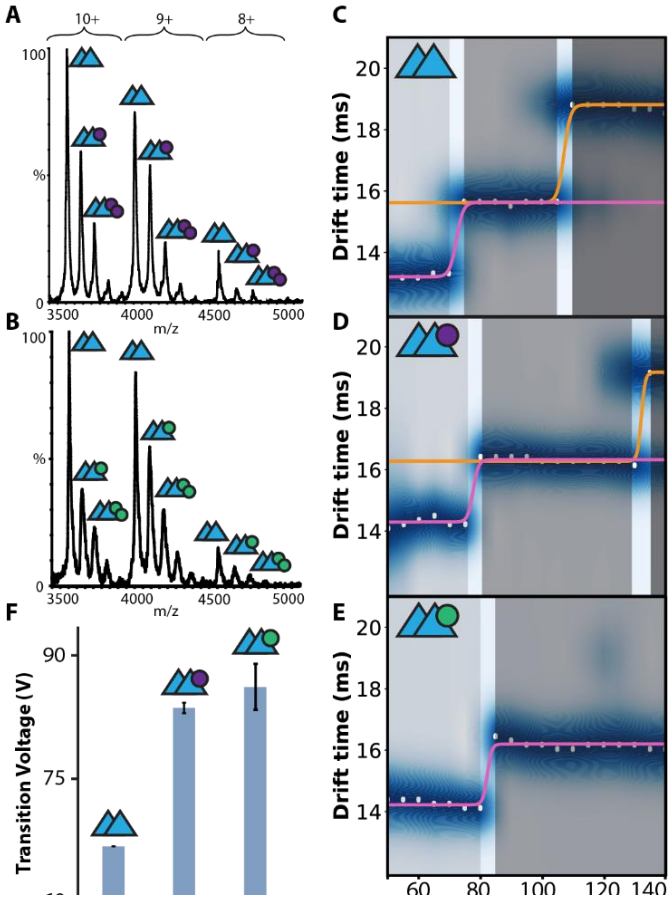

Figure 2-1. CIU analysis workflow for membrane protein complexes. **A,B.** Representative MS data for dimer TSPO:ligand complexes, where triangles denote TSPO monomers and circles represent PI (purple) and PG (green). **C,D,E**. CIU fingerprints of specific protein:ligand complex stoichiometries are analyzed through automated feature detection (shaded boxes) and CIU50 transition fitting (sigmoid traces). Here, 9+ dimer TSPO (MS not shown) and TSPO:ligand complexes are analyzed. **F**. CIU50 values are extracted across replicates, N = 3, in order to quantify complex stability and assess error, in this case clearly indicating the stabilization of TSPO upon binding both PI and PG lipids as opposed to when no lipid is present.

observed, and lastly extract the inflection points for each sigmoid fit, which we term the CIU50 value.

Two example MS datasets for TSPO:ligand complexes are shown in Figure 2-1, both acquired under activating conditions so as to dissociate the majority of interfering surfactant clusters and simplify MS interpretation. While both monomer and dimer states of the protein are observed over multiple charge states, the data shown in subsequent panels focuses on 9+ TSPO dimer signals. CIU experiments involve the stepwise increase of the field strength associated with the injection of ions into the trapping region preceding IM separation. Thus, ions accumulate excess kinetic energy, which is subsequently converted to internal modes via collisions with Ar gas, leading directly to protein ion unfolding. Plots of IM drift time versus the voltage are used to define the ion kinetic energies used to initiate CIU (CIU Fingerprints). These serve to track the structural transitions undergone by the ions as a function of their internal energies (Figure 2-1C). CIU fingerprints typically display multiple features related to protein ion unfolding, and we use CIUSuite2.1 software to detect and define these structural states in terms of their centroid drift times, and the range of collision voltages over which they appear. We then fit the transitions between the CIU features observed and extract the inflection points, or CIU50 values, for these fits from the data collected (Figure 2-1C-E). The example data shown in Figure 2-1 indicates that the transition between the most compact CIU feature and its nearest neighbor occurs at higher collision voltage values when PG or PI lipids are bound to the TSPO ions studied here (Figure 2-1D, E). In order to display and compare CIU stability data, we typically project bar charts that track each CIU50 recorded for the initial unfolding transition, although the CIU50 values for higher-energy transitions can also be readily compared in this fashion if the intensity of the highest energy feature reaches 50% of the maximum signal.

2.3.2 Data Processing Reduces Micellar Chemical Noise

Previous reports have studied the role of detergent in preserving native-like conformations for gas-phase membrane proteins ions,[95,137] indicating that detergent choice can dramatically impact IM-MS data quality and ion stability measurements.[54,122,138] Data quality can also be impacted by surfactant-related chemical noise, where stable surfactant cluster ions can produce high intensity signals and excess detergent adduction can obfuscate protein data. Although the impact of surfactant-related noise on the native MS analyses of membrane proteins can be observed in multiple previous reports,[40,78,119,138] we have found such noise to present a particular challenge for the IM-MS analysis of smaller membrane protein complexes, where significant overlap between micelle cluster ion signals and protein signals can exist. TSPO dimers, at 36 kDa, are an example of one such small membrane protein complex, where CIU analysis is dramatically complicated by chemical noise arising from OG detergent ions, which we have observed to be an optimal detergent

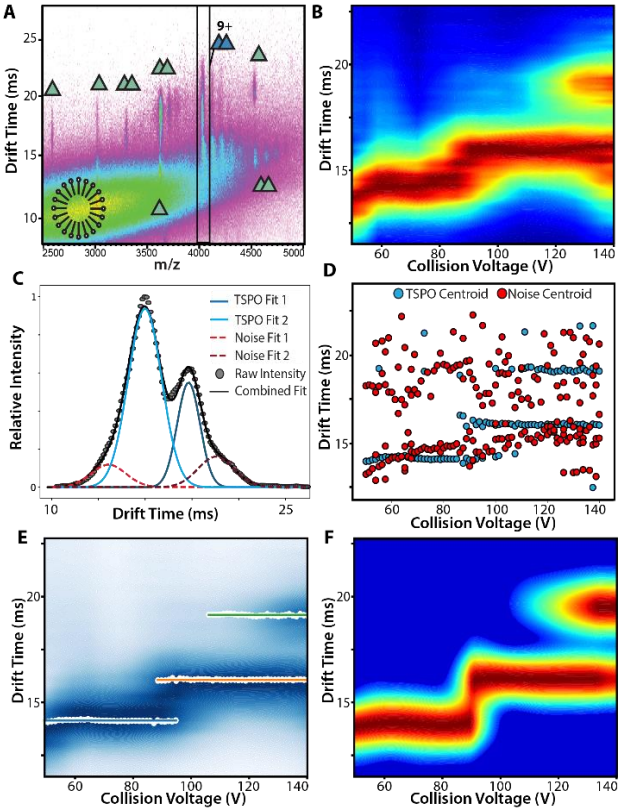

Figure 2-2. A Gaussian noise filter for membrane protein CIU Data. **A**. IM-MS data for TSPO-ligand complexes where surfactant micelle-related noise overlaps with the signals for the 9+ charge state of TSPO (boxed). **B**. Resulting CIU fingerprints recorded for the noise-contaminated 9+ TSPO signal. **C, D**. Gaussian fitting of interpolated TSPO IM drift time data recorded at a single collision voltage indicates that surfactant noise typically exhibits broader peak widths than TSPO signals, allowing informative protein signals (blue) to be sorted from noise (red) across all collision voltages scanned to create the complete CIU fingerprint. **E**. Performing feature detection on the Gaussian fitted peaks in panel D allows for the robust identification of three features within the resulting CIU fingerprint (see Figure S3 for more detail). **F**. Reconstruction of the features detected from the Gaussian fits gives a de-noised fingerprint.

choice for native MS of TSPO (Figure I-2). For the TSPO data discussed in this report, we have developed a CIUSuite 2.1 approach that automatically detects and removes surfactant noise from CIU data using a combination of data interpolation and Gaussian filtering.

Figure 2-2A displays a typical TSPO IM-MS dataset in the presence of OG detergent, where we observe signals identified as TSPO monomer and dimer over a range of charge states indicative of native like structures (5-12+). We chose to evaluate dimeric TSPO as it has been identified as the biologically active form.[129] Furthermore, our native IM-MS data reveals evidence of ligand binding exclusively within dimeric TSPO. We focused our analysis on signals assigned to the 9+ charge state of the TSPO dimer based on their signal intensities relative to other dimer ion signals and their lack of m/z overlap with TSPO monomer signals. Additionally, the collision cross section recorded for 9+ TSPO, 2751 $Å^2$, is in good agreement with theoretical estimates based on the available X-ray structure, 2741 $Å^2$, indicating native-like folding for ions under these conditions (PDB 4UC1).[14] A broad set of signals associated with OG surfactants and clusters is observed to overlap significantly with the TSPO 9+ peak (Figure 2-2A), producing CIU fingerprints contaminated with micellar noise, which increases in intensity relative to protein signals detected at higher collision voltages (Figure 2-2B).

Automated Gaussian fitting applied to interpolated IM data acquired for 9+ TSPO dimer ions detects two main classes of ions readily differentiated based on their IM peak widths and intensities (Figure 2-2C), with broadened, low-intensity features belonging primarily to surfactant noise. We utilized this difference to identify the noise signals detected across all collision voltages scanned during the CIU experiment, developing a scoring algorithm based on peak width and intensity to do so (Figure 2-2D, I-3). While the signals assigned as noise appear at a variety of drift times, the most intense, and thus problematic, noise signals are those that occupy drift times similar

to those observed for protein ions. In the data shown in Figure 2-2, OG-related cluster ion noise signals appear at IM drift times ranging from 13.0 -15.5 ms across all collision voltages probed during CIU. In addition, we observe that the positioning of OG cluster ions in drift time-m/z space changes as a function of the collision voltage setting used to acquire CIU data, further complicating data analysis. Despite the challenge presented by this complex noise distribution, following our Gaussian de-noising procedure, we are able to accurately detect and quantify three TSPO CIU features free of interfering noise (Figure 2-2E, Figure I-4). These features can be used to reconstruct de-noised fingerprints from the Gaussian fits used during the de-noising process (Figure 2-2F), which can be used for CIU50 and classification analysis, in a manner impossible to achieve using the original CIU data alone.[133]

2.3.3 Ligand Binding Generally Stabilizes TSPO

Both lipids and protoporphyrins have known binding sites on the surface of the TSPO dimer, but little is known regarding the biophysical consequences of TSPO ligand binding. Figure 2-3A indicates the locations for these binding sites which are based on a combination of crystal structures and computational modeling. Specifically, protoporphyrins occupy a discrete pocket between transmembrane helicies I and II and underneath loop 1, while phospholipids likely attach to both the hydrophobic dimer interface as well as a series large, hydrophobic grooves in the

membrane embedded region of the protein.[129] Using native IM-MS we were able to evaluate nine TSPO ligands, including a diverse set of eight lipids and PPIX, observing a wide range of binding stoichiometries to TSPO dimers for each ligand tested (Figure 2-3, S4). All TSPO complexes studied here, except PC complexes, exhibited the same number of CIU features when compared to the unbound dimer, and resulted in various degrees of stabilization upon complex formation (Figure II-6). Our CIUSuite2.1 workflow was used to obtain de-noised CIU50 values for each complex observed.

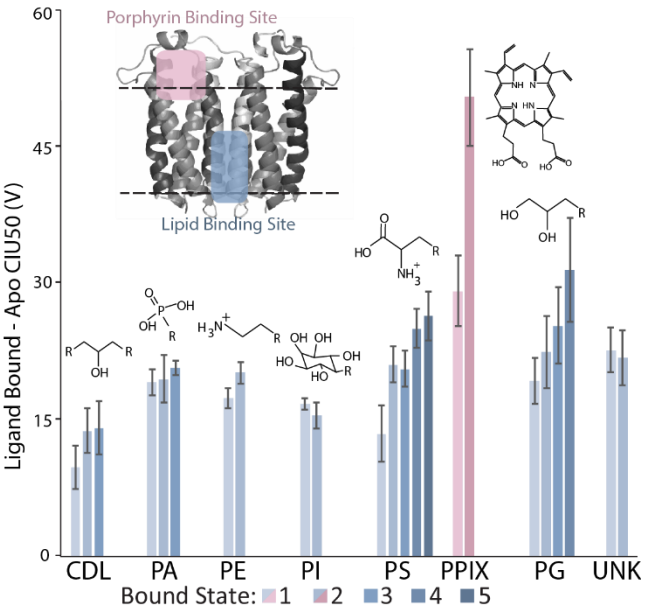

Figure 2-3. Net CIU stabilities of TSPO:ligand complexes. After incubation with a selected ligand, CIU50 values for 9+ TSPO dimer ions were quantified, N = 3, in a range of liganded states (key, bottom). The CIU50 of unbound TSPO (no lipid present) was subtracted from the CIU50 values extracted for TSPO:ligand complexes to generate a measurement of net CIU stability. Among lipids, subsequent binding events tend to increase TSPO stability, with some lipids acting to stabilize the complex to a greater extent than others. The PPIX binding site identified from X-ray structures occupies a separate site from that identified for lipids (inset PDB 4UC1),46 and TSPO:PPIX complexes exhibit enhanced stabilities when compared to the lipids screened here. An unidentified endogenous ligand (UNK) exhibits CIU stabilities with the greatest similarity to the majority of lipids analyzed in this report.

Figure 2-3 displays CIU50 values associated with the first unfolding transition observed, collected in triplicate for PPIX, six lipids (CDL, PA, PE, PI, PS, PG), and a previously unidentified endogenous ligand, labeled UNK. All of these ligands exhibited binding stoichiometries with dimeric TSPO ranging from 1:2-1:5. In order to compare CIU50 values across all complexes measured, Figure 2-3 plots CIU50 values that are a product of subtracting the equivalent unbound TSPO CIU50 value. This net CIU stabilization measurement reveals that PPIX binding stabilizes gas-phase TSPO to a greater extent than any protein-lipid complex detected and discrete lipid bound states appear to differentially stabilize the TSPO dimer. As noted above, TSPO-PC

complexes provide a notable exception to this trend, in that lipid binding appears to destabilize the complex and alter the observed CIU pathway significantly (Figure I-6).

The presence of TSPO-UNK complexes allows us to investigate the ability of CIU to classify an unknown membrane protein ligand, in this case, extracted endogenously from expressed TSPO samples. A cursory analysis of the CIU data shown in Figure 2-3 indicates stability values for UNK complexes similar to most of the lipids screened in our experiments, excluding the lipids PS and CDL. Although we are unable to measure the mass of UNK in native MS data with high accuracy (740 ± 20 Da), such MS data supports the conclusion that the unknown ligand is likely not related to PPIX. Taking the 1:1 and 1:2 TSPO dimer:lipid complex CIU stability values observed as a guide, UNK CIU50 values are within error of those recorded for equivalent TSPO complexes containing both PG and PA. As an MS/MS analysis of UNK could not be achieved due to the presence of significant surfactant-related chemical noise, we performed lipid extraction experiments aimed at identifying all endogenous lipids co-purified with TSPO samples. These experiments revealed a range of PG isoforms as the principle endogenous lipids present within these TSPO samples (Figure I-7).

2.3.4 CIU Classification of Ligand Binding Location

As depicted in Figure 2-3, prior research has determined that TSPO possesses multiple binding surfaces for ligand attachment, including separate sites for drug and cholesterol attachment, not depicted.[128] TSPO shares this quality with a plurality of membrane protein drug targets, and rapidly confirming binding location on a target protein surface often represents a significant challenge associated with the discovery of new therapeutics.[73,139,140] As such, we investigated the ability of CIU to provide such information. We began by utilizing de-noised and

interpolated CIU fingerprints acquired for 1:1 TSPO dimer:ligand complexes as a training dataset which can be sorted into two classes based on proposed binding location: PPIX and lipid-associated complexes. Using CIUSuite 2.1, we built a classification scheme, where each voltage was evaluated for its ability to differentiate the two training classes, generating a score for each test (Figure 2-4A). By considering both the absolute value and error of the score generated for each voltage, the region shown in the gray box in Figure 2-4A was automatically selected for a "leave one out" cross-validation step (Figure 2-4A, inset) which serves to evaluate the robustness of an identified classifier through evaluation with data not included in the training set.[133] Our analysis indicated that when the voltages ranging from 116.25 V to 125 V are used, the maximum classification accuracy of 0.99 is obtained, as measured by the area under the corresponding receiver operating characteristic curve (AUC-ROC).

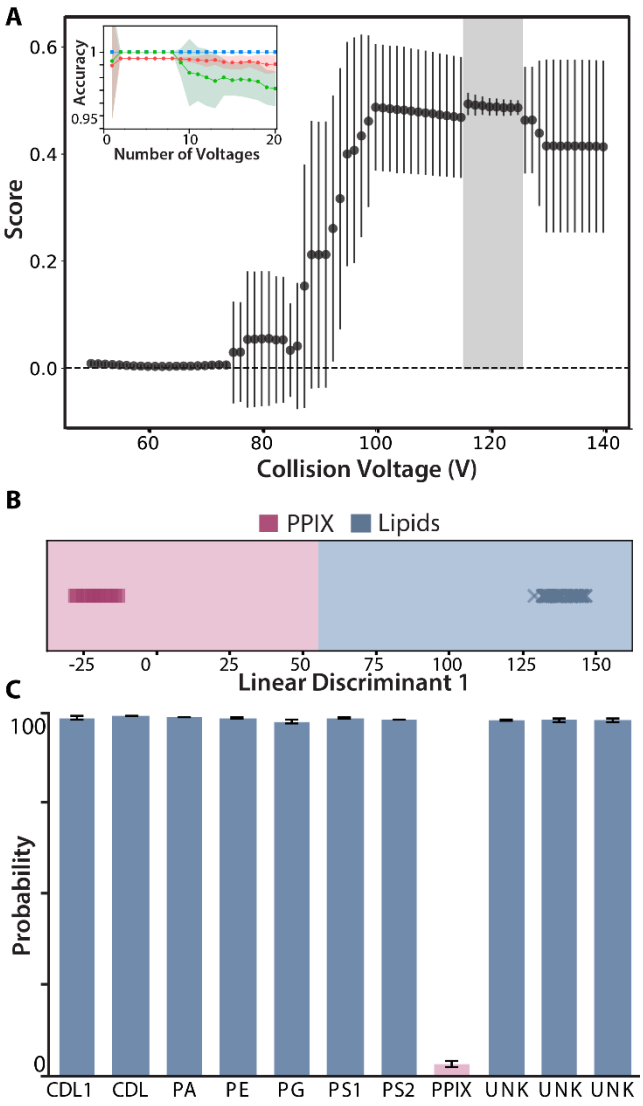

Figure 2-4. CIU based classification of TSPO ligands. **A.** We begin by scoring each voltage for its ability to differentiate all lipids from PPIX, N=3. Voltages shown within the gray box were those selected for a 'leave one out' cross-validation test, shown in the inset plot, where we tracked accuracy achieved with the training data (blue),non-training data (green), and the area under the ROC curve (red) as a function of the number of voltages included in the classification scheme. Based on these tests, two voltages (*) were selected for the final classification. **B.** A plot of linear discriminant space resulting from the classification structure described in A, illustrating clear separation of the CIU responses related to lipids from those recorded from PPIX. **C.** A bar chart tracking the probability of our class assignments for additional replicates that were not part of the dataset used to build the classification engine in A, where error bars shown reflect the assigned probability from each voltage. Three CIU replicates recorded for TSPO-UNK complexes provide high probability lipid assignments.

These eight voltages are used to build a CIU classification scheme and the results of the linear discriminant (LD) analysis are plotted on the resulting LD axis, where we observe clear separation between PPIX and lipid CIU signatures (Figure 2-4B). To further test the strength of our classification scheme, additional replicates not included in our original training set were input into our classification structure as "unknowns". We analyzed the results of this mock unknown classification experiment by assessing the probability associated with the ligand class assignment indicated by CIUSuite 2.1 (Figure 2-4C). All of our mock unknowns analyzed in this way classify correctly. In addition, data for TSPO-UNK complexes resulted in high probability lipid classifications, as predicted by the analysis of our CIU50 values displayed in Figure 2-3. The data shown in Figure 2-4 represents the first time CIU data has been used to classify ligands to previously identified disparate binding locations within a membrane protein complex, and the first instance of using CIU to identify an unknown ligand class with a quantified level of confidence.

2.3.5 CIU Differentiates TSPO:Lipid Complexes Based on Chain Saturation

Lipids can play functional roles upon binding to membrane proteins, engendering the possibility of specific binding pockets, where the detailed structure of the intended binder may affect its ability to interact with the protein.[40,124,139] The structure of lipids can include variations in head group, chain length, and chain saturation, all of which can exert biological effects.[141] Therefore, we proceeded to evaluate the ability of CIU data to classify TSPO:lipid complexes according to their specific chemical structures, finding the greatest correlations in our CIU data to lipid chain saturation levels.

As lipids are often distinguished by their head group, we began by grouping the lipids we analyzed in our study based on their overall physiological charge. However, all attempts to build

a classification scheme based on lipid physiological charge resulted in poor cross validation testing and incorrect assignment of some lipids (Figure I-8). Noting the similarities between the CIU data collected for TSPO complexes containing PE and PA (Figure I-8 C), and that TSPO possesses large hydrophobic lipid binding grooves,[14] we next focused on lipid tail attributes in order to classify TSPO:lipid complex CIU data. To construct our classification scheme, we sorted our 1:1 TSPO dimer:lipid CIU data into groups based on lipid tail saturation levels, including saturated (PG, PI), one double bond containing (PS), and two double bond containing (PA, PE) lipids. Data for CDL and PC complexes were left out of classification efforts, as when compared with the majority of the lipids studied here, they possessed either a different number of chains or exhibited evidence of a significantly altered CIU pathway, respectively. Our general procedure for constructing our three-way lipid saturation-based classification scheme

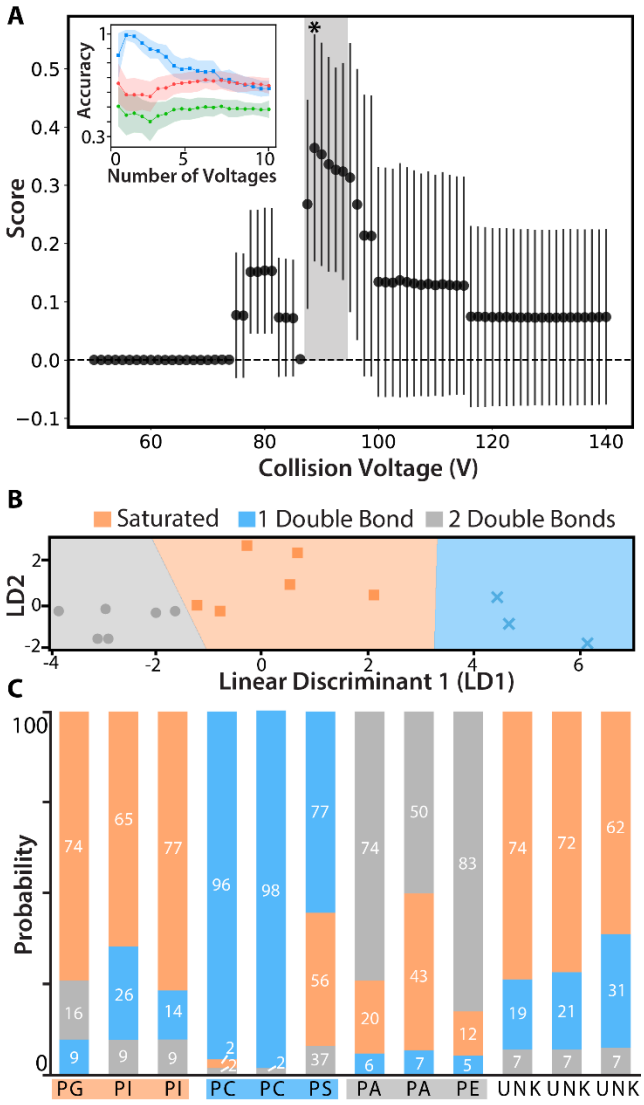

Figure 2-5. CIU based classification of lipid chain saturation in TSPO complexes. **A.** Each voltage was scored for its ability to differentiate the following three classes: saturated lipids (PG, PI), lipids with one double bond (PS) and lipids with two double bonds (PA, PE), N = 3 for each ligand. Voltages in the gray box were selected for cross-validation runs, where the inset plot is formatted identically to the one shown in Figure 4. Based on this analysis, one voltage (*) was found to provide the most accurate classification scheme. **B.** A three-way CIU classification depicted as decision regions on linear discriminant axes, illustrating the separation of all three classes. **C.** A bar chart plotting the probability of correct assignment for lipids replicates not included in the training data evaluated in A, where the true saturation of each lipid is shown using a color code as indicated. Each lipid is correctly assigned, and three replicates of UNK, suspected to have a PG head group, show the most likely assignment to be the saturated isoform, followed by isoforms containing a single double bond

was similar to the approach described above for building our two-way lipid/PPIX classification scheme (Figure 2-4). A notable difference between our lipid saturation level based classification scheme and our two-way lipid/PPIX classification structure is that the optimally differentiating region of our CIU data identified is at lower voltage values than those selected as optimal for lipid/PPIX discrimination. Our cross-validation testing demonstrated that using CIU data acquired at one voltage, 88.75 V, provides a classification scheme of greatest accuracy, resulting in a 0.72 AUC-ROC (Figure 2-5A). Although the accuracies measured in our cross-validation runs for this scheme are lower than those generated in Figure 2-4, this is expected when evaluating a three-way classification, where assignments are not binary (Figure 2-5B).

In our three-way classification scheme, the assignment probabilities for each CIU dataset are shown as a stacked bar graph in Figure 2-5C for nine datasets not included in our training data. While mock unknowns are assigned to chain saturation classes with a range of probabilities, the lowest of which is associated with a PA dataset shown, all CIU datasets are correctly assigned to the appropriate chain saturation class. This includes CIU data for PC, a sample for which no replicates were included in the training dataset. The CIU replicates for UNK can be assigned with moderate confidence as a saturated lipid (with probabilities of 74%, 72% and 62% for the three datasets acquired), where the second most likely lipid class assignment includes one degree of chain unsaturation (Figure 2-5C). The ambiguity of these assignments can be rationalized using HRMS lipid extraction data (Figure I-7). While these data identified saturated PG (16:0) as the dominant endogenous lipid form present within these TSPO samples, other PG isoforms, including many with one degree of unsaturation, were also identified. While the HRMS data cannot assign which of these lipids is bound within the TSPO:UNK complexes detected in our experiments, it is possible that UNK is comprised of a mixture of PG isoforms, containing both saturated chains and

those containing one double bond. Our proposed assignment for UNK is also in agreement with X-ray data analysis reported for a TSPO mutant.[14] Overall, the data shown in Figure 2-5 represents the first example of using CIU data to classify protein complexes in terms of the chemical structures of the bound ligands, and provides a foundation for future membrane protein complex analyses where CIU data is leveraged to aid in the identification of unknown endogenous ligands.

2.4 Conclusions

CIU is a promising technique for studying ligand binding within membrane proteins, despite challenges related to chemical noise. We have designed data analysis workflows aimed specifically at addressing these noise related challenges, and have applied these workflows to gain new insight into the ligand binding behavior of TSPO. Complexes containing ligands known to bind at two separate sites were detected as possessing differential stabilities using CIU, where PPIX binding provides a greater degree of gas-phase stabilization for TSPO than any of the lipids assessed here. While our data alone does not enable us to rule out the presence of rearranged TSPO structures under our conditions, such information suggests that a stronger intermolecular interactions exist within TSPO:PPIX complexes than when compared to equivalent lipid binding events, and furthermore indicates that TSPO ligand complexes retain a strong memory of their native-state structures in the gas-phase.

Furthermore, we constructed two CIU-based classification schemes for TSPO and observed the successful differentiation of lipid binders from PPIX, as well as sub-classes of lipids based on their levels of chain saturation. These include iso-mass lipids (e.g, POPC and POPS, both having an average mass of 760 Da), which we are able to confidently differente based on CIU classifiers. The construction of our classification schemes also revealed that level of lipid chain

saturation, rather than charge, dictates the gas-phase unfolding of TSPO. This observation is congruent with the large, hydrophobic lipid binding grooves on the surface of TSPO observed in X-ray structures for the complex.[31] By combining CIU data at the level of intact protein complexes with HRMS data collected on extracted lipids co-isolated with TSPO, we propose the identity of a detected unknown endogenous ligand to be a heterogeneous population of PG isoforms containing various states of chain saturation. Such a determination would be challenging, if not impossible, to achieve with contemporary MS/MS approaches. Clearly, rapid advancements in CIU technology are acting to position this approach within the broader structural MS field, as a uniquely flexible biophysical tool, capable of assessing the composition, stability, and structure of membrane protein assemblies within complex mixtures.

2.5 Acknowledgments

Membrane protein research in the Ruotolo lab is supported by the National Institute of Health under Grants GM105942 and S10 OD021619. The authors further thank Gabriela Grigorean (UM, Chemistry MS Facility) for her assistance with HRMS data acquisition.

Chapter 3 Ion Mobility - Mass Spectrometry Reveals the Role of Peripheral Myelin Protein Dimers in Peripheral Neuropathy

3.1 Introduction

The misfolding of membrane proteins is implicated in the mechanisms of multiple debilitating diseases, including cystic fibrosis and retinitis pigmentosa.[142–145] Often, specific sequence membrane protein mutations are associated with disease states, with variant forms exhibiting altered stability and cellular trafficking.[146] Unfortunately, due to the challenges associated with preparing and handling pure, highly concentrated membrane protein samples, detailed structural information on such targets is often lacking, especially for disease mutant forms. Further, as some membrane proteins associated with misfolding diseases have hundreds of mutations of interest,[144] there is a clear need for high-throughput methods to assess dynamic structural changes in membrane protein folding.

Peripheral myelin protein 22 (PMP22) is such a membrane protein, for which misfolding and trafficking of mutant variants have been implicated in disease.[20] PMP22 is a tetra-span integral membrane glycoprotein predominately expressed in Schwann cells which are the principal glial cells of the peripheral nervous system (PNS), where they produce myelin.[147–149] In addition to accounting for ~5% of the protein found in the myelin sheath surrounding PNS nerve axons, PMP22 is thought to regulate intracellular Ca^{2+} levels,[150] apoptosis,[151] linking of actin-cytoskeleton with lipid rafts,[152] epithelial intercellular junctions formation,[153] myelin formation,[154] lipid metabolism and cholesterol trafficking.[155] Dysregulation and misfolding of PMP22 has been identified as a key factor in multiple neurodegenerative disorders, such as Charcot Marie Tooth disease types 1A and E, as well as Dejerine-Sottas Syndrome[17,19,20,156]. Like a number of other

disease-linked membrane proteins,[157] the trafficking of PMP22 is known to be inefficient, with only 20% of the wild-type (WT) protein reaching its final membrane destination in Schwann cells.[158] Previously, it has been shown through a range of biophysical analyses that disease associated PMP22 mutations lower protein stability and further reduce trafficking, however the mechanism by which these mutant forms of PMP22 cause disease is still not well understood.[20] Additionally, a high-resolution structure of PMP22 has not yet been published.

Native mass spectrometry has recently been demonstrated to overcome sample purity and concentration barriers to reveal critical details of membrane protein structure and function.[54,95,159] Through the use of nano-electrospray (nESI), intact membrane proteins are ionized within detergent micelles or other membrane mimetics,[119,138,160,161] which can then be removed from the membrane protein ions within the instrument. This method has been used to elucidate oligomeric information,[40,116,162] complex organization,[52,163] and lipid interactions [93,96,118] of diverse membrane proteins. The addition of ion mobility separations to mass spectrometry (IM-MS), provides data on the orientationally-averaged size of analytes,[99] and enables collision induced unfolding (CIU) experiments.[100] In CIU, the energies experienced by gas phase protein ions are ramped in a step wise fashion causing unfolding. These dynamic measurements have been shown to be sensitive to ligand binding,[164,165] glycosylation,[166,167] and disulfide bonding[166] in soluble proteins, as well as selective lipid and small molecule binding in membrane proteins.[94,102,106,168] While CIU can clearly capture subtle structural changes in proteins, its performance in discerning membrane protein stability as a function of protein mutation has yet to be described.

Here, we demonstrate the ability of native mass spectrometry and CIU to detect key differences in the stability and complex formation of PMP22 variants, leading to new insights into the mechanism of PMP22 dysregulation in disease. We quantify the propensity of PMP22 to

dimerize across WT and seven disease-associated point mutations. We find that mutations associated with severe disease states form significantly more dimer than WT. Through CIU, we quantify the stability of gas phase monomeric and dimeric PMP variants and find that PMP variants bearing mutations connected to severe neuropathy exhibit the lowest relative monomer conformational stability. Interestingly, we observe that all dimers of mutant PMP22 forms are more stable than WT PMP22 dimeric complexes. We continue by comparing our results to previously published biophysical datasets and find that our monomeric PMP22 gas phase stability values correlate well with cellular trafficking data.[20] Lastly, we probe the effects of solubilization agents on PMP22 by characterizing its dimerization within sphingomyelin and cholesterol rich (SCOR) bicelles. We find that dimeric PMP22 complexes persist within SCOR bicelles and that the differences we observed in dimer abundance between WT and severe mutant protein forms are intensified. We conclude by proposing a new mechanism of PMP22 dysregulation in severe neurodegenerative diseases by which PMP22 monomers are destabilized and more prone to form poorly trafficked dimeric complexes than WT PMP22.

3.2 Methods

3.2.1 Membrane Protein Sample Preparation

PMP22 WT and mutant variants (S22F, A67T, T118M, G93R, L16P, H12Q, and G150D) were expressed in *E. Coli*. using protocols adapted from Schlebach et al.[169] Octaethylene Glycol Monododecyl Ether (C12E8) and n-Dodecyl-β-D-Melibioside (DDMB) were purchased from Anatrace. Ammonium acetate was purchased from Sigma Aldrich (St. Louis, MO). Detailed

information on the purification of PMP22, preparation of sphingolipid and cholesterol rich (SCOR) bicelles, and incorporation of PMP22 into SCOR bicelles can be found in the supporting information. For detergent-based native MS experiments, 50 µM PMP22 was buffer and detergent-exchanged simultaneously from 50 mM Tris buffer pH 8.0, 0.15% β -n-decyl maltoside (DM), 15 mM imidazole and 1 mM TCEP buffer and 0.1% β -n-dodecylmaltoside (DDM) to 0.02% C12E8 (~4 x CMC), 200 mM ammonium acetate pH 8.0 using 10 kDa Amicon Ultra-0.5 centrifugal filter units (MilliporeSigma, Burlington, MA). Detergent screening was performed following established protocols.[95] For SCOR bicelles native MS experiments, 40 µM PMP22 in 10 mM acetate buffer (pH 5.0) containing 100 mM NaCl, 0.2 % SCOR bicelles (q = 0.33), 1 mM EDTA, 5 mM TCEP, and 0.3 mM DDMB was exchanged into 200 mM ammonium acetate pH 8.0 using 10 kDa Amicon Ultra-0.5 centrifugal filter units.

3.2.2 Native IM-MS

All IM-MS data was collected using a Synapt G2 HDMS IM-Q-ToF mass spectrometer (Waters, Milford, MA), with a direct infusion nESI source set to positive ion mode. Instrument settings were tuned to dissociate solubilization agents with minimal perturbation to protein structure prior to the IM separator, including appropriately tuned settings for the source temperature (30° C), source gas flow (50 mL/min), and the sampling cone (120 V). Trapping cell wave velocity and height were 116 m/s and 0.1 V. IMS wave velocity and height were 250 m/s and 15 V. Transfer cell wave velocity and height were 300 m/s and 10 V, with an accelerating potential of 70 V used to dissociate empty micelles and bicelles. Collision cross section analysis was performed by using IMSCal-19v4, a program written in C. Theoretical collision cross sections of monomeric PMP were calculated using a previously published homology model[170] and

IMPACT.[99,130,131] Percent dimer calculations were performed at trap collision voltages of 80V to minimize the impact of surfactant related noise. Arrival times distributions of charge states that were known to be uniquely monomer or dimer, as indicated in Figure 3-1A and Figure II-3, were extracted using to a text-based format using TWIMExtract.[132] For data where excess detergent and lipid noise was present at low arrival times, the extraction window was set to begin at the drift times indicated in Figure II-8. To calculate the percent dimer, the sum of the extracted dimers is divided by the monomer sum added to the dimer sum. All CIU analyses were performed in triplicate by increasing the trap collision voltage in a stepwise manner from 5 – 150 V in 5 V increments. CIU data from the 9+ and 13+ charge states of monomeric and dimeric PMP22, respectively, were extracted using TWIMExtract[132], then processed and analyzed using CIUSuite 2.[133] Data processing included three rounds of 2D Savitzky-Golay smoothing with a window of five bins and interpolation of the collision voltage axis by a factor of four.

3.2.3 High Resolution MS Lipid Extraction Experiments

PMP22 S22F was prepared as described above for IM-MS experiments. HRMS experiments were performed on an Orbitrap Fusion Lumos Tribid MS (ThermoFisher, San Jose, CA). Instrument settings were optimized to dissociate the detergent micelle to facilitate detection of the PMP22 signal. The in-source fragmentation voltages were enabled and set to 100 V. The transfer tube temperature was also increased to 325 degrees. The ions were subjected to higher-energy C-Trap Dissociation (HCD) at an energy of 22.5% to further liberate the PMP22 from the detergent micelle. All HRMS spectra were collected at a resolution 120000 FWHM at 200 Th and with the rf frequency set to 30 %.

3.3 Results and Discussion

3.3.1 Quantifying Abundance of PMP22 Dimers

The existence of PMP22 homodimers formed by WT and mutant PMP22 variants,[169,171] as well as aggregation-like higher order species,[172–174] has been inferred from low resolution mass separations and biophysical assays. Therefore, we endeavored to use native mass spectrometry, which enables the characterization of oligomeric state through detectable differences in the m/z of protein ions, to analyze PMP22 oligomeric species. While PMP22 has not previously been studied via native MS, we find that PMP22 liberated from C12E8 detergent micelles reveals two distributions of protein related signals with charge states of 10+ to 6+ and 15+ to 11+, as shown in Figure 3-1A. The two distributions indicate that multiple protein oligomeric states are present, and the masses calculated from the charge state distributions,

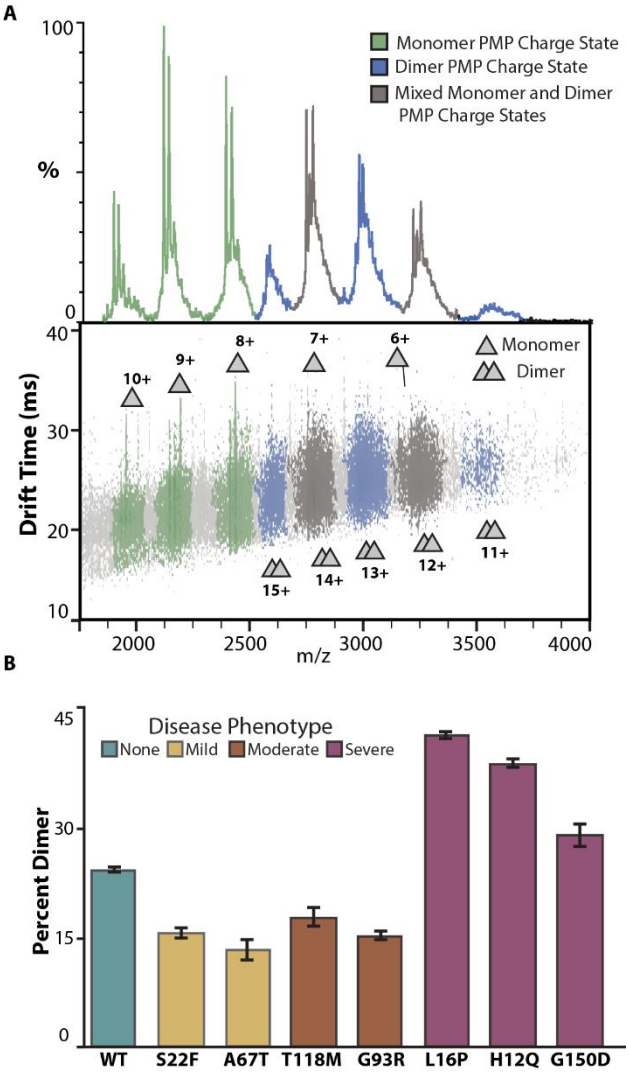

Figure 3-1. Native IM-MS of PMP22 reveals differences in dimeric abundances. **A.** Mass spectra of PMP22 liberated from detergent micelles produces populations of monomer and dimer protein at charge state ranges of 6-10+ and 11-15+, respectively. Charge states were identified as containing only monomer (green), only dimer (blue), or a mix of monomer and dimer (gray). **B.** Intensities for PMP22 signals identified as monomer or dimer were used to calculate the percent dimer in wild-type (WT) PMP22, as well as seven mutant PMP22 variants, N = 3. Notably, L16P, H12Q, and G150D, mutants correlated with severe disease phenotypes, exhibit significantly more dimer (41.0 ± 0.4%, 37.3 ± 0.6%, 28.1 ± 1.5%) than WT (23.2 ± 0.3%).

19.4 kDa and 39.1 kDa, correspond well to those expected for monomeric and dimeric PMP22 (Table S1). The data presented here is the first time a PMP22 homodimer has been confirmed and characterized. Notably, the calculated collision cross section (CCS) for monomeric PMP22 in these experiments was 1595 $Å^2$, which agrees well with the theoretical cross section calculated for folded PMP22 based on a previously published homology model of 1535 $Å^2$.[170] The CCS agreement, as well as the range of charge states observed, indicate that PMP22 monomers retain compact conformations when subjected to IM-MS analysis.

The observation of a PMP22 homodimer prompted us to quantify its relative abundance across a range of disease associated PMP22 mutants. As there is some overlap between the IM-MS signals corresponding to PMP22 monomer and dimer, in order to calculate the abundance of dimer, we filtered all signals detected using both IM and MS data leaving only those signals that could be identified as uniquely resulting from monomer or dimer PMP22 (Figure 3-1A). For example, our measurements indicate a 23.2 ± 0.3% dimer population for WT PMP22.

Of the 44 pathogenic missense mutations known to exist for PMP22, we chose seven mutations occurring throughout its sequence based on their prevalence, disease phenotype relationship, and previous biophysical assay behavior (Figure II-2). Two mutants, S22F and A67T, are associated with hereditary neuropathy with liability for pressure palsies (HNPP), which is considered a mild neuropathy.[17,175] Previous assays have found that S22F and A67T have trafficking efficiencies and folding stabilities comparable to WT PMP22.[20] Another set of mutants, T118M and G93R, are associated with Charcot-Marie-Tooth 1 disease (CMT1), a moderate form of neuropathy.[156,176] These mutations were observed to cause a reduction in cellular trafficking and stability relative to WT,[20] and recent computational modelling of T118M PMP22 indicate that the mutation causes changes to structural stability.[177] Relative quantitation of the dimer states for these

four PMP22 mutations exhibit abundances within error of each other, but significantly lower than that of WT PMP22 (Figure 3-1B).

The final three mutants analyzed in this work, L16P, H12Q, and G150D, are associated with Dejerine-Sottas Syndrome (DSS), a severe neuropathic condition. These mutations were previously found to cause a dramatic reduction in cellular trafficking for PMP22, as well as decreased stability.[20,177–179] Strikingly, we observed that all three of these mutant PMP22 forms exhibited a significantly larger amount of dimer than WT PMP22 or mutations corresponding to mild or moderate disease phenotypes (Figure 3-1B). The severe disease mutations H12Q and L16P exhibited 41.0 ± 0.4% and 37.3 ± 0.6% dimer, respectively, while G150D shows a more moderate increase in dimer relative to WT, at 28.1 ± 1.5%. As H12Q and L16P occur in the first transmembrane spanning helix and G150D in the fourth transmembrane spanning helix (Figure II-2), these differences in dimeric abundance may indicate that the mutation site Q12 and P16 residues are associated with the PMP22 homodimer interface.

3.3.2 CIU Reveals Differential Stabilities of Monomeric PMP22 Mutants

The folding and stability of PMP22 mutants have been linked to their mistrafficking, and role in neuropathies.[145] However, as a high-resolution structure of PMP22 has not yet been determined, measurements related to the folding of PMP22 have largely been made from bulk solution,[169] biochemical, biophysical and cellular assays,[20] and molecular dynamic simulations.[170,177] While chemical modification has been used to specifically study WT PMP22 dimers,[169] many reports convolute the various oligomeric states of PMP22. The gas phase unfolding of membrane proteins has been previously reported to inform on the conformational dynamics and stability of protein structures,[94,168] and it allows for independent analysis of discrete

protein forms that are separated in m/z. Advances in CIU analysis software have also enabled the automated identification of features, and the fitting of sigmoidal transitions between such identified features.[133] We proceeded to use this approach to evaluate the role of mutations in the stability of monomeric and dimeric PMP22.

The feature observed at the lowest drift time and collision voltage represents the most folded and native-like conformational family observed for the PMP22 monomer in our experiments, while features appearing at longer IM arrival times correspond to a more unfolded PMP22 conformations. Sigmoidal curves fit to the transitions between CIU features, allow us to compute inflection points, or CIU50 values, that we use to compare the relative stabilities of PMP22 variants.[133,168]

Figure 3-2A displays averaged CIU fingerprints, N = 3, with detected features (left) and sigmoidal fits (right) for 9+ monomeric PMP22 WT, as well as example mutants from each disease severity phenotype. While a CIU50 of 19.8 ± 0.4V is observed for WT PMP22, variants of PMP22

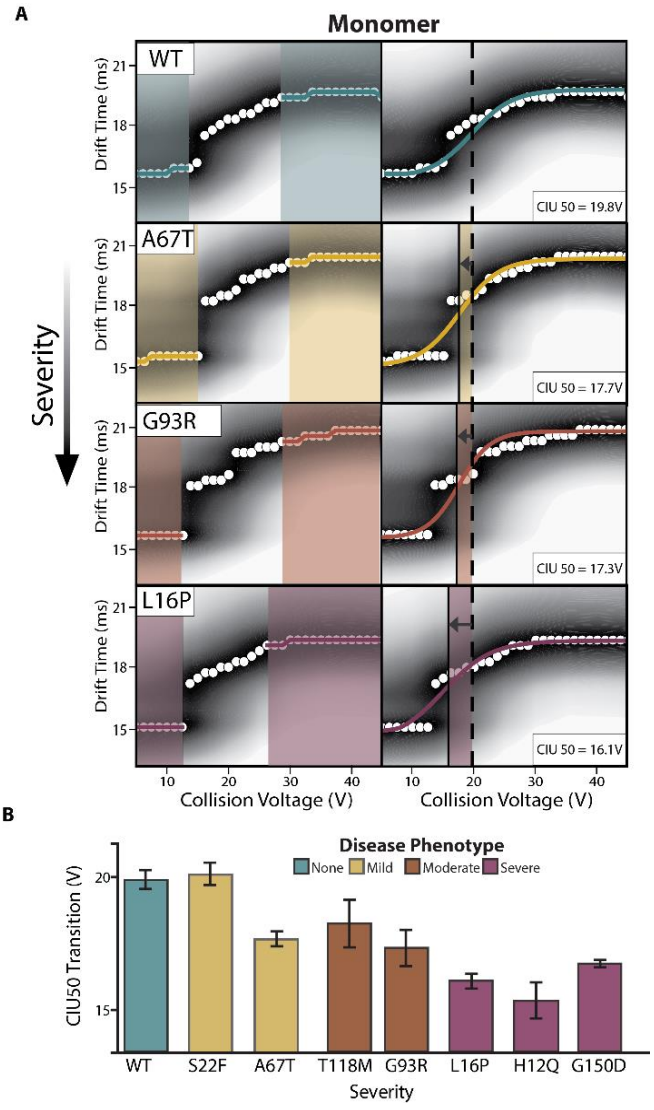

Figure 3-2. CIU data for monomeric PMP22. **A**. Analysis of 9+ monomer WT PMP22 and all mutant PMP22 forms tested identify two features between 5V and 45V. Fitting a sigmoid to these features allows for CIU50 determinations. **B**. CIU50 values recorded for monomeric WT and mutant PMP22, N = 3, ordered by the disease phenotype severity engendered by the indicated mutant.

possessing the A67T, G93R, and L16P mutations all exhibit a range of CIU50 values significantly lower than that observed for WT. These lower CIU50 values indicate that PMP22 mutants are destabilized in a manner correlated with the severity of the neuropathy they induce. A clear example of this trend is apparent upon comparing CIU data collected for WT PMP22 (teal) and L16P PMP22 (pink), where the feature appearing at longer drift times in the L16P PMP22 data begins at significantly decreased collision voltage values when compared to the analogues feature appearing in the WT fingerprint, resulting in an average CIU50 of 16.1 ± 0.3V for the mutant protein.

Destabilization is observed for all PMP22 mutants studied here, except for S22F, where each mutant protein monomer exhibits an average CIU50 value significantly below the lower bound of error defined for our WT CIU50 measurement, (Figures 3-2B and II-5). The observation that monomer S22F PMP22 exhibits a stability comparable to WT PMP22 is interesting in light of previously published data concluding that this variant may be trafficked more efficiently than WT and exhibits unique metal binding properties.[20] We note that the most destabilized class of PMP22 variants are those associated with severe neuropathy (pink). Specifically, H12Q and L16P PMP22 exhibit CIU50 values of 16.1 ± 0.3V, and 15.4 ± 0.7V, which are both significantly less than the stability we record for the WTPMP22. This observation agrees well with previous data recorded for these mutatants;[20,170,177] however, this study represents the first direct stability measurements of mutant PMP22 monomers

3.3.3 Mutant PMP22 Dimeric Complexes Are More Stable Than WT

Both While the mistrafficking of PMP22 is a common theme in its proposed roles in neuropathic disease,[145,171,180] the underlying mechanism involved, as well as the fate of proteins

which do not get properly trafficked, are still active areas of research. Proposed as either a cause or a consequence of mistrafficking, PMP22 oligomerization has been linked to the etiology and possible treatment modalities for its associated importance of PMP22 oligomers to its dysregulation we assessed the stability of PMP22 homodimers detected in our experiments as a function of protein mutation.

To study the CIU of PMP22 dimers, we selected the 13+ charge state, as it appears across all datasets and cannot overlap with monomeric ions. Automated feature analysis detected two conformational populations between 5V and 60V for all WT and mutant PMP22 dimer fingerprints. Sigmoidal fits between these two states yielded reproducible CIU50 values (Figure 3-3A). We interpreted these data similarly to the monomeric fingerprints described in Figure 3-2, where CIU50 values were compared across dimeric fingerprints for all PMP22 variants (Figures 3-3B and II-5).

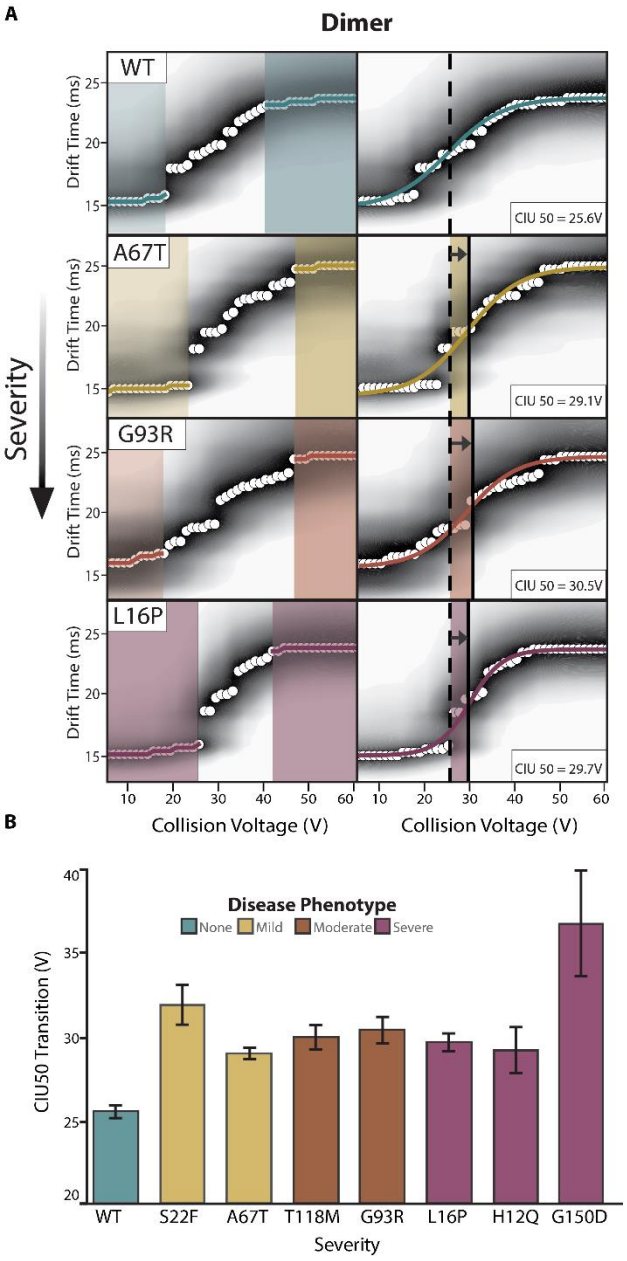

Figure 3-3. CIU data for dimeric PMP22. **A.** Analysis of the 13+ WT PMP22 dimer and all mutant PMP22 forms exhibit two features between 5V and 60. **B.** CIU50 values for WT and mutant PMP22 dimers, N = 3, ordered by disease phenotype severity.

In contrast to monomeric data, we find that all variant dimers exhibit enhanced stability relative to WT. Specifically, we record WT PMP22 dimer CIU50 of 25.6 ± 0.4V, while all mutant

variants studied here exhibit significantly larger values, ranging from 29.3 ± 1.4V to 36.8V ± 3.0V. Note that equal levels of homodimer stabilization is observed for all mutant variants, despite the differences between mutants in relative abundance of homodimeric complexes (Figure 3-1B). Additionally, the stability measurements reported by the CIU50s in Figure 3-3 reflect the unfolding of the intact dimeric complexes, not dissociation events. However, no clear differences between CIU50s of dimeric PMP22 mutants associated with different disease phenotypes is observed. Given the literature precedence for PMP22 oligomers,[172–174] our findings clearly support the notion that destabilized mutant PMP22 monomers are also more apt to form stabilized dimeric states, and this equilibrium plays a role in disease state severity. This insight is further supported by the general correlations we observe between disease severity, monomeric stability, dimer stability, and the quantitative amounts of PMP22 dimers detected in our experiments.

3.3.4 Gas Phase Stability Data Correlate with Cellular and *in vivo* Assays

To probe how our results fit with prior findings associated with PMP22 mutations, we performed a series of correlative analyses where we plotted our native MS and CIU data presented in Figures 3-1 and 3-2 against previously published cellular and *in vivo* data.[20] For these analyses, we included prior data tracking cellular trafficking efficiencies, nerve conduction velocities (NCV), and Gibbs free energy of Zn(II) binding ($\Delta\Delta G_{app,total}$), which informs directly on the solution phase folding of PMP22.

We first assessed how our monomer stability data and dimer abundance values correlated to patient NCVs, as such data presents the most direct link to disease phenotype. Previously, NCVs have been shown to correlate well with both trafficking efficiency ($R^2 = 0.88$) and $\Delta\Delta G_{app,total}$ ($R^2 = 0.64$) in an analysis of a larger set of ten PMP22 mutants.[20] While we find no significant

correlation between NCVs and the abundance of dimeric complexes measured here (Figure II-6B), we do observe a correlation between the CIU50 monomeric stability of PMP22 and patient NCVs (R^2 = 0.57, Figure II-6A). The nature of this relationship is proportional, where a lower monomeric CIU50 value is associated with slower nerve conduction velocities. We next explored relationships between the $\Delta\Delta G_{app,total}$ and our measurements, but found no significant correlations for either the PMP22 monomeric stability values or our abundance measurements for PMP22 dimers (Figures II-6D,E). We interpret this result as stemming from the foundational differences underpinning the two measurements analyzed in these separate studies. First, the prior data relies specifically on Zn(II) binding to PMP22, and Zn(II) was not present in our native IM-MS experiments. Secondly, in the CIU experiments we are extracting stability information from specific oligomeric species, where the solution phase experiments are likely sampling complex mixture of PMP22 oligomers.

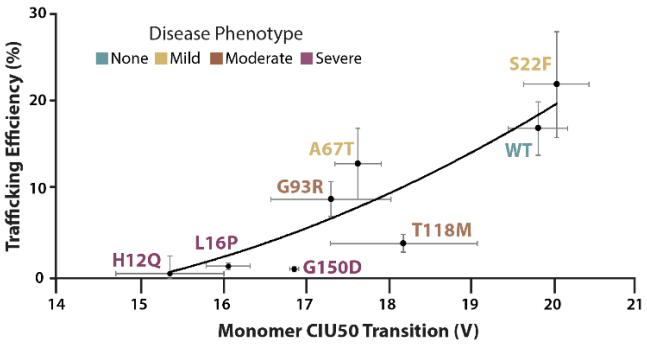

Figure 3-4. Correlation between monomer CIU50 stability values and PMP22 trafficking efficiencies, N = 3.[20] CIU50 values are from the 9+ monomeric charge state of PMP22 and describe the relative stability of the protein forms. Mutations associated with different disease severities are color coded. A quadratic regression ($y = 0.3538x^2 - 8.4629x + 47.218$) fit is shown for the data, along with errors for each dataset, producing a correlation coefficient of 0.80, which shows a positive correlation between monomeric protein stability and trafficking efficiency

We find that cellular trafficking efficiency data correlates strongly with our monomer CIU50 values (R^2 = 0.80, Figure 3-4) . The quadratic regression shown in Figure 3-4 exhibits a relationship wherein less stable PMP22 monomers are associated with reduced cellular trafficking. Additionally, we find that PMP22 trafficking efficiency is anti-correlated with the dimer abundance values recorded from our native IM-MS measurements (R^2 = 0.66, Figure II-6F). Notably, the mutants associated with severe forms of neuropathy, H12Q, L16P and G150D, cluster at the extremes of these plots, all exhibiting low trafficking efficiencies, low monomeric stabilities,

slow NCV values, and abundant dimers. The above correlations clearly indicate the importance of PMP22 dimers, an oligomeric state ultimately fed by the stability of the preceding monomers, in predicting PMP22-associated disease phenotypes.

3.3.5 Encapsulation within SCOR Bicelles Engenders Mutant PMP22 Homodimer Formation

Lipids are recognized as key regulators of membrane proteins through selective interactions or association with the membrane itself.[40,181–183] Many biochemical processes are thought to be driven by partitioning of membrane proteins into nanodomains containing saturated lipids, sphingolipids, and cholesterol.[184,185] The myelin membrane in which PMP22 natively resides is rich in cholesterol and sphingomyelin, and the protein is thought to play an important role in cholesterol homeostasis.[155,186,187] Recently, a novel bicelle construct aiming to imitate the ordered membrane environment encountered natively by PMP22 has been described.[188] These sphingomyelin and cholesterol rich (SCOR) bicelles have been found to be compatible with a wide range of biophysical measurement techniques, including native IM-MS.[188] To evaluate the effect of a more native-like lipid environment on the formation of PMP22 homodimers, we employed SCOR bicelles to prepare samples for native IM-MS analysis.

In Figure 3-5A, an example plot of IM drift time versus m/z for PMP22 L16P liberated from SCOR bicelles is shown, and we observe signals associated with monomer (green), dimer (blue), and overlapping PMP22 oligomeric states (gray). Additionally, noise signals associated with lipid and detergent clusters can be observed in the lower IM drift times and m/z values (not

shown). Due to this noise, the SCOR bicelle-related data shown in Figure 3-5C,E were background subtracted for visualization purposes, and prior to our quantitative analysis of the PMP22 dimers observed, data were de-noised (Figure II-8). MS data collected for these samples exhibited evidence of enhanced levels of adduction (Figure 3-5B-E), which likely stems for the sample preparation procedures associated with SCOR bicelles used in these experiments (also observed in Figure II-7). Figure 3-5F displays our quantitative analysis of PMP22 dimers for WT, L16P, and G150D variants liberated from micelles and SCOR bicelles. While minor differences exist in the amount of dimer observed for the mutants between our micelle and bicelle in our measurements, we detect significant differences in the relative abundance of WT dimers ejected from SCOR bicelles when compared with our micelle data, which produced values of 22.7 ± 0.3% PMP22 dimer in micelles and only 5.6 ± 0.6% dimer in SCOR bicelles. This data further demonstrates that the increased levels of dimer we detect are not an

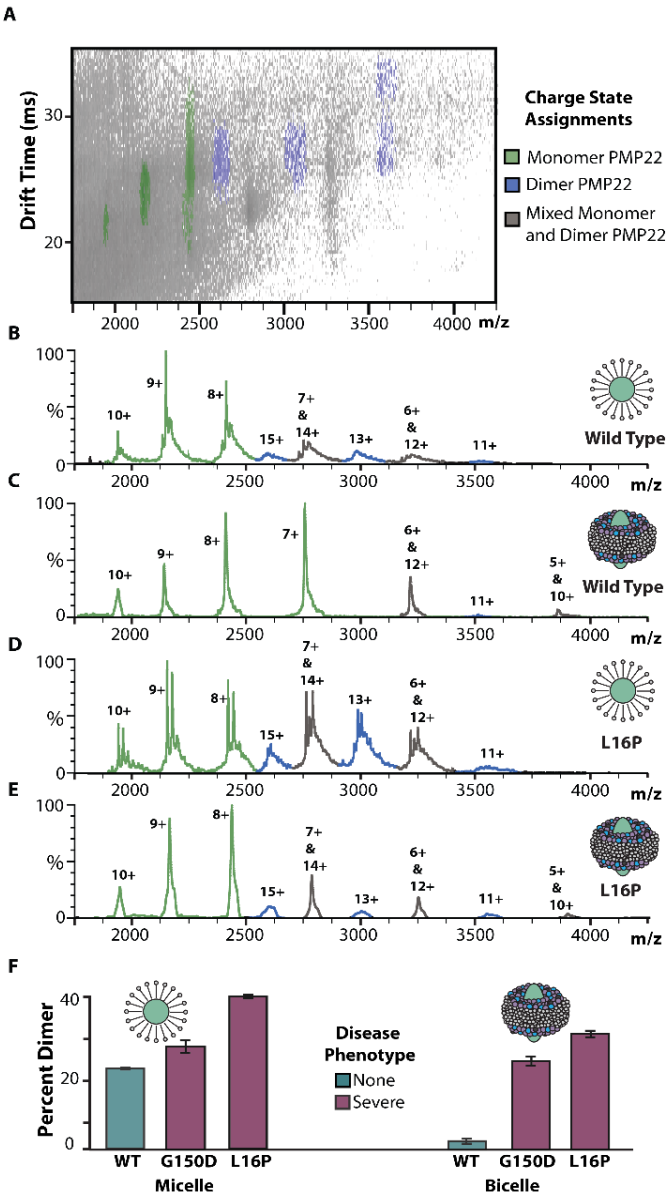

Figure 3-5. WT and L16P PMP22 liberated from SCOR bicelles. A. IM-MS of L16P liberated from SCOR bicelles reveals populations of monomer and dimer PMP22 as well as lipid and detergent related noise (low m/z noise not shown). Such noise was systematically excluded from all further analysis (Figure II-8). B-E. Mass spectra of WT PMP22 liberated from either detergent micelles or SCOR bicelles (as indicated), followed by similar data collected for L16P PMP22. Monomeric PMP22 (green) and dimeric PMP22 (blue) is present in each spectra, but in varying amounts. F. Measured percent dimer values from native IM-MS. Significant differences in these values are observed between mutants and solubilization techniques

artifact associated with the location of the mutation studied, but rather a function overall structure and stability of the PMP22 variant. Overall, these observations demonstrate the importance of the native membrane and lipid environment on PMP22 dimer formation, and are further evidence that PMP22 homodimers complexes are strongly correlated with mutations that give rise to severe neuropathies.

3.4 Conclusions

The role of PMP22 in neuropathies is complex, with its folding, stability, trafficking, and aggregation all previously indicated as important predictive links to its role in disease.[20,172,177,178,189] Native IM-MS and CIU have allowed us to identify PMP22 homodimers as a new candidate for inclusion among these factors. Our measurements enabled the stabilities of PMP22 monomers to be untangled from their associated oligomers, and they suggest that those mutations that engender the least stable monomers also produce the most stable dimers. These results indicate that those monomers that are least stable, and thus most likely to form dimers, are also those most likely to accumulate a significant dimeric population. The likelihood that these results bear relevance under native cellular environments is high, as our SCOR bicelle measurements reveal a more dramatic difference in dimer amounts in WT and disease-associated mutants. Furthermore, if these observations are combined with our analyses that highlight correlations between our results and previous measurements of PMP22 trafficking efficiency, a mechanistic link can be formed between protein misfolding, dimer production, and a loss of functional PMP22 within the cellular membrane, leading to direct consequences in NCV and the overall severity of the neuropathy experienced by patients.

While we present no mechanistic evidence that explains how the formation of PMP22 dimers may negatively impact their cellular trafficking here, we can postulate multiple potential mechanisms based on what is known about membrane protein trafficking. The trafficking of WT PMP22 is known to be tightly regulated,[158] and it is likely that the molecular machinery which recognizes properly folded PMP22 may not similarly interact with dimeric PMP22. Additionally, PMP22 dimers may be the first step in a more extensive aggregation pathway leading to disruption of neuronal function. Beyond the implications for PMP22 biology, we suggest that the methods described here could be more broadly applied to membrane proteins misfolding disorders, paving the way for future insights into a host of debilitating human ailments.

3.5 Acknowledgements

Membrane protein research in the Ruotolo lab is supported by the National Institute of Health under GM105942. HRMS research on the Orbitrap Fusion Lumos was supported by the National Institutes of Health under S10OD021619. This work was also supported by NIH R01 NS095989 granted to the Sanders and Ohi lab.

Chapter 4 Collision Induced Unfolding Differentiates Functional Variants of the KCNQ1 Voltage Sensor Domain

4.1 Introduction

Membrane proteins maintain cellular homeostasis through regulating signaling, facilitating ion transport across biological membranes, and performing a wide array of enzymatic reactions.[107] Disruption of any of these vital functions can cause cellular dyshomeostasis and disease.[190] The point mutations can significantly alter membrane protein folding, trafficking, and activity, and thus many such sequence variants have been linked to disease.[145] With recent advances in genome sequencing, more human membrane protein variants are being discovered, but the significance of these variants is not always readily apparent.[191,192] Methods for identifying structural differences in these variants is complicated by their insolubility in aqueous solutions and difficulty in obtaining pure, high concentration samples.[33–35] As there can be hundreds of variants of interest for a particular membrane protein,[144,145,193] there is an active need for high-throughput compatible methods to obtain such structural information on variants of unknown significance (VUS) and classify them according to their functional and disease associations.

Mass spectrometry (MS) is readily capable of overcoming the complexity and challenges associated with studying membrane proteins structure.[85] In particular, native MS has enabled insights into the ligand binding,[74,117,194] local lipid environment,[118] and protein-protein interactions of membrane proteins.[162] Coupling ion mobility (IM) separation with mass spectrometry (IM-MS) produces information on the orientationally averaged size of ions which allows for the evaluation of structural changes within membrane proteins.[52,122] The use of IM-MS also facilitates collision

induced unfolding (CIU) experiments, in which protein ions are subjected to ramped collision energies within the instrument, causing unfolding.[100] CIU experiments have been shown to be sensitive to lipid binding in membrane proteins,[102,168] and classification schemes built on CIU experiments,[101,195] which have differentiated disulfide bonding patterns[134] and ligand binding modes.[196] However, the use of CIU to differentiate classes of membrane protein variants associated with disease have yet to be completely explored.

Here, we report the first use of CIU to differentiate variants of the 18 kDa voltage sensor domain of the KCNQ1 voltage-gated potassium channel. KCNQ1 plays a key role in the tightly controlled process of repolarization in cardiac cells[197] and the pore opening is modulated by the tetraspan integral voltage sensor domain (VSD).[198] Specific heritable mutations in the KCNQ1 VSD are associated with losses or gains of pore activity, where a loss of function (LOF) causes Long QT Syndrome (LQTS)[199,200] and a gain of function (GOF) causes Short QT Syndrome (SQTS).[201,202] Both conditions are associated with an increased risk for cardiac arrhythmias and sudden cardiac death.[203,204] There are over 600 human variants of KCNQ1 associated with LQTS.[193] However, the significance of many of these variants and the mechanism by which the mutations cause disregulation is not well understood.

In this work we use native IM-MS and CIU to differentiate KCNQ1 VSD variants, each having different disease and functional significance. We first show that wild-type (WT) KCNQ1 VSD can be liberated from detergent micelles in a manner that retains native like folding in the gas-phase. Then we explore the CIU of WT KCNQ1 alongside three mutants: R231C which is associated with SQTS and GOF,[205,206] E115G which is associated with LQTS and LOF,[207] and H126L which was originally a VUS associated with LQTS, but has recently been shown to exhibit LOF.[146] Our feature detection and comprehensive difference analysis approach highlights changes

in the unfolding pathway of WT, GOF, and LOF KCNQ1 VSD variants, where our data support the finding that H126L behaves most similar to the LOF type. We demonstrate the ability of CIU classification to differentiate WT, GOF, and LOF KCNQ1 VSD variants and show that the H126L fingerprints classify as LOF. Further, we quantify the relative stability of the KCNQ1 VSD variants through CIU-based stability measurements and find preliminary evidence that the R231C mutation is destabilized relative to WT, which we discuss in the context of cellular data showing increased expression and cellular trafficking. Lastly, we project the utility of CIU methods in screening large numbers of VUS and discuss the future of such technologies for advancing our understanding of membrane protein function.

4.2 Methods

4.2.1 Membrane Protein Sample Preparation

WT KCNQ1 and the mutant variants H126L, E115G, and R231C were expressed in *E. Coli* and purified using protocols described elsewhere.[146] Octaethylene Glycol Monododecyl Ether (C12E8) was purchased from Anatrace and ammonium acetates was purchased from Sigma Aldrich (St. Louis, MO). For mass spectrometry experiments, 150 µM KNQ1 in buffer of 50mM MES, 0.5 mM EDTA, 2 mM TCEP, and 56 µM DDM at pH 5.5 was simultaneously buffer and detergent exchanged to 200 mM ammonium acetate buffer, pH 8.0, containing 0.02% C12E8 (~4 x CMC) using 10 kDa Amicon Ultra-0.5 centrifugal filter units (MilliporeSigma, Burlington, MA). The final protein concentration used in mass spectrometry experiments was approximately 25 µM.

4.2.2 IM-MS and CIU Experiments

All IM-MS and CIU data was collected using a Synapt G2 HDMS IM-Q-ToF mass spectrometer (Waters, Milford, MA), with a direct infusion nESI source set to positive ion mode. Our instrument settings were tuned to generate intact protein ions while completely dissociating detergent adducts prior to the IM separator, including appropriately tuned settings for the source temperature (30° C), source gas flow (50 mL/min), and the sampling cone (120 V).). Trapping cell wave velocity and height were 116 m/s and 0.1 V. IMS wave velocity and height were 250 m/s and 15 V. Transfer cell wave velocity and height were 300 m/s and 10 V, with an accelerating potential of 70 V used to dissociate empty micelles and salt clusters. All CIU analyses were performed by increasing the trap collision voltage in a stepwise manner 5 – 50 V in 5 V increments, as severe signal loss was observed at voltages above 50V for mutant KCNQ1. CIU data from the 8+ charge state of KCNQ1 was extracted into a text-based format using TWIMExtract[132], then processed and analyzed using CIUSuite 2.1.[134] Data processing included three rounds of 2D Savitzky-Golay smoothing with a window of five bins and interpolation of the collision voltage axis by a factor of four.

Protein stability values, referend to here as CIU50 values, were extracted as described previously.[168] For mutant KCNQ1, excess surfactant related noise was noted in the lower drift time region from 35 - 50V. This noise was removed from the text-based files manually to eliminate any biased contributions to RMSD analysis, as shown in Figure IV-2. To ensure the noise removal did not alter the reported protein unfolding trajectory, CIU50 analysis was performed before and after noise removal, and reported CIU50 values were identical. Classification of these denoised data was performed using the "All Data" mode of CIUSuite 2.1, with automatic feature selection. The algorithms that enable our CIU classification analysis are described in detail elsewhere.[101,134,168]

4.3 Results and Discussion

4.3.1 IM-MS of KCNQ1 VSD

Sequence variants associated with toxic LOF or GOF occur throughout the sequence of the KCNQ1 VSD, as shown in Figure 4-1A.[146] For our assay, we chose mutations from disparate locations in the proteins sequence, with E115G in soluble helix S0, H126L in membrane spanning helix S1, and R231C in membrane spanning helix S4. The KCNQ1 potassium channel functions as a tetrameric complex, where four pore domains interact to form a functional channel.[198] The tetraspanning VSD of KCNQ1 is connected to the pore domain by a flexible linker between helices S4 and S5, as well as a 99 residue soluble domain before helix S1. The VSD is positioned at the periphery of the pore, and as such is not thought to be heavily involved in pore oligomerization.[198]

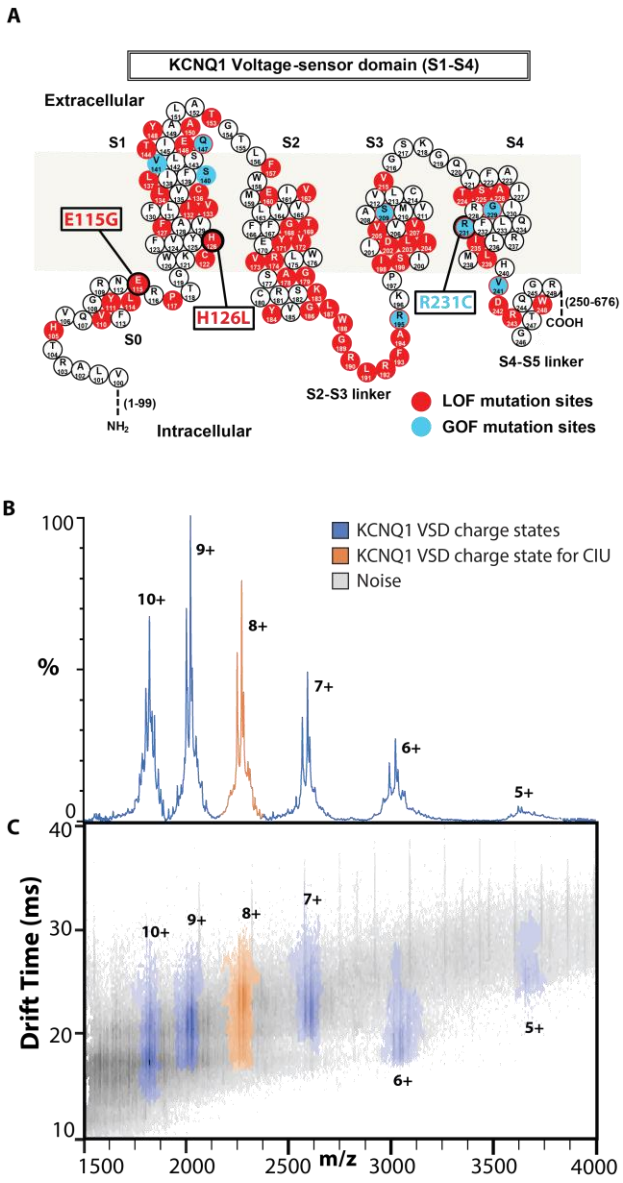

Figure 4-1. Native IM-MS of the KCNQ1 VSD. **A.** Sequence and structure of the KCNQ1 VSD with residues known to mutate and produce disease phenotypes associated with LOF (red) or GOF (blue). Variants E115G, H126L, and R231C are studied in this work. **B,C.** Representative mass spectra and IM-MS data for WT KCNQ1 at 80V trap collision voltage. One distribution of charge states, 5-10+, is observed for all variants corresponding to KCNQ1 VSD monomers (Figure IV-3). The charge state envelope and compactness of signals in IM space indicate native like folding, and the 8+ charge state (orange) was chosen for further analysis.

Figure 4-1B shows an example native IM-MS dataset for WT KCNQ1, where a range of signals can be detected after successful dissociation of the detergent from the associated membrane protein of interest. MS analysis and charge state

assignment of these signals reveals an ion population with an intact mass of 18208 ± 44 Da, which agrees well with the expected sequence mass of a KCNQ1 VSD monomer, 18158 Da. The range of charge states we observe for KCNQ1, 5-10+, as well as the compactness of the signals in the two-dimensional IM-MS data set, indicate native like structure for the gas phase protein (Figure 4-1C, IV-3). Notably, for all four sequence variants studied here we observe monomodal IM distributions at low collision energies for KCNQ1 VSD and find no evidence of oligomeric species (Figure IV-1). As has been documented for other IM-MS experiments of membrane proteins,[40,78,119,138,168] we observe noise signals related to detergent and salt clusters that persist despite sample and instrumental optimization (Figure 4-1C). These noise signals were more intense in the IM-MS datasets recorded for the mutant protein forms, perhaps due to lower protein concentrations used in these experiments in comparison to those involving WT. The 8+ charge state of the KCNQ1 VSD (orange) was chosen for subsequent CIU experiments as it exhibits a good balance between high signal intensity, low noise overlap, and low charge state.

4.3.2 Differences in CIU features correlate to KCNQ1 VSD variant function

CIU analysis often focuses on defining regions of stability, or features, within the recorded data. CIU feature analysis has been used to characterize domain level unfolding[208] and noncovalent interactions[106] within proteins, and we employed such an approach to interrogate the CIU data collected for the four KCNQ1 VSD sequence variants discussed above. CIU fingerprints of the KCNQ1 VSD 8+ charge state were collected in triplicate using a collision voltage range of 5-50V. While the interference from noise signals was minimal for WT KCNQ1 VSD, each variant dataset contained noise appearing primarily at larger collision voltage (CV) values and at drift times shorter than those occupied by the protein ion signals tracked in our experiments (Figure IV-

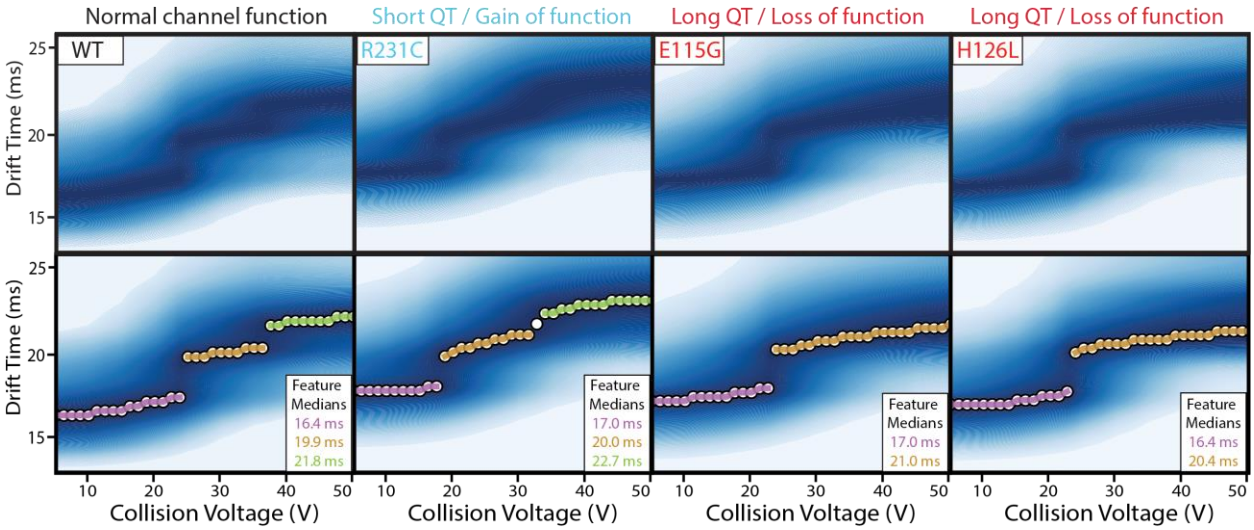

Figure 4-2. CIU fingerprints of KCNQ1 VSD variants. (Top) Fingerprints were collected for the 8+ charge state in triplicate for WT, R231C, E115G, and H126L variants, from 5-50V, and then denoised and averaged to produce the images shown here. (Bottom) Automated feature detection of the fingerprints finds three discrete features between 5-50V for WT and R231C, and two discrete features for E115G and H126L. The similarity in starting drift times of the four fingerprints indicate all forms begin at similar orientationally averaged sizes.

2). These noise signals did not alter the CIU features detected in our fingerprints; however, to eliminate any bias in our workflow, detergent noise was uniformly removed as described in the methods section for all datasets interrogated in this study. Figure 4-2 shows the de-noised CIU fingerprints (top) and detected features (below) for each KCNQ1 variant. For WT KCNQ1 VSD fingerprints, three features are observed with median drift times of 16.4 ± 0.3 ms, 19.9 ± 0.0 ms, and 21.8 ± 0.0 ms. Similarly, the GOF variant R231C fingerprints also exhibit three features with median drift times comparable to those observed for WT at 17.0 ± 0.3 ms, 20.0 ± 0.3 ms, and 22.7 ± 0.2 ms respectively. The two LOF variants, E115G and H126L, show two features each at drift times of 17 ± 0.0 ms and 21.0 ± 0.0 ms, and 16.4 ± 0.3 ms and 20.4 ± 0.5 ms, respectively.

The most compact CIU feature occurs at similar starting drift times for each KCNQ1 VSD form studied here, which indicates that there are no significant differences in the orientationally averaged size of the gas phase structures at low collision energies. However, differences were detected in the overall number of features occurring between 5-50V across the mutants studied here, where specifically we detect two CIU features over the above voltage range for E115G and

H126L, and three features for WT and R231C. The similarity observed in the CIU data recorded for E115G and H126L is further apparent in the similar range of voltages values over which their features persist, as seen in Figure 4-2. These results mimic previous cellular assays which classify these mutants together in terms of their LOF severity.[146]

4.3.3 Quantifying differences in CIU fingerprints of KCNQ1 VSD variants through comprehensive analysis (RMSD)

To further investigate the relative differences between KCNQ1 VSD variants, we performed comprehensive difference analyses using a root mean squared deviation (RMSD) approach. Such analyses have been used previously to quantify differences detected in CIU data recorded for antibodies based on glycosylation state.[166] Figure 4-3A shows a difference plot for the averaged CIU fingerprints recorded for E115G and H126L KCNQ1 mutants. We find an overall RMSD across these CIU fingerprints to be 6.6%, which is similar to RSMD values encountered for comparisons between KCNQ1 technical replicates alone (a range from 6.4% to 9.5%). In contrast, Figure 4-3B shows a similar CIU difference plot generated for averaged H126L and R231C fingerprints. In this comparison, the regions of differences are more pronounced, with stronger R231C signal apparent at longer drift times relative to H126L. Unsurprisingly, we find a larger overall RMSD for this comparison, resulting a value almost three times that for the analysis shown in Figure 4-3A, at 18.6%.

A comprehensive comparative analysis of our KCNQ1 variant CIU datasets reveals those mutants that are most and least similar to one another (Figure 4-3). In this cross-comparison, we see differences 2-3 times greater than the baseline RMSD values computed through the analysis of technical replicates of identical samples (think bordered boxes, values range from 6.4% for WT

to 9.5% for H126L), as indicated by the color scale in Figure 4-3C. Specifically, we observe no significant differences for the CIU data recorded for H126L and E115G, while the RMSD recorded for the comparisons of H126L and R231C CIU data is significant. Interestingly, the datasets exhibiting the largest RMSD value in our cross-comparison CIU analyses are WT and R231C, producing a value of 24.8%, despite the similarities observed between these data in Figure 4-2(also see Figure IV-4). A careful analysis of our comparative CIU data indicates that R231C is destabilized relative to WT and other KCNQ1 VSD variants. Overall, the comprehensive difference analysis shown in Figure 4-3C indicates significant differences between R231C and all other variants, as well as similarities between E115G and H126L. These results align with the fact that R231C is the lone GOF mutant studied here, as well as prior trafficking and cellular expression data recorded for this variant (Figure IV-5).

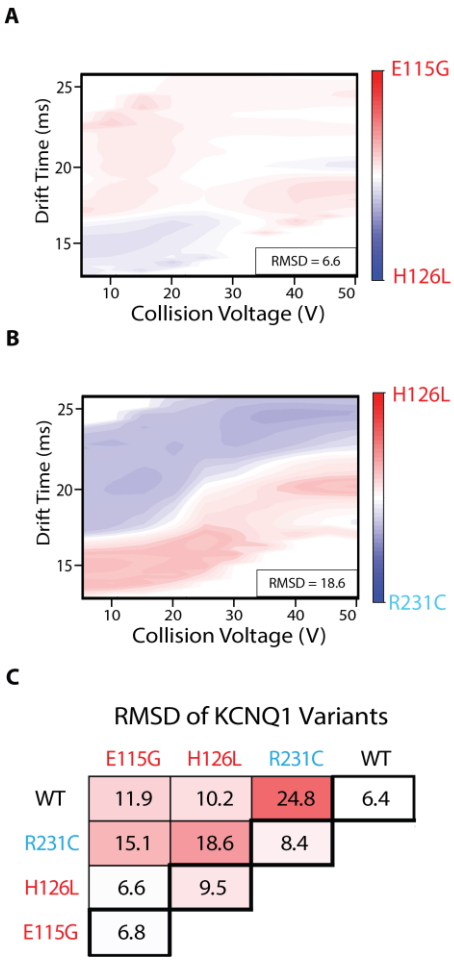

Figure 4-3. Comprehensive difference analysis of KCNQ1 VSD variants CIU fingerprints, N = 3 for each variant. **A, B.** Example difference plots of a low RMSD comparison (E115G and H126L, 6.6%) and high RMSD comparison (H126L and R231C, 18.6%). **C**. Pairwise comparisons of WT, E115G, H126L, and R231C. RMSD baseline values are shown in the thick bordered boxes. Differences of 2-3x above the baseline are considered significant. All difference plots are shown in Figure IV-4.

4.3.4 CIU Classification of phenotypic function

As shown in Figure 4-1A, there are many possible disease-associated variants of the KCNQ1 VSD. Additionally, some former VUS forms of KCNQ1 were found to exhibit WT-like function, indicating they may not contribute to disease.[146] To probe the ability of CIU to differentiate these phenotypes, we endeavored to build and test a three-way CIU classification scheme with the KCNQ1 variants studied here, grouped into the following classes: WT, gain of function (GOF), and loss of function (LOF). We started with at least three denoised replicates each of WT, E115G, and R231C KCNQ1 VSD used in the training dataset to represent each functional phenotype: WT, LOF, and GOF, respectively (Figure 4-4A). Our efforts identified the most differentiating region within the datasets as those between 10-20V, and cross validation revealed that classifiers using only the 15V datasets produced the greatest assignment accuracy (0.81 AUC-ROC, Figure 4-4A inset).

A classification scheme was built using the chosen 15V data and the results plotted linear

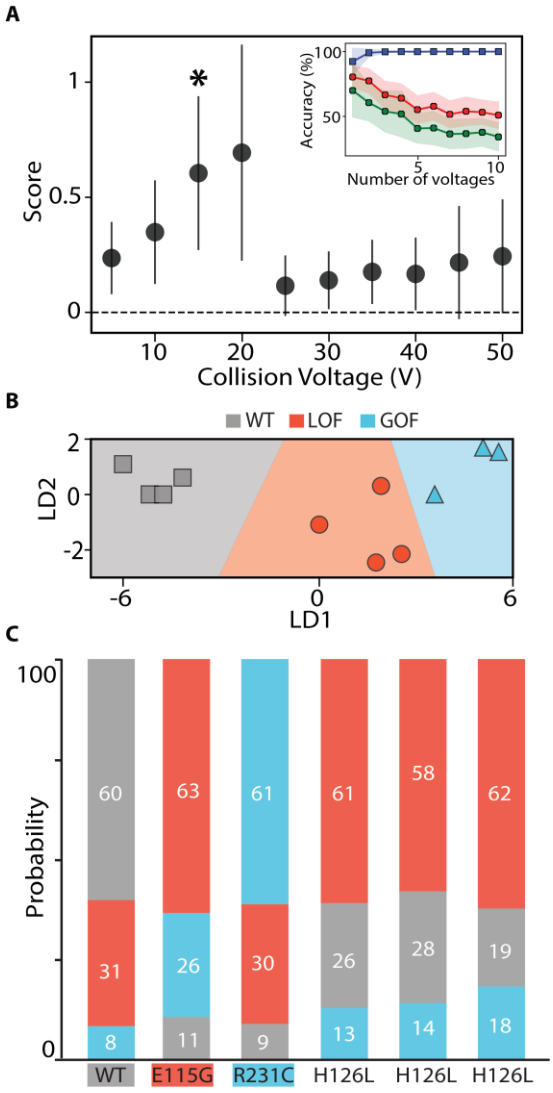

Figure 4-4. CIU based classification of KCNQ1 VSD functional variants **A.** Each voltage is scored on its ability to differentiate the three classes: WT, LOF, and GOF, using at least three WT, E115G, and R231C replicates. All voltages were used in a 'leave one out' cross-validation test, shown in the inset plot, where we tracked the accuracy achieved with the training data (blue),non-training data (green), and the area under the ROC curve (red) as a function of the number of voltages included in the classification scheme. These tests indicated one voltage (*) was best for classification **B.** Using the voltage indicated in A, the training data set is plotted in linear discriminant and shows clear separation of the data into the three classes. **C.** The probability of assignment for replicates not part of the training data set is displayed in a bar chart. Each replicate is correctly assigned, including replicates of the H126L variant for which no example was included in the training set.

discriminant space in Figure 4-4B, where gray data points indicate the WT phenotype, red indicate the LOF phenotype, and blue indicate the GOF phenotype. To test the performance of our classification scheme, we input CIU replicates to our classifier that were not part of the training dataset as unknowns. The classification scheme was then used to analyze these unknowns, and the resulting probabilities associated with assigning each of these datasets to one of the three KCNQ1 classes defined above is shown in Figure 4-4C. Data for WT, E115G, and R231C all classify as expected, producing an assignment probability of at least 60% for the correct class. In order to further test our classifier, we employed it to analyze CIU data collected for the H126L variant, none of which was in our training data. Our H126L CIU data classifies, as expected, as an LOF KCNQ1 variant,[146] with a total class assignment probability of 60 ± 2%. As such, our results both support prior assignment of H126L as an LOF KCNQ1 variant, and demonstrate the ability of CIU fingerprints to classify KCNQ1 VSD mutants based on their associated phenotypic function.

4.3.5 CIU reveals evidence of mutant destabilization

To further probe and quantify differences among KNCQ1 VSD variants relative to WT, we performed CIU50 analysis in order to quantify stability differences among the mutants studied here. Due to the different number of CIU features detected across KCNQ1 VSD variants, WT and R231C fingerprints were fit with two sigmoidal curves and E115G and H126L fingerprints were fit with one sigmoidal curve, as shown in Figure 4-5A, during our analysis. As such, all comparisons here will focus on the first CIU50 value recorded for WT and R231C variants. Figure 4-5B summarizes the average CIU50 values extracted for the first CIU transitions shown in Figure 4-5A. We record the stability of the WT as 25.8 ± 1.5V, which is within error of the stability values we observe for the two LOF mutants E115G and H126L at 24.8 ± 0.3V and 24.0 ± 1.3V

respectively. The GOF mutant, R231C, exhibits CIU50 values for the first transition that occur at significantly lower energies relative to the other three variants, producing a value of 17.7 ± 1.6V. In addition, the CIU50 we record for the second CIU transition for R231C also occurs at lower average voltages than WT, but the experimental error we compute for these values renders the difference insignificant , with R231C producing a value of 30.0 ± 4.9V and WT having a CIU50 of 37.0 ± 3.3V.

Overall, we take these CIU50 data to support the observations made in Figures 4-2 and 4-3, in that the GOF mutant R231C produces CIU data that is significantly different from that observed from either WT or LOF KCNQ1 mutants, primarily driven by the destabilization of R231C. This data is interesting to consider in the context of cellular assays which demonstrate that R231C KCNQ1 is a super-trafficker, and is highly-expressed relative to the other variants studied here (Figure IV-5). It is important to note that the construct used in our MS assays only includes the VSD of KCNQ1, whereas cellular assays track the entire KCNQ1 sequence, which is tetrameric in its native state. As such, we propose that the destabilization of the GOF R231C VSD variant may be somewhat ameliorated in the tetrameric state of KCNQ1, the

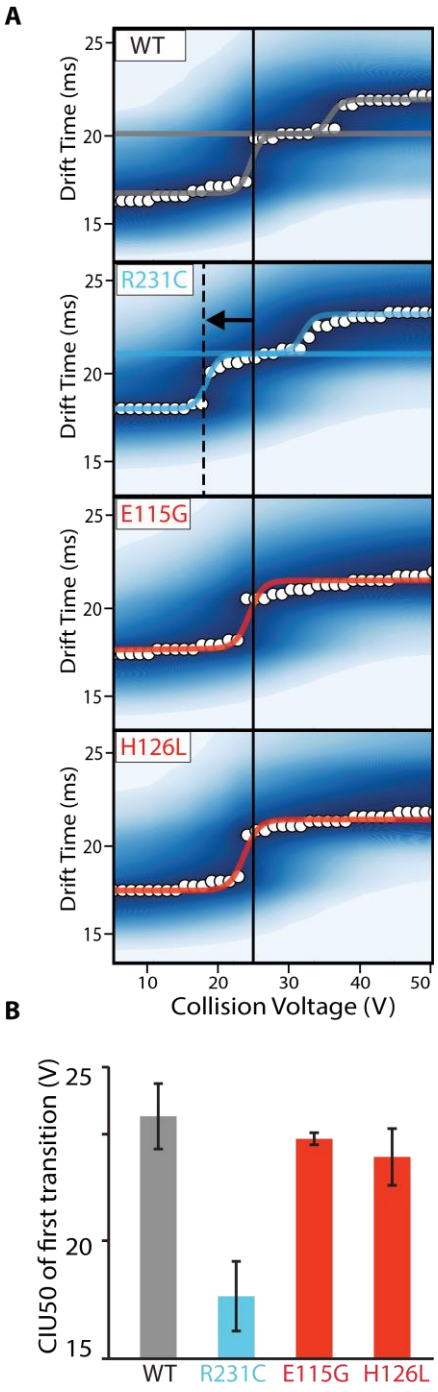

Figure 4-5. CIU50 stability analysis of KCNQ1 variants. **A**. Sigmoidal curves are fit between identified features to describe the transition, where WT and R231C have two transitions and E115G and H126L have one transition**. B**. A bar chart of average CIU50 values extracted from the inflection point determined for each first transition detected, N = 3. Gray is WT, blue is GOF, and red is LOF.

formation of which would be favored at high cellular concentrations. In addition, the protein-protein interactions formed with cellular trafficking machinery may not depend on the VSD region of KCNQ1. Future native IM-MS assays of full length KCNQ1 may further reveal how the destabilization of R231C KCNQ1 results in increased trafficking and GOF.

4.4 Conclusions

CIU is amenable to the characterization of membrane proteins and many past studies have focused mainly on its ability to characterize lipid binding behavior in such systems.[94,106,168] Here we describe a workflow for using CIU to classify membrane protein variants according to their associated functional role within a given phenotype using the KCNQ1 VSD as a model system. We first performed comprehensive difference analysis on three types of KCNQ1 VSD variants: WT, GOF, and LOF, and found strong evidence to support clustering our CIU data along the known functional consequences associated with each variant analyzed. We then used an established machine learning approach to build a classification scheme using CIU data recorded for example WT, GOF, and LOF variants, and found that this approach was reliably able to group further CIU data into classes based on the above phenotypes. Lastly, we explored the stability differences associated with these variants using CIU50 analysis and found evidence that the GOF R231C mutant is destabilized relative to WT and LOF variants for the KCNQ1 VSD.

Additionally, Native IM-MS is a method compatible with high throughput analyses, as is our classification workflow. Using the methods presented here, we estimate that the native IM-MS data collection needed to classify a VUS with our scheme could be performed in triplicate within 6 minutes. Clearly, our KCNQ1 classification scheme would be strengthened by the addition of more known variants for each class prior to testing such high throughput analyses, where we

envision that only the most differentiating regions of CIU space need be collected for each unknown interrogated. We also project that this same type of CIU classification scheme could be deployed to detect the five LOF classes described previously for KCNQ1.[146] Overall, we project that the workflow presented here could be expanded to broadly classify KCNQ1 VUSs, or be expanded to accommodate a wide range of membrane protein systems possessing variants that lack functional annotation.

Chapter 5 Systematically Studying Membrane Proteins Liberated from Different Solubilization Techniques

5.1 Introduction

Membrane proteins are important biological and pharmacological analytes that play key roles in maintaining cellular homeostasis and represent 60% of current therapeutic targets.[3,107] Membrane proteins exist in complex cellular environments where their function can be modulated by interactions with other proteins and lipids.[141,209] Due to their complex native environments and intrinsic hydrophobicity, the purification of membrane proteins requires the use of solubilization agents and often results in low protein yields.[35] These factors make obtaining high resolution structural information on membrane proteins difficult and they are currently underrepresented in structural databases in comparison to their soluble counterparts.[33,34] Further, the function and structure of membrane proteins has been shown to be sensitive to the solubilization technique used,[69,210,211] resulting in increased use of more native-like lipid mimetics for membrane protein solubilization.

Recently, native mass spectrometry (nMS) has been demonstrated capable of analyzing membrane protein samples from a variety of solubilization methods to reveal structural and functional information.[54,85] In nMS workflows, membrane proteins undergo nano-electrospray ionization (nESI) while protected by their solubilization agents before being collisionally ejected from these agents inside the instrument. While the majority of the work in this field has been performed by liberating gas phase membrane proteins from detergent micelles,[40,78,93,96,97,122,137,168,194] membrane proteins housed in other solubilization techniques such

as amphipols,[119] bicelles,[119] nanodiscs,[52,118,162,212,213] styrene maleic acid lipid particles (SMALPs),[214] and biological membrane vesicles[161] have all been successfully detected through nMS. The use of these techniques with nMS has enabled the study of elusive membrane protein characteristics such as annular lipid binding,[118] oligomerization pathways,[212] and complex formation which span both the inner and outer cellular membranes.[215] The addition of ion mobility (IM) separations to mass spectrometry allows for the measurement of the orientationally averaged size of gas phase membrane proteins,[99] and also enables collision induced unfolding experiments (CIU). In CIU experiments, the protein is collisionally activated by the stepwise ramping of an accelerating potential, which causes it to unfold. The trajectory by which a gas phase analyte unfolds can be tracked through the generation of CIU fingerprints, and recent software advancements have presented workflows for the automated analysis of CIU fingerprints to inform on the relative gas phase stability of ions.[133] While multiple studies have used IM-MS and CIU to elucidate the structural changes and relative stability of membrane protein complexes,[76,94,106,122,136,168,216] to our knowledge the use of CIU on membrane proteins has been previously limited to membrane proteins solubilized in detergent micelles.

Here, we employ IM-MS and CIU to systematically study membrane proteins solubilized by multiple techniques. We first describe the use of IM-MS to screen nanodisc lipid conditions towards the optimization of dimeric complex incorporation for the 18 kDa integral translocator protein (TSPO). Given the ability of IM-MS and CIU to access structural information about membrane proteins, we our second set of experiments probe the ability of IM-MS and CIU to capture differences in the membrane proteins associated with their solubilization agents. We do this through collecting CIU fingerprints for model protein systems which have been solubilized using at least two different methods, including detergent micelles, sphingomyelin and cholesterol

rich (SCOR) bicelles,[188] and lipid nanodiscs. Our three model protein systems, chosen based on their representation of diverse membrane protein structure, are the monotopic cytochrome P450 3A4 protein, the dimeric L16P variant of the integral peripheral myelin protein (L16P PMP22), and a form of WT PMP22 which contains a large, 11 kDa soluble tag (WT_{tag} PMP22). We find that significant differences exist between all CIU datasets collected for these proteins as a function of their solubilization agents. Further, we find that these differences emphasize the importance of careful evaluation of solubilization agents on a protein to protein basis to best preserve gas phase structure. We conclude by discussing how the data presented should be considered during membrane protein nMS optimization and proposing further experiments to elucidate the role of solubilization agents in gas phase protein structure.

5.2 Methods

5.2.1 Membrane Protein Sample Preparation

TSPO bearing the A138F point mutation was purified and expressed using established protocols from *Rhodobacter sphaeroides*.[128] CYP 3A4 was expressed in *E. Coli* and purified using protocols described elsewhere.[217–219] PMP22 WT_{tag} and the L16P mutant variant were expressed in *E. Coli*. using protocols adapted from Schlebach et al.[220] For WT_{tag} PMP22, the protein did not undergo the final thrombin cleavage step. Octaethylene glycol monododecyl ether (C12E8) and n-Dodecyl-β-D-Melibioside (DDMB) were purchased from Anatrace, Octyl β-D-glucopyranoside (OG), membrane scaffold protein 1D1(-), ammonium acetate, sodium chloride, tris(hydroxymethyl)aminomethane (tris), sodium azide, and ethylenediaminetetraacetic acid (EDTA) were purchased from Sigma Aldrich (St. Louis, MO). The lipids 1,2-dioleoyl-sn-glycero-

3-phosphoethanolamine [DOPE], 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine [POPC], 1,2-dipalmitoyl-sn-glycero-3-phospho-L-serine [DMPS], and 1,2-dimyristoyl-sn-glycero-3-phosphocholine [DMPC] were purchased from Avanti Polar Lipids (Alabaster, AL). All membrane proteins were screened for appropriate detergent conditions.[85]

Samples housed in detergent micelles were simultaneously detergent and buffer exchanged using 10 kDa Amicon Ultra-0.5 centrifugal filter units (MilliporeSigma, Burlington, MA), except TSPO which was exchanged using a 100 kDa Amicon Ultra-0.5 centrifugal filter units (MilliporeSigma, Burlington, MA). Starting and ending buffers and detergents conditions prior to native MS are as follows: 30 μM TSPO was exchanged from 50 mM Tris, 150 mM NaCl , 0.20% DM (N-dodecyl β-D-maltoside), pH 8.0 to 40 mM OG , 200 mM ammonium acetate, pH 8.0, 50 μM PMP22 was exchanged from 50 mM Tris, 0.15% DM, 15 mM imidazole, and 1 mM TCEP, and 0.1% DDM, pH 8.0, to 0.02% C12E8 (~4 x CMC), 200 mM ammonium acetate, pH 8.0, 36 μM CYP 3A4 was exchanged from 40 mM potassium phosphate, 20% glycerol, pH 7.4 to 40 mM OG, 200 mM ammonium acetate, pH 7.4. Samples housed in SCOR bicelles were prepared as described in the supplemental methods and were then buffer exchanged using 10 kDa Amicon Ultra-0.5 centrifugal filter units. Bicelle samples were not detergent exchanged and the DDMB concentration was held at 1 x CMC to preserve the bicelles q ratio (0.33) Specifically, 40 μM PMP22 in 10 mM acetate buffer (pH 5.0) containing 100 mM NaCl, 0.2 % SCOR bicelle (q = 0.33), 1 mM EDTA, 5 mM TCEP, and 0.3 mM DDMB was exchanged into 200 mM ammonium acetate, 0.3 mM DDMB, pH 8.0 using 10 kDa Amicon Ultra-0.5 centrifugal filter units. TSPO samples were incorporated into nanodiscs as described in the supplemental methods and were then buffer exchanged using Micro Biospin6 spin columns (BioRad, Hercules, CA) from standard MSP buffer to 200 mM ammonium acetate, pH 7.4. Add DMPC, DPPS. CYP 3A4 samples were

incorporated into POPC nanodiscs with the scaffold protein MSPE3D1 using a microfluidic device.[221]

5.2.2 Native MS and CIU Experiments

All IM-MS and CIU data were collected using a Synapt G2 HDMS IM-Q-ToF mass spectrometer (Waters, Milford, MA), with a direct infusion nESI source set to positive ion mode. Our instrument settings were tuned for each protein system and mimetic to generate intact protein ions while completely dissociating detergents, lipids and scaffold protein prior to the IM separator, including appropriately tuned settings for the source temperature (30-40° C), source gas flow (50 mL/min), and the sampling cone (120 V). The traveling wave height and wave velocities in the trap, IM, and transfer region, as well as the helium cell flow rate, were identical for each protein system across mimetics. For TSPO, the trapping cell wave velocity and height were 116 m/s and 0.1 V. IMS wave velocity and height were 700 m/s and 32.5 V. For PMP and CYP, trapping cell wave velocity and height were 116 m/s and 0.1 V, IMS wave velocity and height were 250 m/s and 15 V, transfer cell wave velocity and height were 300 m/s and 10 V. An accelerating potential of 70 V in the transfer region was used to dissociate empty solubilization agents for all systems except CYP in nanodiscs, which only required 10 V. Experimental collision cross section analysis was performed by using IMSCal-19v4, a program written in C, and, where possible, theoretical cross sections were calculated from crystal structures[14] and homology models[178] using IMPACT.[99,222] All CIU analyses were performed by increasing the trap collision voltage in 5 V increments across ranges tuned for each system. CIU data from selected charge states were extracted into a text-based format using TWIMExtract,[132] then processed and analyzed using CIUSuite 2.1.[133,221] Data processing included two or three rounds of 2D Savitzky-Golay smoothing with a window of five bins and interpolation of the collision voltage axis by a factor of four.

5.3 Results and Discussion

5.3.1 IM-MS of TSPO from micelles and nanodiscs

TSPO is a penta-span integral mitochondrial protein proposed to be involved in cholesterol transport and lipid transport.[127,223] This protein is highly expressed in areas of inflammation and is associated with multiple disease states, including traumatic brain injuries and cancer.[16,129,224] TSPO is thought to perform its function as a homodimeric complex and previous native mass spectrometry studies have identified lipids endogenously bound to the dimer protein.[168,194] Although it was found to bind multiple lipids across different lipid classes,[168] all previous nMS studies of TSPO have been performed using detergent micelles. Given its proposed roles in lipid transport, we considered TSPO to be a good model system for investigating the effects of solubilization agents on nMS datasets. Here, we report the first nMS studies of TSPO liberated from membrane mimetics, describe how nMS can be used to inform optimization process for obtaining biologically relevant protein complexes and discuss the noise related challenges presented by each method.

Figure 5-1A, B shows example nMS and IM-MS datasets for TSPO liberated from micelles composed of the detergent OG. From these spectra, monomeric and dimeric TSPO can be observed across the charge state ranges of 4 - 7+ and 8 – 11+, respectively. The CCS for 9+ TSPO dimer, 2751 $Å^2$, agrees well with the expected CCS based on the available X-ray structure, 2741 $Å^2$ (PDB 4UC1).[14] While there is literature detailing the purification of TSPO into detergent micelles,[128] to our knowledge no protocol has been optimized for incorporating TSPO into a nanodisc. We screened a range of lipid conditions for TSPO incorporation and found that despite optimizing the protein:scaffold:lipid ratio, nanodiscs formed from the standard net neutral head group lipid,

DMPC, only allowed for the detection of monomeric TSPO (Figure V-2).[225] This is an interesting observation in the context of previous studies that observed potential destabilization of TSPO dimeric complexes upon POPC binding.[168] Additionally, nMS of nanodiscs formed with the net negatively charged lipid DPPS showed significant amounts of dimeric complex but required excessive trap collision voltages of 200 V to dissociate the protein from the nanodisc. Nanodiscs formed with TSPO using the net neutral lipid DOPE were both amenable to nMS analysis and exhibited dimeric complexes, and were thus determined to be the optimal composition for further experiments.

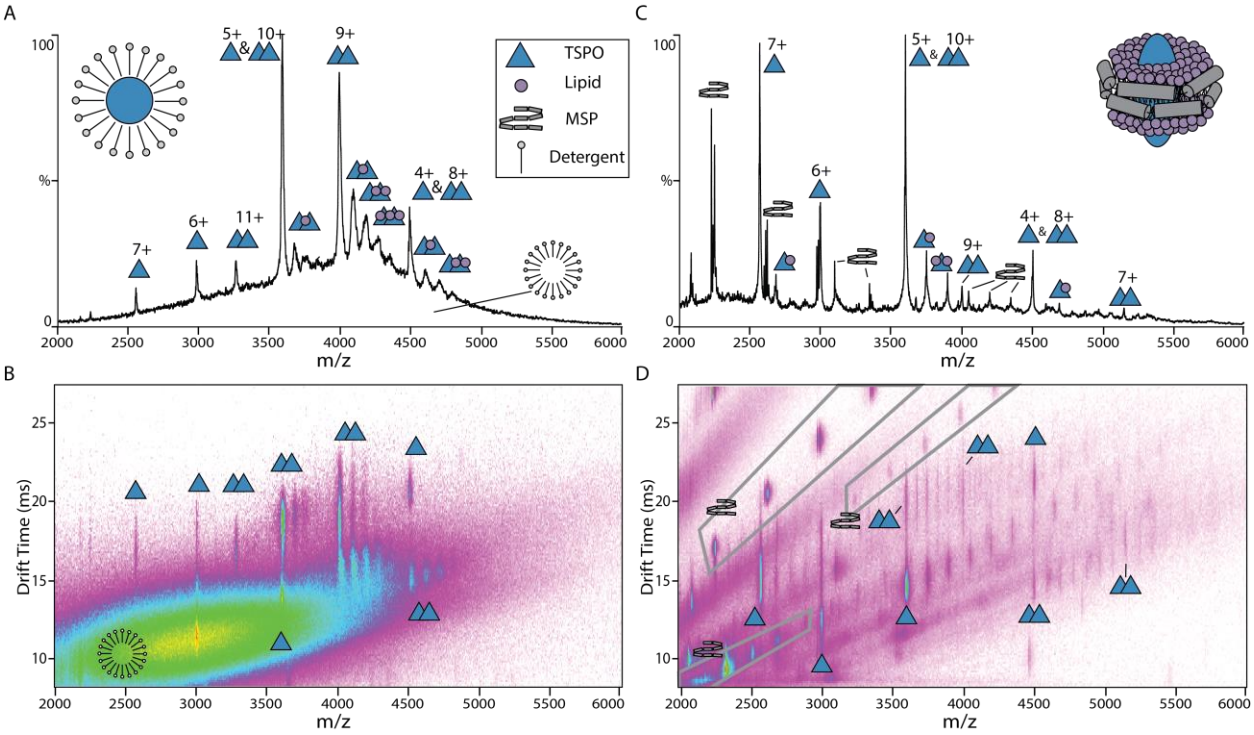

Figure 5-1. Native ion mobility-mass spectrometry of the TSPO from micelles and nanodiscs. **A, B**. Mass spectra and IM-MS data of TSPO liberated from OG detergent micelles shows mainly dimeric protein from the charge states 8-11+ and some monomeric protein at charge states 4-7+. Endogenously bound lipid can also be observed from the dimeric complexes, as well as broad noise signals related to empty detergent micelles. **C, D**. Mass spectra and IM-MS of TSPO liberated from the optimized nanodisc of POPE lipids shows relatively less dimer than the OG micelles at reduced charge states of 7-10+, and more monomeric protein at charge states of 4-7+. Lipids are bound to both monomeric and dimeric protein in high stoichiometries, indicating they are retained from the nanodisc as it dissociates. Noise related to overlapping to the scaffold protein and its retained lipid can interfere with the TSPO signals.

The nMS and IM-MS data for TSPO liberated from DOPE nanodiscs are shown in Figure 5-1C, D. Comparing these optimized nanodisc samples to the detergent micelle nMS data, key

similarities and differences can be observed. First, the arrival times in IM space are generally comparable for monomeric and dimeric species, indicating similar levels of compactness in the gas phase structures. Second, while dimeric complexes can be observed, the intensity of these complexes in the optimized nanodisc conditions is much lower. Additionally, these dimer signals for the nanodisc nMS experiments are slightly shifted to lower charge states, which can be explained considering the potential differences in solvent accessible areas of the protein between the two methods. Lastly, vastly different noise signals exist between the two methods which present unique challenges. The micelle data shows a large region affected by noise signals which creates an elevated baseline in the nMS, and is caused by the polydisperse empty micelles (Figure 5-1AB). The nanodisc data has more discrete noise signals related to the membrane scaffold protein and various lipid:scaffold complex stoichiometries, which can overlap with signals of interest in both mass to charge and arrival time space (Figure 5-1D). These observations highlight the different challenges associated with each technique as well as the importance of optimizing nanodisc incorporation protocols for protein complexes.

5.3.2 Studying the effects of solubilization technique on CIU of monotopic membrane protein CYP 3A4

While the field of membrane protein nMS has largely focused on multi-pass integral proteins, such as TSPO, recent efforts have explored how these techniques can be extended to peripheral proteins.[226] These proteins possess large aqueous domains in addition to regions which interact with but do not span, the lipid bilayer and their dual nature can complicate their biophysical characterization.[227] For this reason, single pass and monotopic membrane proteins, which are embedded in the membrane with a single alpha helix, are commonly studied in forms where the

sequence is truncated to exclude the alpha helix which can alter the structure of membrane associated domains.[227] However, to study the more biologically relevant protein form with the full sequence, a method must be capable of handling the solubilization agents needed to retain the native helix structure. Here, we use the full length monotopic protein cytochrome P450 as a model system for testing the effects of micellar and nanodisc based solubilization techniques on the gas phase stability of proteins possessing large aqueous domains.

The CYP family of proteins are enzymes important for the process of drug metabolism.[217] They bind a wide variety of drugs and have been noted in the past to also bind detergent molecules.[228] We therefore screened for nMS detergent conditions in which there was no bound detergent peak present, which led us to use the detergent OG for nMS of CYP 3A4. Figure V-3A shows monomeric CYP 3A4 liberated from OG micelles at charge states of 13-17+, and while the peaks are broad relative to soluble systems of comparable size,[229] no distinct detergent binding is observed. Mass analysis of centroid of these peaks corresponds to the mass of CYP 3A4 plus its heme cofactor, indicating we mainly see the holo protein state. CYP 3A4 has been previously studied in nanodiscs, and multiple incorporation protocols have been published.[230,231] In Figure V-3B, we show CYP 3A4 incorporated into POPC nanodiscs with the MSP3ED1 scaffold protein which was produced using a microfluidic device.[219] Peaks for CYP 3A4 liberated from the nanodisc are significantly more resolved than those from micellar conditions, and allow for the identification of apo and holo CYP 3A4, as well as retained lipid bound states. These monomeric CYP 3A4 peaks occur at charge states identical to the micellar data at 13-17+ and both IM-MS data sets show compact signals indicating native-like folding. CCS measurements from the 16+ holo peak do show small differences in orientationally averaged size, with micellar CYP 3A4 at 5022 ± 26 $Å^2$ and nanodisc CYP 3A4 at 4785 ± 25 $Å^2$. We propose that these differences, as well

as the broadness of the micellar nMS peaks, are related to increased glycerol, salt, or detergent adduction in micellar CYP 3A4 samples.

To further investigate the differences between CYP 3A4 solubilized in micelles and nanodiscs, we next performed CIU experiments on the 16+ holo charge state as shown in Figure 5-2. CIU fingerprints were generated in triplicate for nanodisc CYP 3A4 and, while multiple replicates of CYP in OG micelles were collected, data quality was too poor to analyze in all but one. Figure 5-2A shows the five features and four transitions observed for the CYP 3A4 micellar replicates, with features beginning at 17.0 ms and

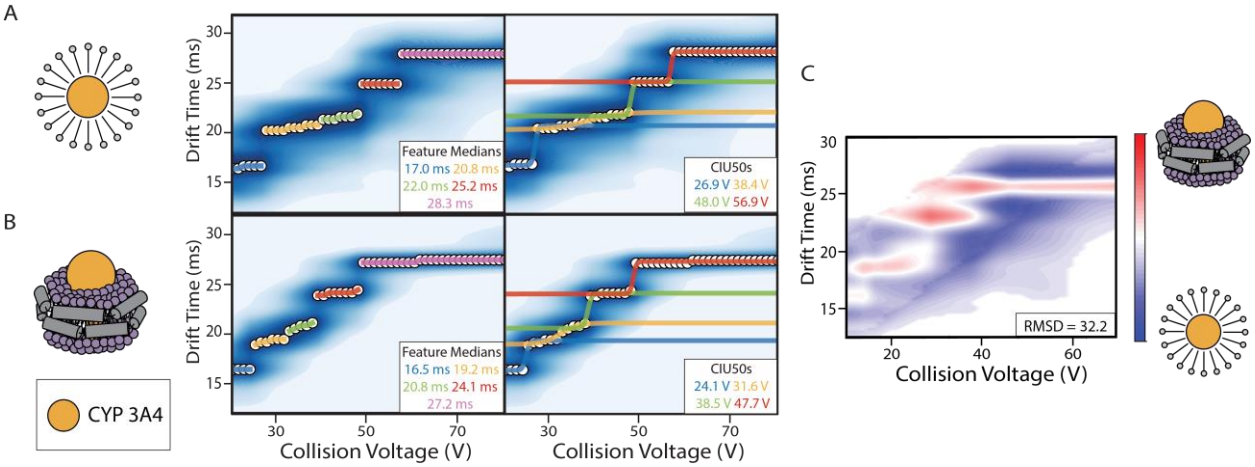

Figure 5-2. CIU of CYP 3A4 liberated from detergent micelles and nanodiscs. All fingerprints shown are for the 16+ holo state. **A**. Feature detection analysis of CYP liberated from OG micelles, N = 1, shows features of 17.0, 20.8, 22.0, 25.2, and 28.3 ms. CIU 50 transitions fit to these features occur at 26.9, 38.4, 48.0, and 56.9 V. **B**. Feature detection analysis of CYP liberated from POPC nanodiscs with the MSPE3D1 scaffold, N = 3, show features of 16.5 ± 0.3, 19.2 ± 0.2, 20.8 ± 0.2, 24.1 ± 0.5, and 27.2 ± 0.3 ms. CIU 50 transitions fit to these features occur at 24.1 ± 1.5, 31.6 ± 1.5, 38.5 ± 0.7, and 47.7 ± 0.7 V. **C**. RMSD plot of the micellar replicate subtracted from the averaged nanodisc fingerprint for CYP 3A4 shows an RMSD of 32.2%.

a first transition from the most native-like state occurring at 26.9V. Figure 5-2B shows the average fingerprint for CYP 3A4 generated from the three POPC nanodisc replicates, which also fit to five features and four transitions. The first CYP 3A4 nanodisc features occurs at 16.5 ± 0.3 ms and the transition from this first, most native-like feature to the first unfolded feature occurs at 24.1 ± 1.5V. These data show that the protein undergoes a similar unfolding trajectory regardless of it being solubilized in OG micelles or POPC nanodiscs, however, the starting drift time for the CIU

fingerprint is slightly longer for the micelles. This aligns with the CCS difference observed above and could be a result of either increased adduction or a more unfolded starting state. However, the CIU50 transitions for the micellar CYP 3A4 are all shifted to higher CV relative to the nanodisc CYP 3A4, which indicates that the micellar CYP 3A4 is stabilized relative to nanodisc. This can also be observed in the comprehensive difference analysis plot in Figure 5-2C, where the blue regions characteristic of the micellar CYP fingerprint are concentrated to the right of the nanodisc fingerprints (red). As the lipid environment of CYP 3A4 is known to be important for its function,[211] we would expect that the protein released from the lipid environment be more stable than that from the micelle. Instead, what we propose to observe here is the increased adduction of CYP 3A4 under the detergent conditions causes enhanced stability relative to the nanodisc. Adduction of soluble proteins has been previously shown to stabilize soluble proteins,[232] and thus it extends that the same phenomenon can occur for membrane proteins. Further work with this system will include additional detergent screening and buffer conditions to minimize adduction.

5.3.3 Comparing the unfolding trajectories of transmembrane protein complexes between SCOR bicelles and detergent micelles

Membrane proteins exist in a complex lipid environment where they can interact with a variety of different lipid classes.[2,141] While the nanodiscs we have discussed so far have been comprised of one lipid, more biologically relevant bilayer mimetics have been recently described.[188,211,214] One example is sphingomyelin and cholesterol rich (SCOR) bicelles which aim to mimic the more ordered, cholesterol dense regions of membranes.[188] As lipids have been known to influence the oligomeric complex formation of transmembrane proteins,[40] we chose to study a dimeric protein complex of peripheral myelin protein 22 (PMP22) liberated from SCOR bicelles

and compare the stability of these complexes to those liberated from detergent micelles. We report the first CIU measurements of protein liberated from SCOR bicelles.

Drawing from previous findings in Chapter 3 related to the relative amount of dimer complex present in nMS of PMP22 SCOR bicelle samples, we chose to evaluate the L16P mutant of PMP22. We also adopted the detergent screen and bicelle preparation protocols from Chapter 3 but extended the analysis to include performing CIU experiments with the L16P PMP22 liberated from SCOR bicelles. Figure V-4 shows example nMS and IM-MS data sets for L16P PMP22 liberated from C12E8 micelles and SCOR bicelles. In both samples, signals corresponding to monomeric protein, dimeric protein, or both are present at identical charge states. For PMP22 L16P liberated from the SCOR bicelles, considerable noise is present in the lower mass to charge range. The 13+ dimeric PMP22 charge state was chosen for CIU analysis due to its intensity and the minimal noise overlap. CCS analysis of the for the 13+ dimeric charge state of PMP22 showed the protein liberated from SCOR bicelles to be more compact, at 2653 ± 13 $Å^2$, versus 2953 ± 9 $Å^2$ from C12E8 micelles (N = 1). This CCS analysis is preliminary evidence showing that gas phase L16P PMP22 dimers adopt a different, more compact starting conformation when liberated from SCOR bicelles than those liberated from C12E8 micelles, presumably due to interactions with the lipids or different solution phase structures.

The 13+ dimeric L16P PMP22 charge state was chosen for CIU analysis due to its intensity and the minimal noise overlap. Figure 5-3 shows the average of three replicates for both micellar and bicellar L16P PMP22 with detected features and CIU50 transitions. Interestingly, the two fingerprints are quite different in terms of the number of features and feature median drift time. The micellar L16P PMP22 fingerprint shows two features, at 15.0 ± 0.3 ms and 23.5 ± 0.7 ms, and the bicellar L16P PMP22 fingerprints show three features, at 17.3 ± 0.4 ms, 21.4 ± 0.2 ms, and

25.5 ± 0.2 ms. These starting feature medians seem to contradict the CCS values above that show bicellar L16P PMP22 dimers to be more compact than micellar L16P PMP22 dimers. However, as CCS analysis calibrates for day-to-day instrumental variations in arrival time distributions, it is likely the more accurate depiction of the relative starting compactness and we therefore conclude that bicellar L16P PMP22 dimers are more compact that micellar L16P PMP22 dimers.

It is interesting to note that the difference between the first and second feature in micellar L16P PMP22 dimers (8.5 ms) is very similar to the difference between the first and last feature of bicellar L16P PMP22 dimers (8.27 ms). Additionally, the CIU50 of the micellar transition (28.9 V) is very similar to the midpoint between the two transitions of 18.3 ± 1.0 V and 38.6 ± 0.8 V (28.4 V). This could be evidence that the bicellar unfolding trajectory possesses similar starting

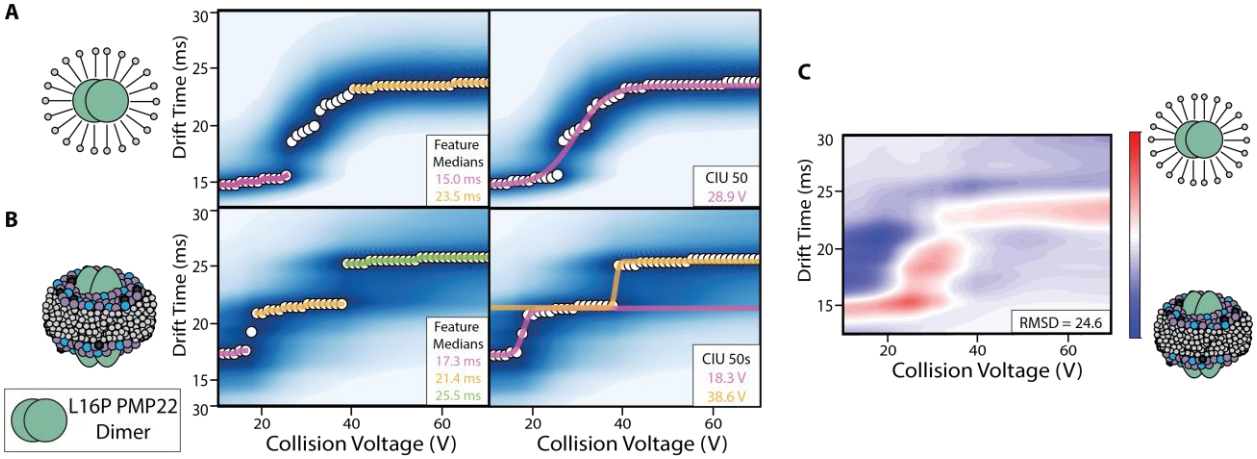

Figure 5-3. CIU of L16P PMP22 dimeric complexes liberated from detergent micelles and SCOR bicelles. All fingerprints shown are for the 13+ dimeric charge state. **A**. Feature detection analysis of L16P PMP22 dimers liberated from C12E8 micelles, N = 3, shows two features of 15.0 ± 0.3 ms and 23.5 ± 0.7 ms. CIU 50 transitions fit to these features occur at 28.9 ± 0.8 V. **B**. Feature detection analysis of L16P PMP22 dimers liberated from SCOR bicelles, N = 3, show three features of 17.3 ± 0.4 ms, 21.4 ± 0.2 ms, and 25.5 ± 0.2 ms. CIU50 transitions fit to these features occur at 18.3 ± 0.9V and 38.6 ± 0.8 V. **C**. RMSD plot of the averaged bicellar replicate subtracted from the averaged micellar fingerprint for L16P PMP22 dimers shows an RMSD of 24.6%, which is almost 2x the replicate baseline RMSDs of 13.2% for L16P PMP22 dimers liberated from micelles and 4.5% for bicelles.

and ending conformations to the micellar protein, but also experiences an intermediate phase which either does not exist or is short lived in terms of voltage space for the micellar protein. This intermediate phase would presumably occur in the voltage space between features one and two of the L16P PMP22 micellar fingerprints, in which the centroids are not flat enough to detect a

reliable feature. Regardless, the unfolding trajectories and starting gas phase structures are different enough that a direct stability comparison using the CIU50 data is not appropriate. These differences are highlighted in Figure 5-3C which shows the comprehensive difference analysis between micellar and bicellar L16P PMP22 dimer CIU fingerprints. As these homodimeric complexes were not previously reported before Chapter 3 of this thesis, further work in this system should focus on computational methods for generating dimeric structures to compare the CCS values found here.

5.3.4 Comparison of CIU from PMP22 containing soluble construct liberated from micelles, bicelles, and nanodisc

Noting the marked differences between PMP22 dimers liberated from micelles and bicelles, we next aimed to compare PMP22 protein liberated from nanodiscs to our previous data sets. However, as shown in Figure 5-1 with TSPO monomers, it can be difficult to distinguish signals between the membrane scaffold protein, 22 kDa, and proteins of similar size such as the 19 kDa PMP22. Additionally, we wanted to first begin with a monomeric system to compare across all three solubilization techniques to avoid the challenges we found in generating sufficient complex signal for CIU, as was the case with TSPO. For these reasons we chose to study a version of WT PMP22 which has an added 11 kDa of soluble tags, denoted throughout as WT_{tag} PMP22. The WT_{tag} PMP22 also proved more amenable to nanodisc sample preparation for nMS, as the penta-histidine tag included in the sequence allowed us to use a nickel affinity resin to separate incorporated nanodiscs from empty nanodiscs. Note that for this reason, the histag free construct of MSP1D1(-) was utilized, allowing specificity in Ni affinity resin step.

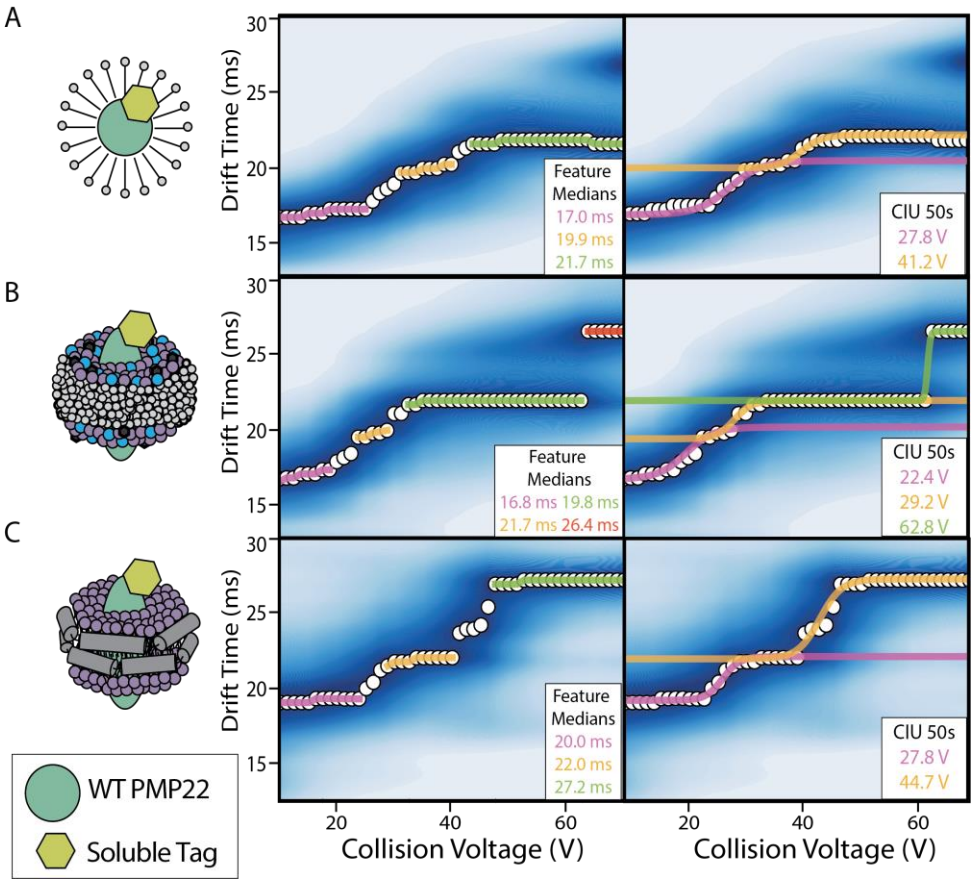

Figure 5-10. CIU of WT_{tag} PMP22 monomers liberated from detergent micelles, SCOR bicelles, and POPC nanodiscs. All fingerprints shown are for the 9+ monomeric charge state. **A**. Feature detection analysis of WT_{tag} PMP22 monomers liberated from C12E8 micelles, N = 3, shows three features of 17.0 ± 0.2 ms, 19.9 ± 0.1 ms, and 21.7 ± 0.2 ms. CIU 50 transitions fit to these features occur at 27.8 ± 1.5 V and 41.2 ± 0.5 V **B**. Feature detection analysis of WT_{tag} PMP22 monomers liberated from SCOR bicelles, N = 3, shows four features of 16.8 ± 0.2 ms, 19.8 ± 0.3 ms, 21.7 ± 0.2 ms, and 26.4 ± 0.2 ms. CIU 50 transitions fit to these features occur at 22.4 ± 1.3 V, 29.2 ± 2.0 V, and 62.8 ± 0.7 V. **C**. Feature detection analysis of WT_{tag} PMP22 monomers liberated from POPC nanodiscs, N = 3, shows three features of 20.0 ± 0.5 ms, 22.0 ± 0.3 ms, and 27.2 ± 0.2 ms. CIU 50 transitions fit to these features occur at 27.8 ± 1.1 V and 44.7 ± 3.3 V

As shown in Figure V-5, WT_{tag} PMP22 can be successfully liberated and detected from C12E8 micelles, SCOR bicelles, and POPC MSP1D1(-) nanodiscs. Monomeric WT_{tag} PMP22 is observed at charge states 7-15+ when solubilized in both micelles and bicelles, but the WT_{tag} PMP22 nanodisc IM-MS shows significant overlap with MSP1D1(-) signals in the higher m/z range, reducing the observable charge states to 7-12+ (Figure V-25C). At lower trap collision voltages, signals related to lipid bound MSP1D1(-) can also interfere with the WT_{tag} PMP22 signals. The 9+ charge state of WT_{tag} PMP22 was found to be most suitable for CIU analysis due to its distance in mass to charge space from the MSP1D1(-) signals. CCS analysis of the 9+ charge

state for micelles and bicelles was found to yield similar values, at 2683 ± 12 $Å^2$ and 2659 ± 12 $Å^2$ respectively. Contrastingly, the CCS for the 9+ charge state liberated from nanodiscs was found to be larger at 2793 ± 13 $Å^2$. Given that the SCOR bicelles possess more biologically relevant lipids than the POPC nanodiscs, one could hypothesize that the protein should be more compact or folded when ejected from the bicelle. While we do observe the bicellar WT_{tag} PMP22 monomer to be most compact in terms of CCS, there is no available high-resolution structure for the WT_{tag} PMP22 construct to which we can compare our CCS values in order to assess their proximity to native folding. It is also interesting to note that while the L16P PMP22 data shows compaction of dimers liberated from the SCOR bicelles, the same trend is not observed here for WT_{tag} PMP22 monomers.

CIU fingerprints for the 9+ of WT_{tag} PMP22 solubilized in the micelles, bicelles, and nanodiscs were generated in triplicate and their averages with detected features and CIU50 transitions are shown in Figure 5-4. The WT_{tag} PMP22 fingerprints from micelles and SCOR bicelles possess three similar features at approximately 17.0 ms, 19.9 ms, and 21.7 ms (Figure 5-4 A, B). However, the most unfolded feature which occurs at 26.4 ms in the bicelle fingerprints is not detected in the micelle fingerprint, as it does not reach the necessary relative intensity within the bounds of the CV range. This information can be combined with the CIU50 values for the first and second transitions of the datasets to show that the WT_{tag} PMP22 liberated from detergent micelles is more stabilized relative to SCOR bicelles. Interestingly, there are no apparent adduction differences in these data as there were for the CYP 3A4 data in Figure 5-2. To further explore the solubilization agent dependent changes in the CIU fingerprints of monomeric WT_{tag} PMP22, comprehensive differences analysis was performed between all three datasets as shown in Figure V-6. The micellar and bicellar fingerprints are found to be the most similar between the three

solubilization techniques, with an RMSD of 16.0%. This RMSD is still 3x the highest baseline replicate RMSD of 5.8%.

The features detected in the WT_{tag} PMP22 nanodisc fingerprints are much different than those detected in the micellar and bicellar fingerprints, with features at 20.0 ± 0.5 ms, 22.0 ± 0.3 ms, and 27.2 ± 0.2 ms (Figure 5-3C). The differences in starting CCS and starting features indicate that the WT_{tag} PMP22 monomers liberated from POPC nanodiscs are beginning at a more unfolded state relative to the micelle and SCOR bicelles, and it is possible that the three features detected to for the nanodisc fingerprint correspond to the second, third, and fourth feature of the micellar and bicellar fingerprints. Comprehensive difference analysis also shows that the nanodisc fingerprints are significantly different from the micellar and bicellar fingerprints, with RMSDs of 25.4% and 34.8% (Figure V-6), which are at least 2x greater than the nanodisc baseline replicate RMSD of 16.0%. While there are clear differences between the unfolding trajectories between all three solubilization techniques, more experiments would help elucidate the cause of these differences. For example, the addition of WT_{tag} PMP22 liberated from POPC bicelles or nanodiscs with more complex lipid mixtures could allow for the independent analysis of the lipid composition effect from the solubilization technique and vice versa. Further, the partly soluble protein construct, CYP 3A4, presented here would be an interesting comparison point for another full data set with micelles, bicelles, and nanodiscs of a mostly hydrophobic transmembrane protein, such as TSPO or PMP22 without soluble tags.

5.4 Conclusions

NMS methods have been extended to membrane proteins solubilized using many solubilization techniques, and here we have demonstrated that IM-MS and CIU methods can also

be extended from detergent micelles to more complex membrane mimetics. We show that IM-MS can inform nanodisc optimization protocols pertaining to protein complexes using TSPO as a model system. We report CIU fingerprints from SCOR bicelles and nanodiscs for multiple proteins which are readily analyzed by existing software workflows. In the case of CYP 3A4, we observe excess detergent, salt, and glycerol adduction cause increased stability of micellar samples relative to nanodisc samples. We show that these methods can be used to study protein complexes with L16P PMP22 and find preliminary evidence that SCOR bicelles may support more compact gas phase dimers than C12E8 micelles. Additionally, we find that for WT_{tag} PMP22, CIU fingerprints from three different solubilization techniques exhibit differences which are not explained by the presence of lipids alone. However, from our data we cannot determine whether the differences are a function of the lipid composition, SCOR mixed lipids versus POPC, or the solubilization methods used.

Overall, we find that CIU is sensitive to the solubilization technique a protein is housed in, as all CIU fingerprints gathered here for the same protein, PMP22, liberated from different solubilization techniques were different to some extent. To further this area of research, studies to parse out of the contribution of lipid composition to the unfolding trajectories will be key, as our current data cannot rule out the possibility that a protein liberated from a POPC bicelles and POPC nanodisc do not present the same CIU fingerprint. Additionally, extension to more diverse protein systems, such as GPCRs, could be vital for defining the role of membrane protein nMS in the assessment of solubilization agents for applications such as cryo-EM or pharmaceutical screens.

5.5 Acknowledgments

Membrane protein research in the Ruotolo lab is supported by the National Institute of Health under Grants GM105942. The author thanks collaborators Kristine F. Parson and Iliana L. Hampton for their contributions to this chapter.

Chapter 6 Conclusions and Future Directions

6.1 Conclusions

Due to their important roles in cells, many disease states can be linked to the dysfunction of membrane proteins and the majority of pharmaceuticals approved by the FDA target membrane proteins. Dynamic structural measurements of membrane proteins can help elucidate the role of these proteins in disease and how pharmaceutical interventions can be designed to treat them. However, the challenges associated with their hydrophobic nature and complex native environments has severely limited the biophysical characterization of membrane proteins. This thesis adds important workflows and analysis methods to the growing field of membrane protein mass spectrometry aimed at overcoming the challenges associated with membrane protein characterization. Specifically, the comparative stability analyses and classification workflows presented in this thesis demonstrate novel ways to study the effects of mutation, disease association, and ligand binding on gas phase membrane protein structure, which could be broadly applied to membrane proteins beyond those described in this work.

One hinderance to the broader adoption of CIU techniques for membrane proteins is the large amount of noise which can be generated by solubilization agents. In Chapter 2, we demonstrate a workflow for removing this noise from CIU fingerprints using Gaussian fitting and automated feature detection, and we use this workflow to screen lipid and protoporphyrin binding to a protein of therapeutic interest, TSPO. We further extend our method to characterize the CIU of a previously unidentified endogenous lipid bound to TSPO. Through the use of CIU classification schemes, stability analysis, and high-resolution MS (HRMS), we are able to identify

the endogenous lipid as a set of lipids containing the PG head group with various chain lengths and saturations. As the observance of endogenously bound lipids is common in membrane protein MS, this workflow for identifying endogenous lipids through CIU and HRMS contributes to the larger field of membrane protein MS. Our identification of the endogenous lipids as PG with varying chain length is supported by recently published tandem HRMS workflows, in which lipids ejected from native TSPO dimers were fragmented to yield detailed head group and chain length distributions.[194] Further, the denoising workflow presented in this chapter can be utilized to manage noise related issues in CIU datasets from any protein origin.

There are multiple membrane proteins for which sequence mutations can cause pathogenic misfolding and the stability of these mutant protein variants has been implicated in disease etiology. While much progress has been made using CIU to characterize membrane protein ligand binding, no studies have previously attempted to detect CIU differences in membrane proteins based on their sequence associated disease pathology. In Chapter 3 and 4, we demonstrate the ability of IM-MS and CIU to detect differences in sequence variants of two disease associated membrane proteins, PMP22 and KCNQ1 VSD. We show that sequence variants of PMP22 can exhibit altered stability and dimeric complex formation, and that differences in dimeric complex formation are intensified when the protein is liberated from SCOR bicelles, which mimic the proteins native lipid environment. Our data adds key pieces of information to previously published cellular biology assays and allows us to construct a mechanism of PMP22 disregulation leading to neuropathic disease. For KCNQ1 VSD, we build a classification scheme capable of differentiating variants based on disease association and set up a framework for rapidly evaluating the hundreds of disease-associated protein variants. We also find preliminary evidence of destabilization of the KCNQ1 VSD variant R231C, which has been shown to exhibit unique cellular trafficking and

expression properties. Together, these chapters add to the knowledgebase of each respective protein and lay the groundwork for extension to high-throughput screens of protein variants in misfolding diseases.

Finally, as all CIU of membrane proteins reported to date have used detergent micelles as solubilization agents, we endeavored to assess the utility and compatibility of CIU with proteins liberated from more complex membrane mimetics. We show that different membrane mimetics present different noise related challenges, and that IM-MS can be used to inform nanodisc incorporation protocol optimization for protein complexes. We also present the first CIU of membrane proteins liberated from bicelles and nanodiscs in multiple protein systems. We find that CIU is sensitive to the solubilization agent used for each membrane protein and describe the importance of this data in the context of membrane protein solubilization agent screening.

Overall, this thesis presents multiple workflows related to denoising, membrane protein classification, and stability analysis which can be adopted by the growing field of membrane protein MS. Throughout our experiments, we have also added valuable information to the fields of membrane protein misfolding diseases, endogenous lipid binding, and mimetic incorporation protocols.

6.2 Future Directions

6.2.1 High-throughput membrane proteins screens

At odds with the intrinsic challenging nature of membrane proteins are the many types of information which could be valuable to their characterization. Screens of small molecule binding to membrane proteins, stability and disease association of the hundreds of known protein variants

in membrane protein misfolding diseases, and lipid screening of membrane proteins to illuminate biological interactions can all be imagined to add a dearth of information to our understanding of membrane protein interactions and roles in disease. In this thesis, we have presented workflows which demonstrate the capability of obtaining the information content that is garnered by each of those screens. The experiments detailed here have been relatively low throughput due to limitations in manual sample handling, with each experiment requiring a few hours of sample preparation before the data can be collected. The use of automated sample handling methods interfaced to the mass spectrometer would greatly increase the throughput of experiments, where one can imagine many samples are prepared in tandem by researchers and then stored in well plates for automated injection and analysis. Additionally, software advances now allow for batch processing of CIU data with the denoising processes described in this thesis, which could allow for even shorter data collection times on the scale of seconds per instrument voltage scanned. The demonstration of a high throughput membrane protein screens based on native IM-MS and CIU could enable the adoption of these technologies as routine analyses in research and industry.

6.2.2 CIU of membrane proteins from diverse solubilization methods

The utility of detergent micelle-based solubilization agents for native mass spectrometry (nMS) of membrane proteins has been thoroughly demonstrated in this thesis and work by others. However, as micelles do not form bilayers similar to the membranes in which these proteins

natively reside, their wide use and the biological conclusions drawn from proteins solubilized in micelles use has been, rightfully, called into question. However, as the field moves to embrace lipid-based membrane mimetics, we ask if an artificial bilayer formed of only one lipid is any "better" at mimicking the native folding and structure of a membrane protein? Similarly, how does the ability to preserve structure change when bilayers are formed with a complex mixture of lipids that are constrained by an artificial boundary provided by membrane scaffold proteins or detergents in nanodiscs and bicelles? These questions are difficult to address and remain largely unanswered. We see nMS and CIU as uniquely suited to explore these questions and inform on the practical use of each of these solubilization methods. In this thesis we have begun to address these questions and demonstrated that CIU is indeed sensitive to the solubilization method of membrane proteins, but we assert that a more comprehensive analysis of lipid composition and solubilization method for a diverse set of membrane proteins would greatly aid to the development of this field.

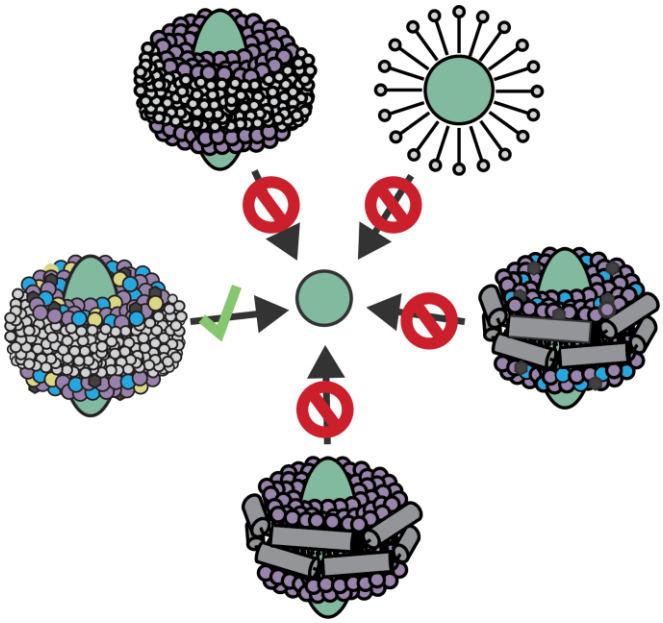

Figure 6-1. Example of optimized solubilization agents across different mimetics. Both lipid composition and mimetic type are optimized to result in the most compact gas phase structure.

Figure 6-2. Example of optimized solubilization agents across different mimetics. Both lipid composition and mimetic type are optimized to result in the most compact gas phase structure.

Appendices

I. Chapter 2 Supporting Information

	Expected	Observed
TSPO dimer	35.952	36.033
TSPO monomer	17.976	17.964

TSPO:Ligand Complexes Mass Analysis (in kDa)

	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs
	1 Bound		2 Bound		3 Bound		4 Bound		5 Bound	
TSPO: 14:0 CDL	37.296	37.316	38.558	38.557	39.821	39.812	-	-	-	-
TSPO: 18:1 PA	36.733	36.762	37.433	37.477	38.133	38.183	-	-	-	-
TSPO: 18:1 PE	36.777	36.843	37.521	37.559	-	-	-	-	-	-
TSPO: 16:0 PI	36.841	36.911	37.649	37.729	-	-	-	-	-	-
TSPO: 16:0-18:1 PS	36.794	36.819	37.555	37.587	38.316	38.349	39.077	39.131	39.838	39.888
TSPO: 16:0-18:1 PC	36.793	36.822	37.553	37.558	38.313	38.321	-	-	-	-
TSPO: 16:0 PG	36.778	36.771	37.523	37.506	38.268	38.244	39.013	39.013	-	-
TSPO: Endogenous Ligand	-	36.806	-	37.478	-	-	-	-	-	-
TSPO: PPIX	36.596	36.683	37.158	37.235	-	-	-	-	-	-

Figure I-1 . Mass Analysis of TSPO and TSPO Ligand Complexes. Expected (Exp) and observed (Obs) are listed for all complexes used in analysis (kDa) and the small differences shown are attributed to the nature of native mass spectrometry experiments. Expected TSPO:ligand complex masses were calculated by adding the ligand weight to the value listed in the observed TSPO dimer mass (36.033 kDa). Note that when ligands are added to the solution, the peak corresponding to the TSPO dimer alone tends to increase in mass slightly, indicating increased adduction. This increased mass was used to measure the weight of the endogenous ligand across replicates (740 ± 20 m/z).

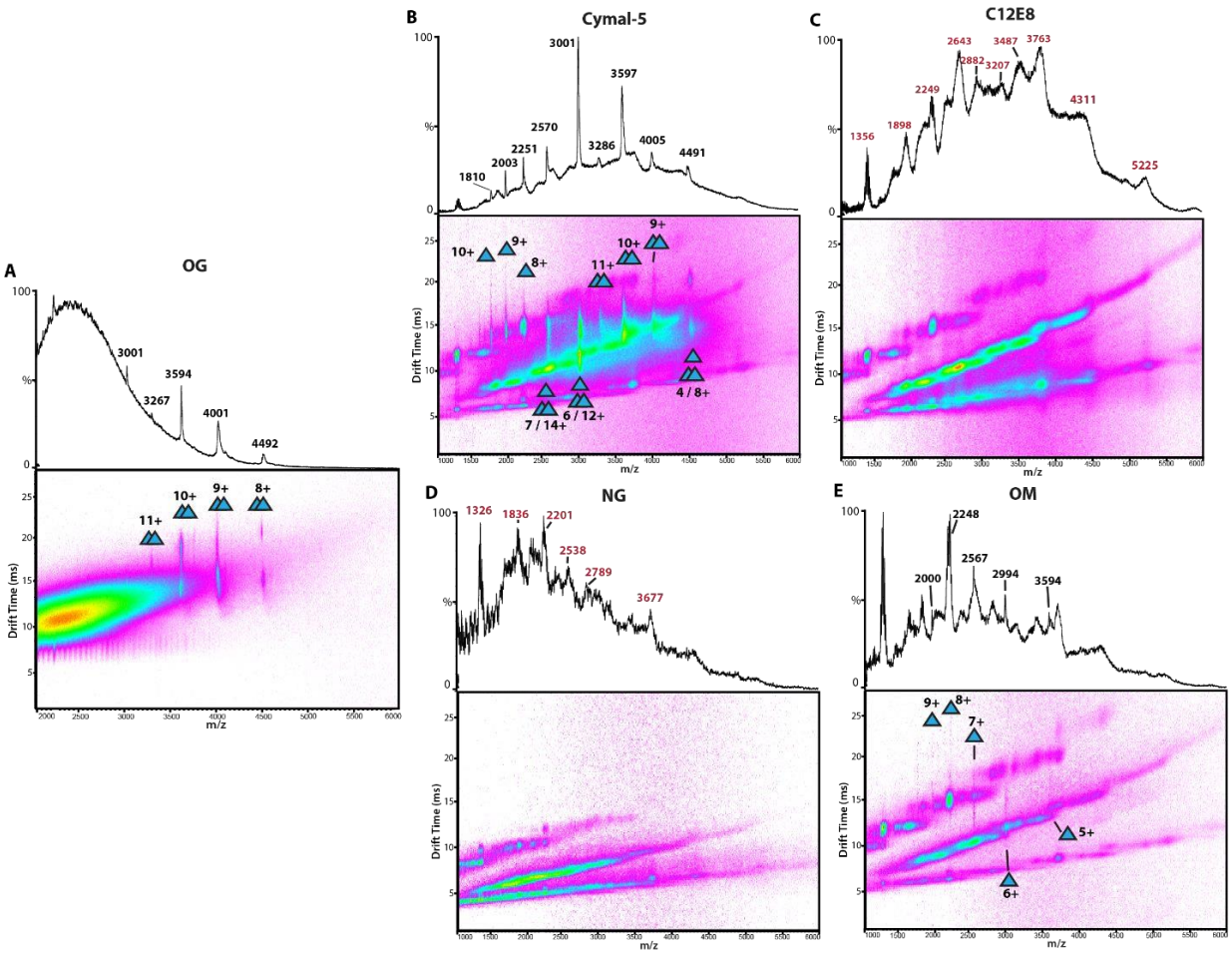

Figure I-2 Detergent screen of TSPO. All IM-MS data shown in squared scale. Peaks labeled with black lettering are consistent with masses expected for TSPO monomer or dimer. Peaks labelled with red lettering are not consistent with TSPO complexes and are most likely detergent signals. A. IM-MS of TSPO solubilized in OG (Octyl β-D-glucopyranoside) shows dimer (blue triangles) at multiple charge states, including the 9+. B-E. Other detergents screened include Cymal-5 (5-Cyclohexyl-1-Pentyl-β-D-Maltoside), C12E8 (Dodecyl Octaethylene Glycol Ether), NG (n-Nonyl-β-D-Glucopyranoside), and OM (n-Octyl-β-D-Maltopyranoside). C12E8 and NG datasets showed no signals corresponding to TSPO monomer or dimer, while OM showed only low-intensity monomer signals that appear to be non-native. Cymal-5 IM-MS does show TSPO dimer, but heavy adduction is present even at high energies. Additionally, the Cymal-5 data show highly charged monomer signals, indicative of non-native folding. Due to these factors, OG was chosen as the ideal detergent over Cymal-5.

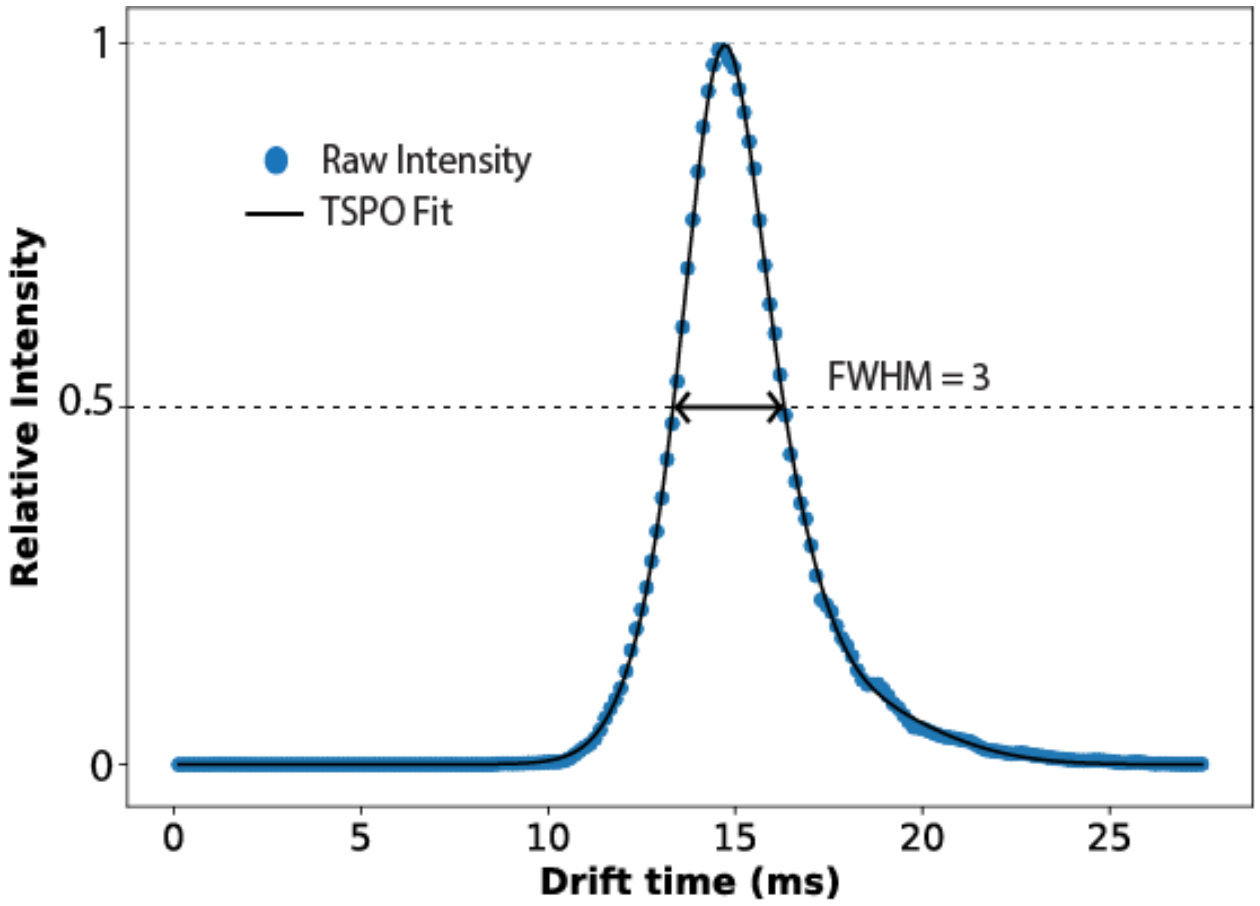

Figure I-3 Width analysis for Gaussian denoising. The arrival time for the apo 9+ TSPO dimer can be extracted without interfering micellular noise at the lowest collision voltage energies. From this data, a Gaussian distribution was fitted and the full width at half maximum can be measured. These measurements (3ms) and a manually optimized tolerance of (0.4 ms) were used as parameters in Gaussian fitting of full CIU fingerprints using CIUSuite 2.1, and in combination with relative peak intensity was used to differentiate noise signals from TSPO signals, where noise signals were over 3ms in width and lower intensity.

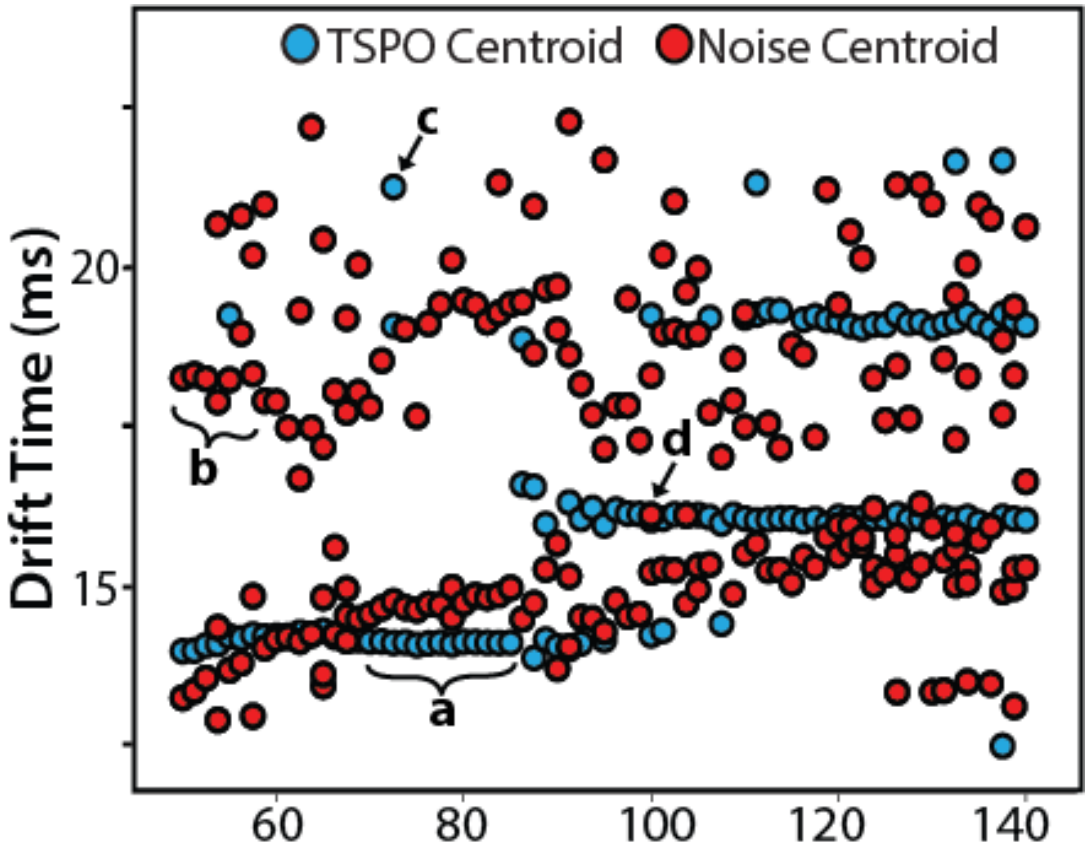

Figure I-4 Feature detection of Gaussian fitted CIU data. Gaussian centroids are originally assigned as corresponding to either protein (blue) or noise- related signals (red). These assignments are all automated and are a function of the input parameters from Figure I-3. Feature detection is performed on these Gaussian peaks to obtain Figure 2-2E and this reduced data set is used for all CIU50 and classification analysis. The feature detection parameters of feature length and tolerance were manually optimized for each ligand set to obtain fits which most accurately reflected the fingerprint, and automated re-assignments as noise or protein were made after feature detection. Highlighted by letters a-d in this figure are the four possible outcomes for centroids after automated feature detection is performed. **a**. These Gaussian centroids assigned as TSPO signals combine to form a feature, defined by the following input parameters: feature length (10 steps), allowed width (1 ms), and minimum collision voltage step gap (2 steps). **b**. These centroids assigned as noise signals are not incorporated in any features and are excluded from the final dataset. **c**. This centroid was assigned as a TSPO related signal during Gaussian fitting but is excluded from the final de-noised dataset as it is not part of a feature. **d**. This centroid originally assigned as a noise signal is re-assigned as a protein related peak due to its good agreement with surrounding TSPO peaks which form a feature.

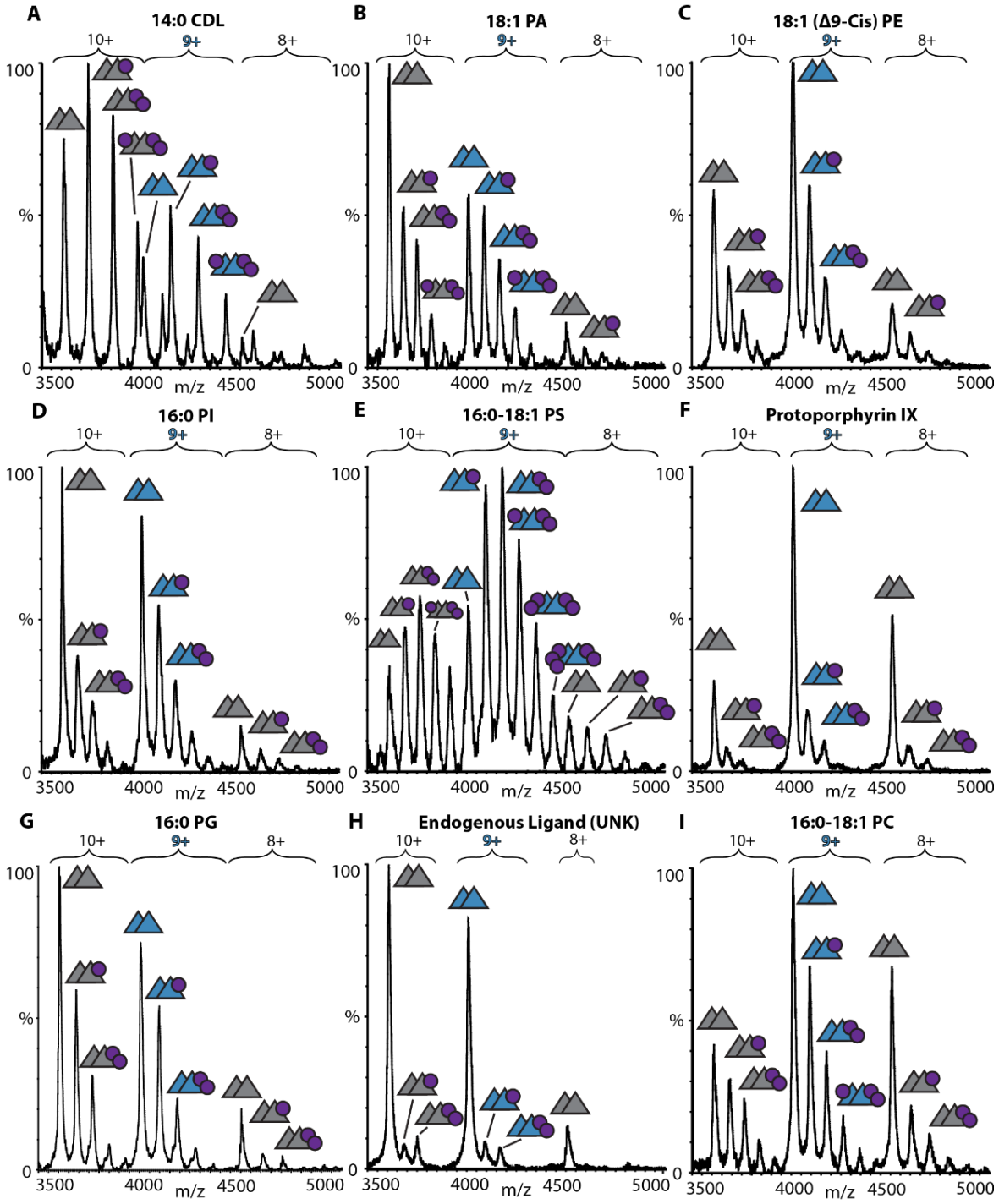

Figure I-5 Mass spectra of ligand bound TSPO. After incubation with the selected ligand (except in the case of the endogenous ligand), dimer TSPO can be observed in complex with 2-5 ligands at charge states 8-10+. The 9+ charge state mainly used in our analysis is labeled in blue. For clarity of peak assignment, these spectra were

taken at high voltages and baseline corrected in MassLynx. A. 1',3'-bis[1,2-dimyristoyl-sn-glycero-3-phospho]-glycerol [14:0 CDL]. B. 1,2-dioleoyl-sn-glycero-3-phosphate[18:1 PA]. C. 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine [18:1 (Δ9-Cis) PE]. D. 1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-myo-inositol)[16:0 PI]. E. 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine [16:0-18:1 PS]. F. Protoporphyrin IX (PPIX). G. 1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-rac-glycerol) [16:0 PG]. H. A previously unknown endogenous ligand, proposed in this manuscript to be a group of lipids possessing the PG head group. I. 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine [16:0-18:1 PC]

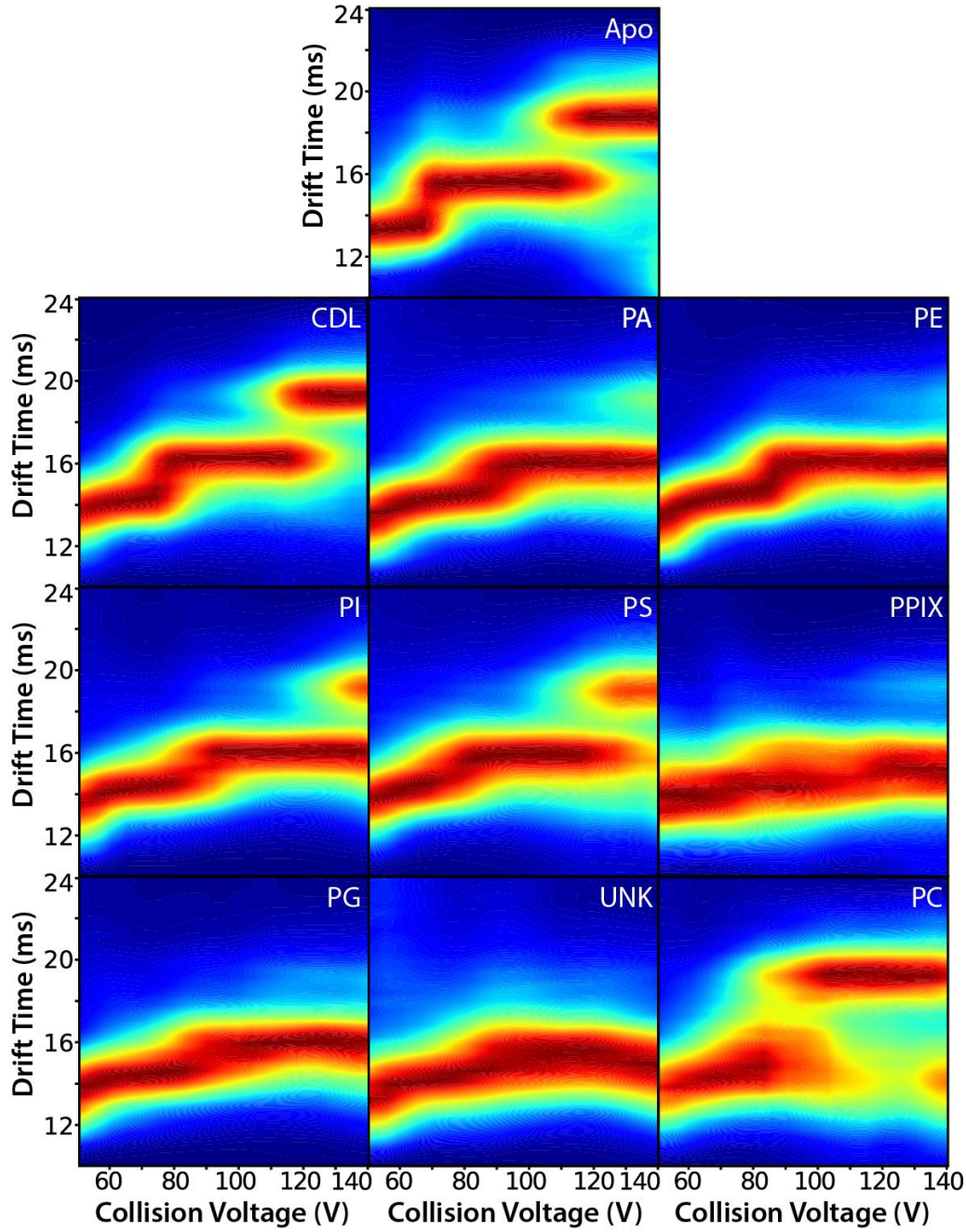

Figure I-6 CIU fingerprints of TSPO:ligand complexes. Averaged fingerprints, n=3, of the one ligand bound state for each ligand. Fingerprints were collected in 5 V increments from 50 -140 V, with each increment comprising scans averaged over 2 minutes. Above 140 V, signal quality was severely degraded and no additional unfolding was able to be observed. Ligand binding causes some differences that are observable by eye, such as stabilization of the first and second features by ligands PPIX, PG, UNK, as well as the unique unfolding pathway of the lipid PC. In all PC replicates, it seems that the most unfolded state occurs at lower voltages than Apo TSPO, indicating destabilization.

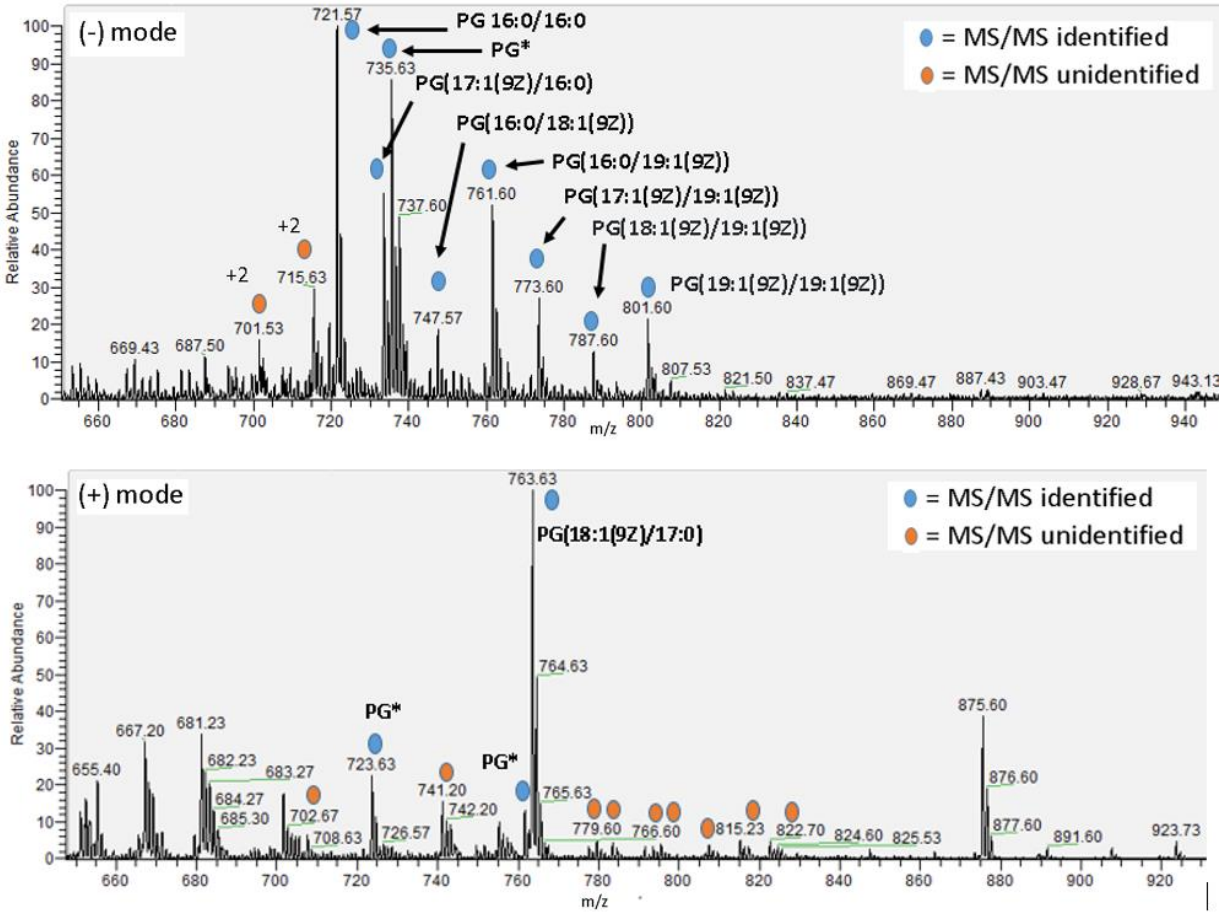

Figure I-7 HRMS of lipid extraction from TSPO with endogenous lipid present. Each labeled peak from positive and negative mode electrospray was isolated and fragmented at varying HCD energies. Intact mass and fragment masses were considered when using LipidMaps to attempt to identify peaks. Lipid matches were identified for peaks labeled with blue, and exact lipid chain information was obtained for all except those marked with (*). Orange labeled peaks yielded no lipid identification information.

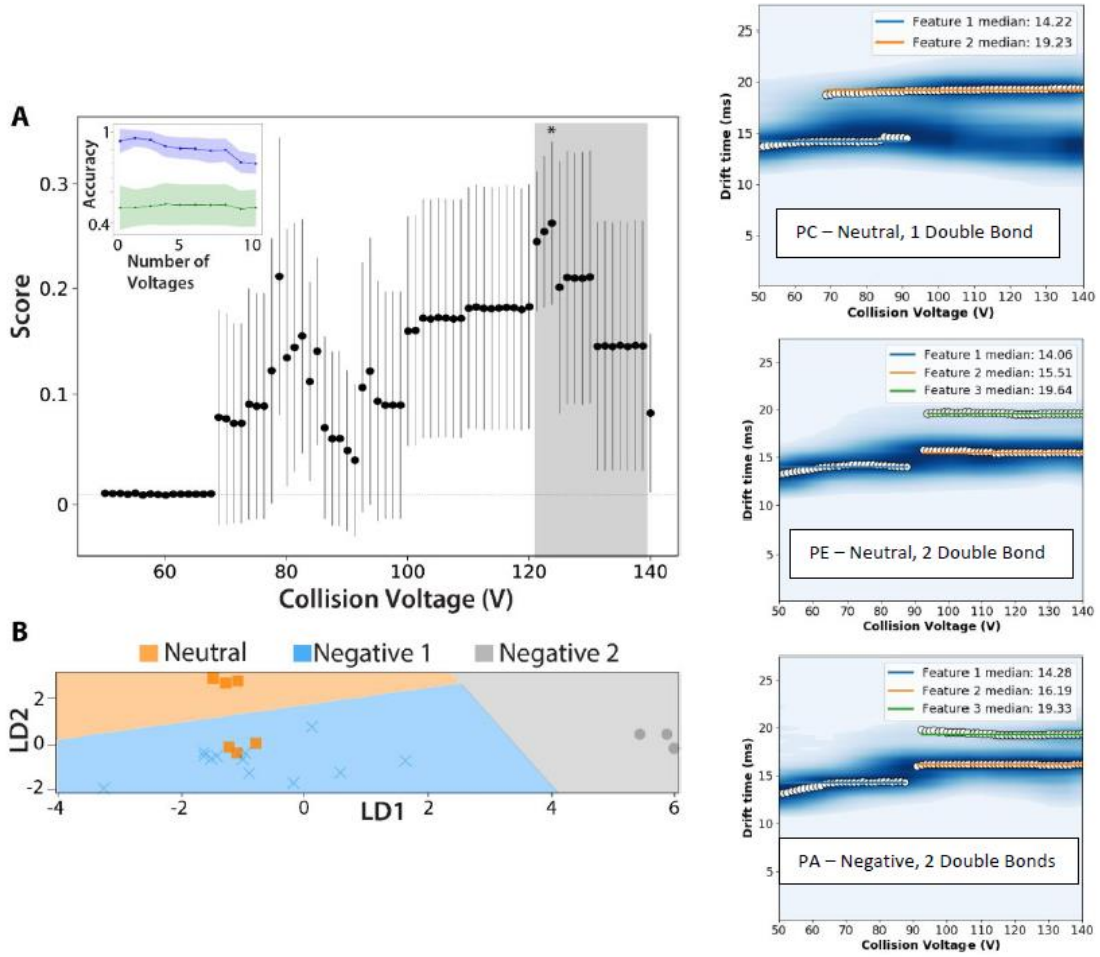

Figure I-8 Classification of TSPO complexes by lipid head group charge. **A**. Each voltage was scored for the ability to differentiate three classes: neutral lipids (PC, PE), lipids with a net charge of -1 (PA, PG, PI, PS) and lipids with a net charge of -2 (CDL), N = 3 for each ligand. Voltages in the gray box were selected for cross-validation, inset plot, and one voltage (*) was chosen to build the scheme based on accuracy. **B**. Three-way classification depicted as decision regions on linear discriminant axes shows that some neutral replicates were classified as negatively charged, indicating that the scheme built is not capable of reliably classifying TSPO:lipid complexes based on head group charge. Attempts to manually identify voltage combinations that may yield better separation gave similar or worse results, where neutral lipids were often misclassified as negative (data not shown). **C**. Three replicates that often classified incorrectly corresponded to the neutral lipid PE, where the classification scheme grouped PE with the -1 charge lipids (PA, PI, PG, PS). Looking at the features detected for PC, PE, and PA, it is clear that PE and PA fingerprints possess very similar features. Although the typical physiological charge of DOPE and DOPA is different, their shared attribute of chain unsaturation (2 double bonds) can be used effectively to classify these fingerprints. (Figure 2-5).

II. Chapter 3 Supporting Information

PMP22 released from C12E8 Micelles				
	Expected	Observed	Expected	Observed
	Monomer		Dimer	
PMP22 WT	19276	19407	38552	39139
PMP22 S22F	19336	19340	38672	38734
PMP22 A67T	19306	19310	38612	38685
PMP22 T118M	19306	19306	38612	38689
PMP22 G93R	19375	19378	38750	38815
PMP22 L16P	19260	19265	38520	38800
PMP22 H12Q	19267	19268	38534	38635
PMP22 G150D	19334	19345	38668	38698

PMP22 released from SCOR Bicelles				
	Expected	Observed	Expected	Observed
	Monomer		Dimer	
PMP22 WT	19276	19292	38552	38575
PMP22 L16P	19260	19480	38520	38998
PMP22 G150D	19334	19399	38668	38797

Figure II-1 Mass Analysis of PMP22 Variant Monomers and Dimers. Expected (Exp) and observed (Obs) are listed for all complexes used in analysis (Da). Note that due to the presence of two extra amino acid residues related to variations in thrombin cleavage, the signals corresponding to the lowest mass to charge at each charge state were chosen for mass analysis. More information on the additions can be found in Figure II-4.

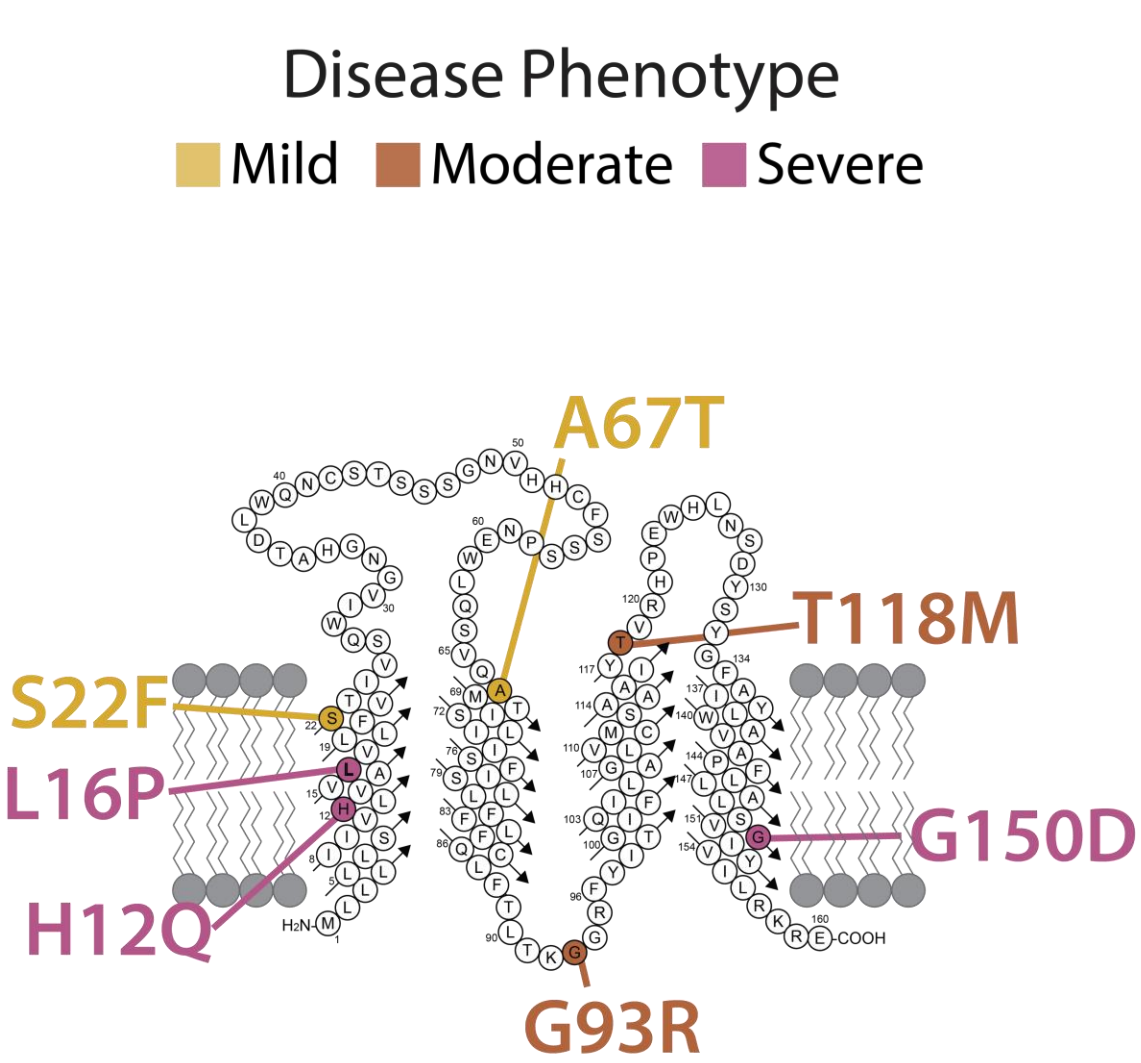

Figure II-2 Sequence location of PMP22 mutations studied in this work. In addition to WT PMP22, two mutations associated with the mild neuropathy HNPP (A67T and S22F), two mutations associated with the moderate neuropathy CMT1 (T118M and G93R), and three mutations associated with the severe neuropathy DSS (L16P, H12Q, and G150D) were characterized in this work.

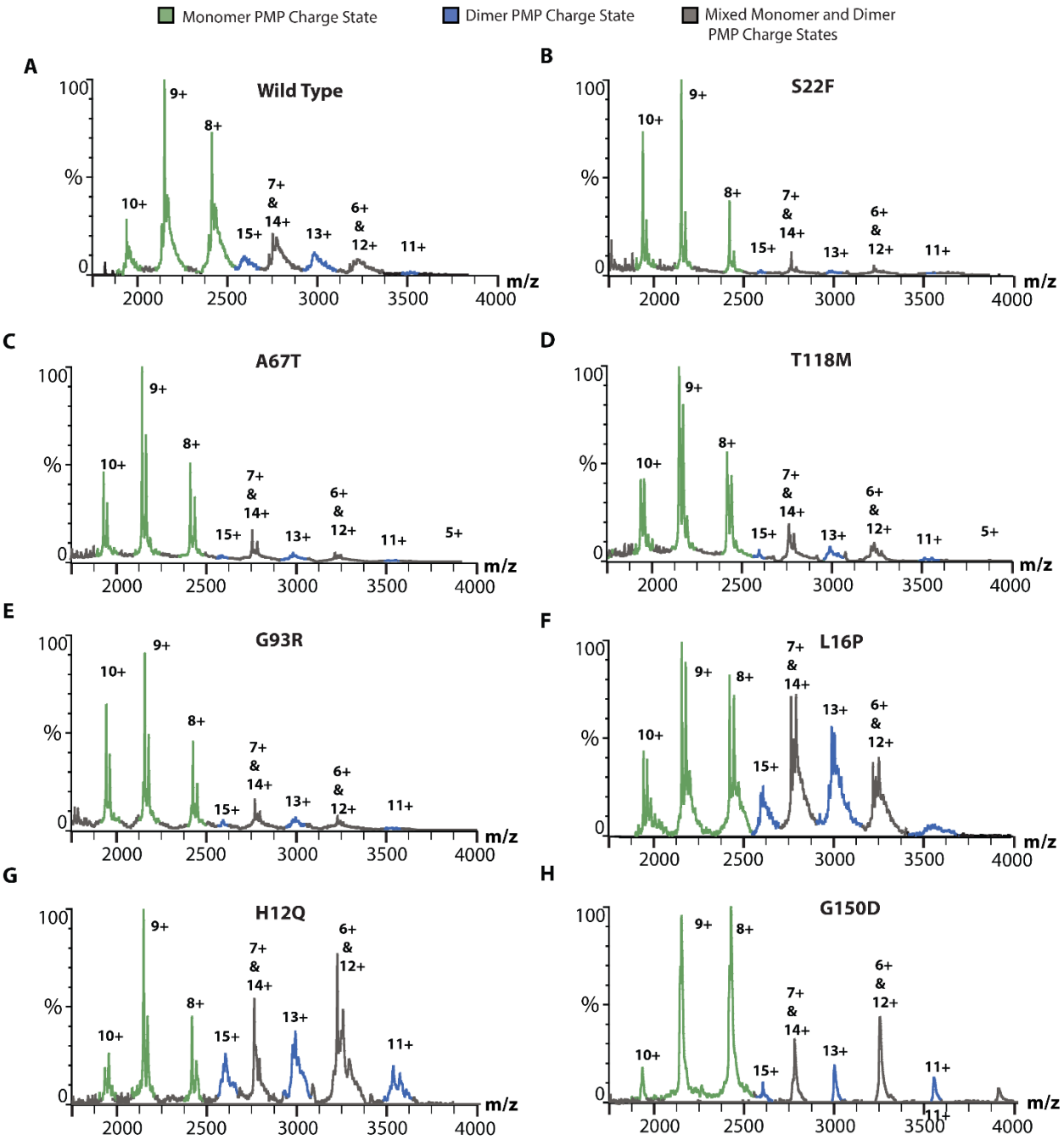

Figure II-3 Mass spectra of PMP22 variants released from C12E8 detergent micelles. For all variants released from detergent micelles, charge states associated with monomer (green) and dimer (blue) were observed. Based on the charge states envelope, some signals were determined to contain indistinguishable amounts of both monomer and dimer (gray). These signals were excluded from percent dimer calculations. Spectra shown here are background subtracted for visualization purposes, but data de-noising for further analysis is only performed as described in Figure II-8.

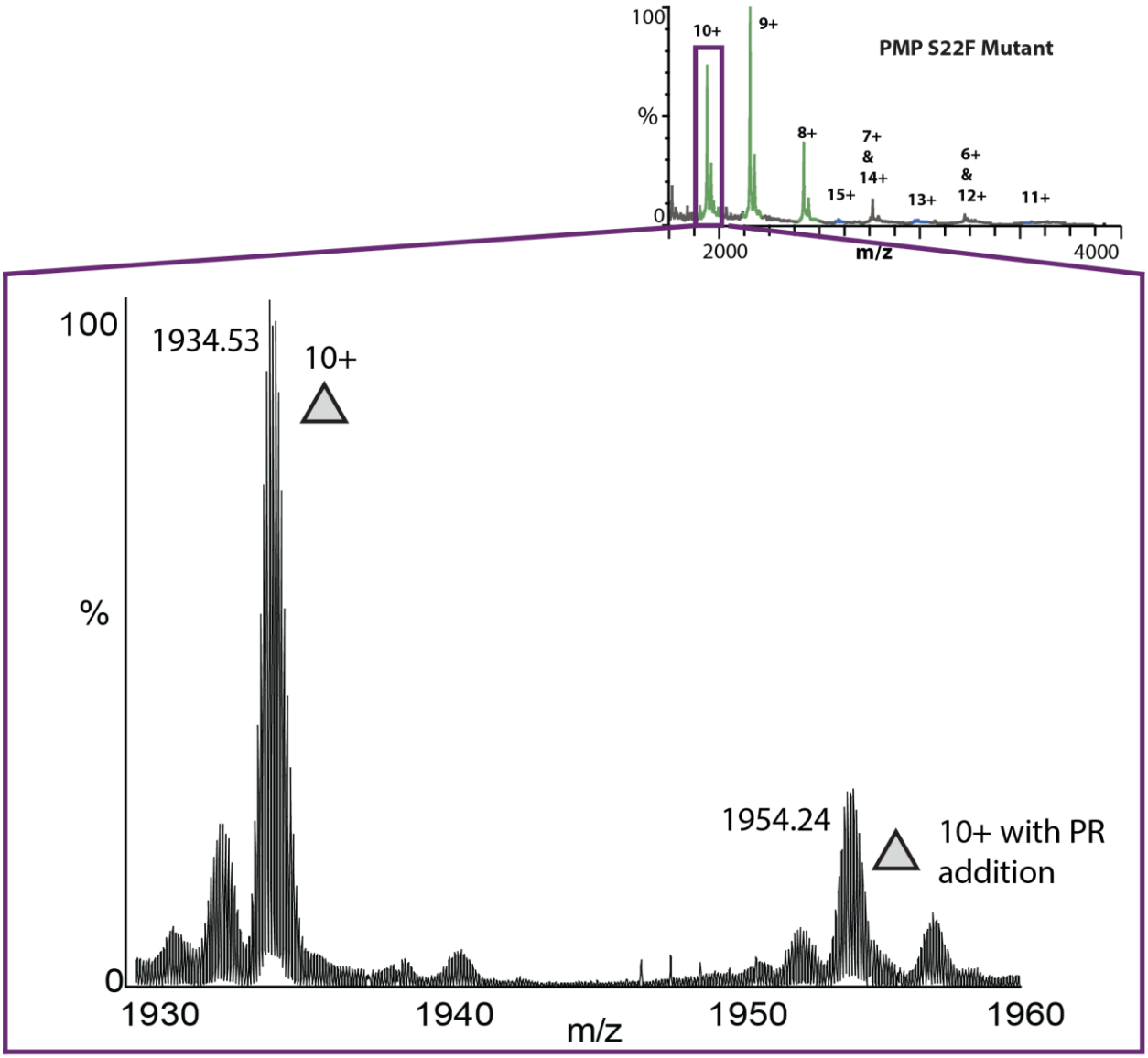

Figure II-4 HRMS identifies thrombin cleavage related additions to PMP22. PMP22 S22F without any additions can be observed at the major peak of 1934.53 Th. PMP22 S22F with an additional two residues, proline and arginine, can be observed at the minor peak of 1954.24 Th. The presence of the additional residues is likely a result of promiscuous cleavage at the thrombin cleavage site. Present in all micelle samples except G150D, which were prepared using thrombin from a different company. CIU of data extracted with narrow windows from each individual amino acid addition exhibited no significant differences, so to maximize signal intensity, all CIU include the base protein and additions.

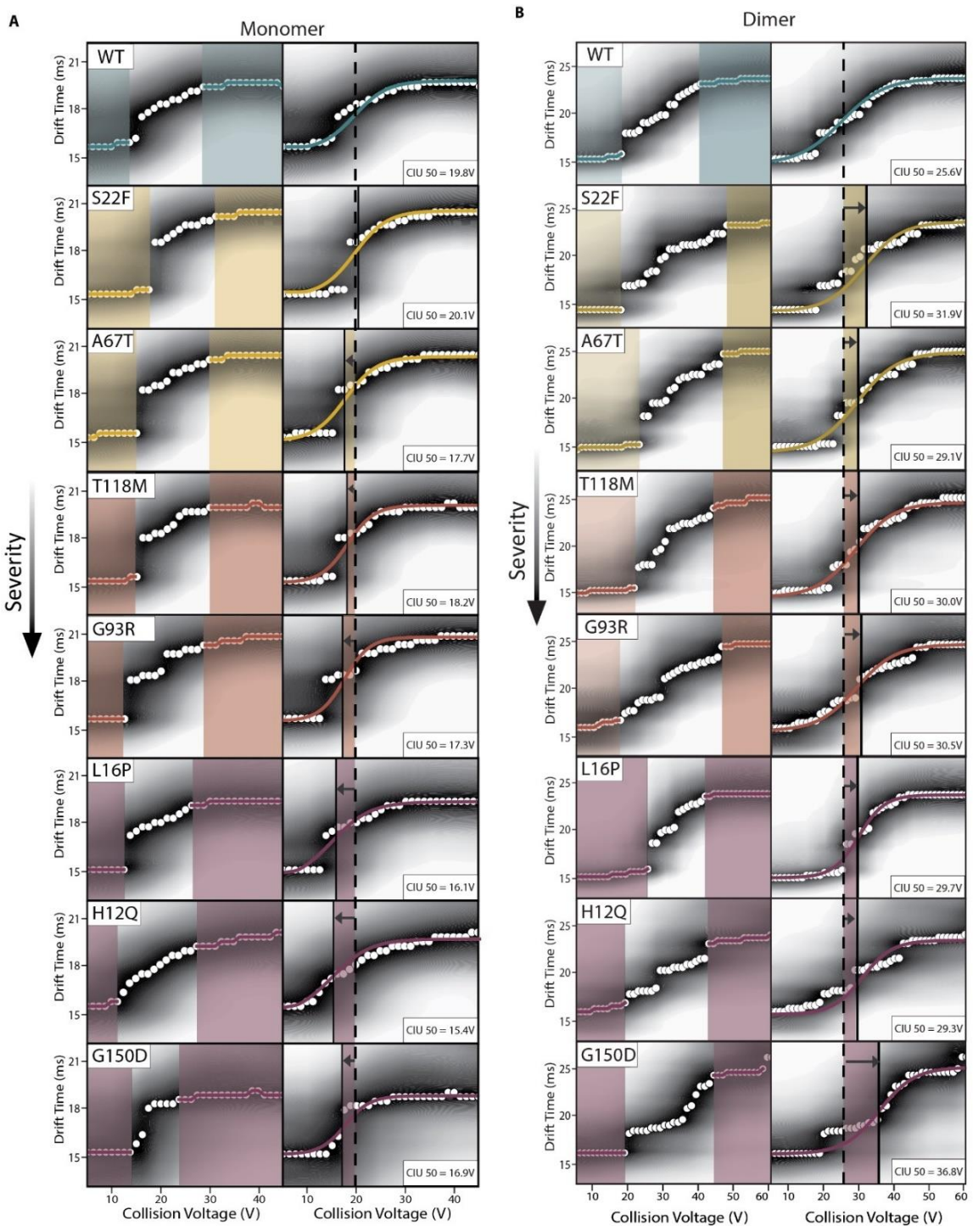

Figure II-5 9+ monomer and 13+ dimer CIU fingerprints for all variants studied in this work. Averaged, N=3. **A.** Automated CIU feature detection (left) and CIU50 sigmoidal transitions (right) for monomeric PMP22 reveal that all PMP22 mutations except for S22F show some degree of destabilization relative to WT. **B** Automated CIU feature detection (left) and CIU50 sigmoidal transitions (right) for dimeric PMP22 reveal that all PMP22 mutants are stabilized relative to WT.

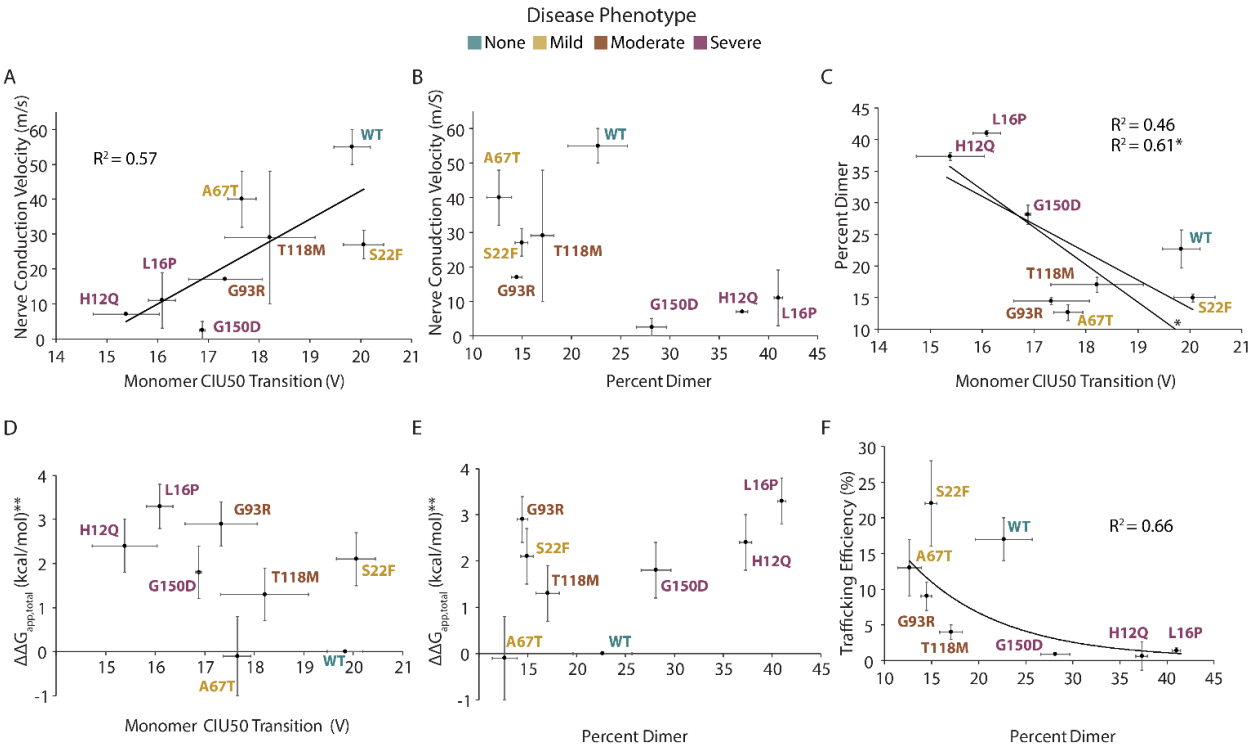

Figure II-6 Correlative plots. Nerve conduction velocities (NCV), cellular trafficking efficiencies, and $\Delta\Delta G_{app,total}$ from Schlebach et al. 2015. **A, B.** NCVs (m/s) versus monomeric PMP22 CIU50 transition values (V) and percent dimer. Linear regression analysis of NCV versus CIU50 yields R^2 = 0.57. No significant relationship between NCV and percent dimer is observed. **C.** Monomer CIU50 transition (V) versus percent dimer shows a linear inverse correlation with R^2 = 0.46 when all data are included. *When the WT PMP22 data point is removed from this data R^2 = 0.61. **D,E.** $\Delta\Delta G_{app,total}$ (kcal/mol) versus monomer CIU50 transition (V) and percent dimer show no significant correlations. **$\Delta\Delta G_{app,total}$ is a measure of Zn(II) binding free energy associated with protein folding. **F.** Percent dimer versus trafficking efficiency. An exponential fit yields R^2 = 0.66.

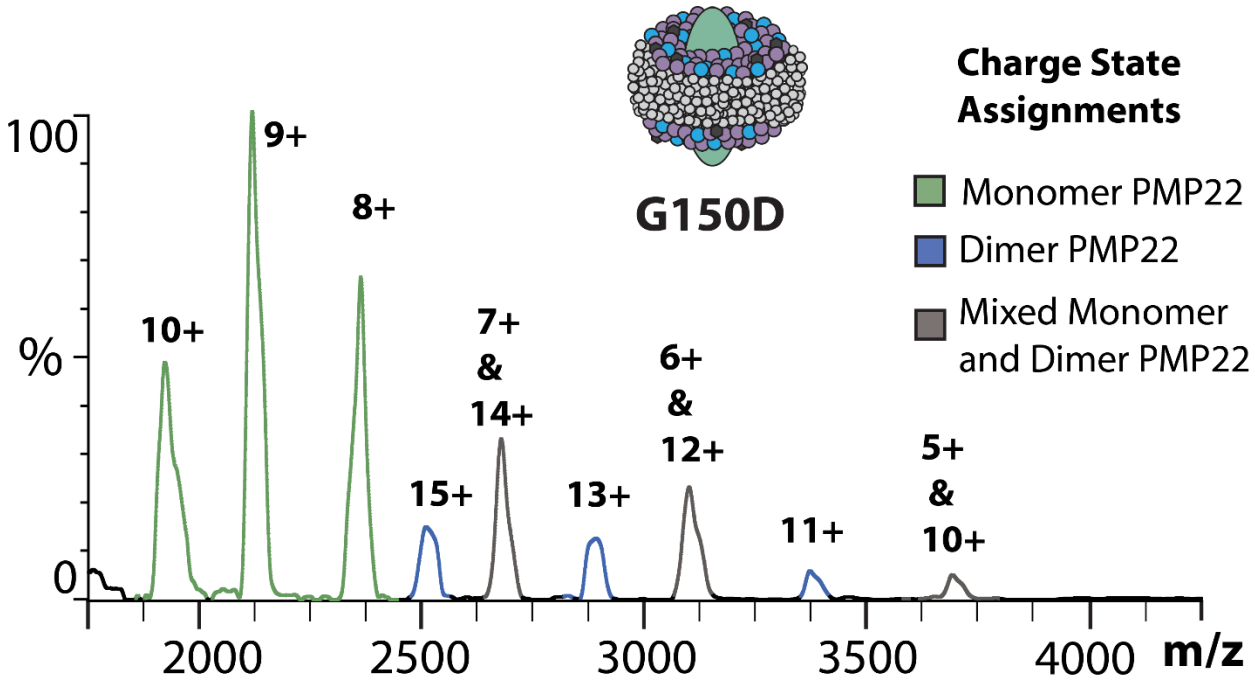

Figure II-7 Mass spectra of G150D PMP22 released from SCOR bicelles. Signals corresponding to monomer (green), dimer (blue) and mixed monomer and dimer signals (gray) are observed.

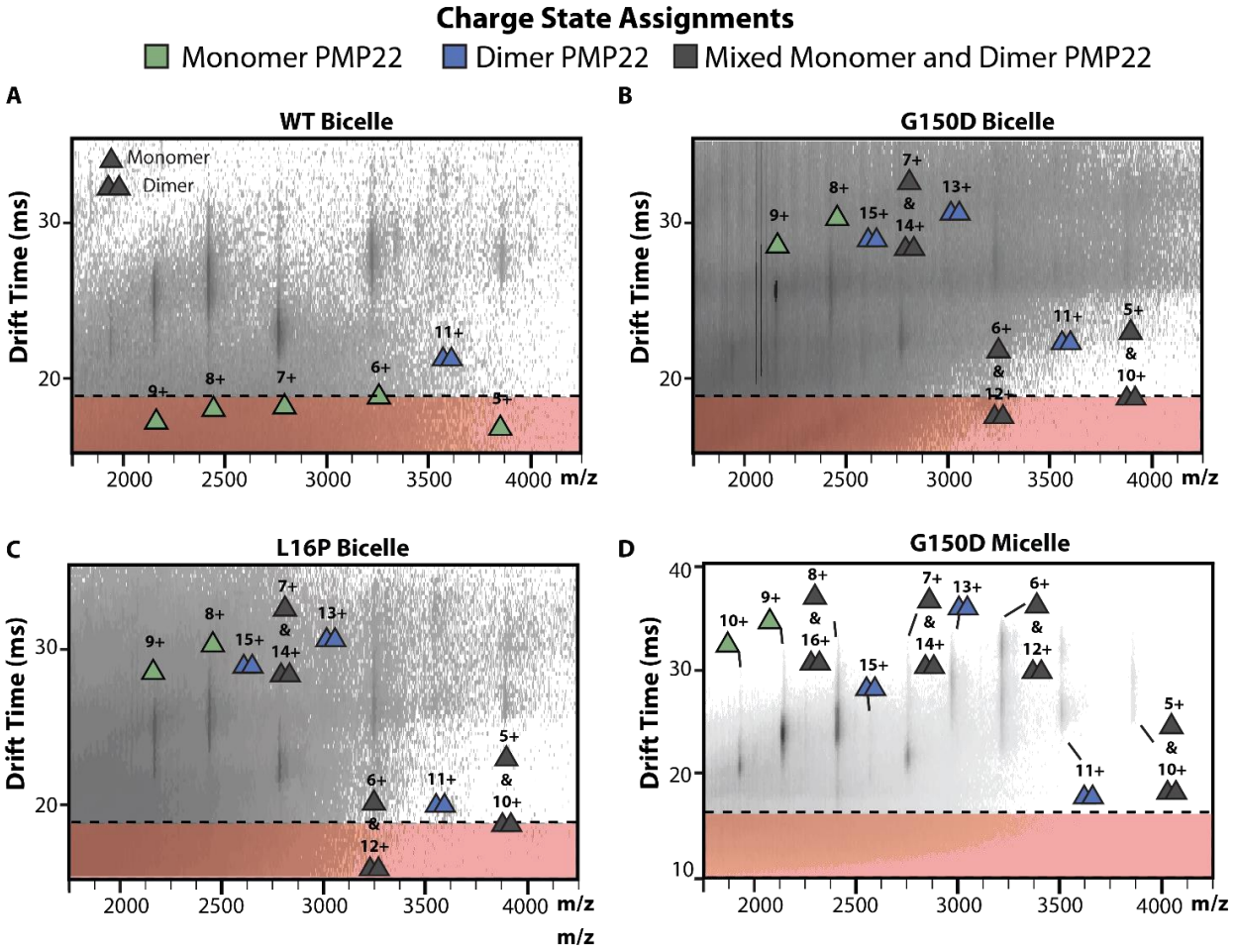

Figure II-8 Denoising of bicelle and G150D micelle data. **A, B, C.** Mass to charge versus arrival time plots of PMP22 variants liberated from SCOR bicelles exhibited noise signals associated with lipid and detergent clusters in the lower arrival time and mass to charge space. To minimize any bias towards monomeric PMP22 in the percent dimer analysis, the data extract was parameterized to extract only the arrival time distribution starting at 18 ms. **D.** G150D micelle samples were prepared and stored in glycerol, which led to a similar clustering issue in the low mass to charge range. G150D micelle arrival time distribution data was extracted starting at 17 ms. Note, this noise was not present in other micelle samples as there was no glycerol present, so no de-noising was performed.

III. Chapter 3 Supporting Methods

Preparation of SCOR bicelles: The sphingolipid and cholesterol-rich (SCOR) bicelle used in this study is composed of a lipid mixture consisting of 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), egg sphingomyelin (eSM), and cholesterol in a mole ratio of 4:2:1, and the detergent β-n-dodecylmelibioside (DDMB), such that the q ratio (lipid-to-detergent mole ratio) is 0.33. Powdered samples of DMPC (CardenPharma), egg SM (NOF Corporation), and cholesterol (Sigma-Aldrich) necessary to make a 10-mL stock solution of 6% (w/v) SCOR bicelles were first weighed out into a 15-mL glass vial and the mixture was dissolved in chloroform by vortexing. The solvent was then evaporated at 45 °C in a Smart Evaporator C1 (BioChromato) until a gel-like substance remained. The lipid mixture was then lyophilized overnight to remove the remaining solvent. The lyophilized mixture was resuspended in 15% (w/v) stock solution of DDMB (Anatrace) in milliQ water at a volume necessary to bring the q ratio to 0.33. The volume of the mixture was brought to 10 mL by adding milliQ water. The resulting mixture was then subjected to a series of freeze-thaw cycles consisting of freezing in liquid nitrogen, followed by thawing in warm sonication bath, and extensive vortexing before freezing again, until it turned clear (usually at least three cycles). The formation of SCOR bicelle was confirmed by dynamic light scattering.

Expression of PMP22: Wild type and mutant PMP22 (i.e S22F, A67T, T118M, G93R, L16P and H12Q) proteins were expressed in E. coli BL21(DE3) as described.1 Briefly, the plasmid containing cDNA encoding recombinant human PMP22 was transformed in the E. coli BL21 star (DE3) cells and grown in minimal M9 media at 20 °C with agitation. 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added when OD600 was ~0.8, to induce the expression

of the PMP22 protein. Cells were harvested by centrifugation ~22 hours after induction and the cell pellets were flash frozen until use in purification.

Purification of human PMP22 in detergent micelles: Recombinant PMP22 proteins (i.e. WT, S22F, A67T, T118M, G93R, L16P and H12Q) were purified on Ni-NTA column (Qiagen) as described1 with some modifications. To refold the protein, the EMPIGEN BB detergent (N,N-Dimethyl-N-dodecylglycine betaine, N-(Alkyl C10-C16)-N,N-dimethylglycine betaine, Sigma-Aldrich) was exchanged with Decyl-Maltoside (DM, Anatrace) by rinsing PMP22 bound to Ni-NTA resin with 200 ml of 50 mM Tris buffer (pH 8.0) contained 0.5% DM and 1mM TCEP (Tris(2-carboxyethyl)phosphine, COMPANY). PMP22 and PMP22 mutants were eluted with 300 mM Imidazole dissolved in refolding buffer (5mM Imidazole, 50mM Tris buffer pH 8.0, 0.15% DM and 1mM TCEP). Finally, after overnight thrombin cleavage the 50 ml protein solution was passed over a second Ni-NTA column that was pre-equilibrated with 50mM Tris buffer (pH 8.0) containing 0.15% DM and 1mM TCEP. Flow through was discarded and bound PMP22 was eluted with 15 mM Imidazole dissolved in 50mM Tris buffer pH 8.0, 0.15% DM and 1mM TCEP. The PMP22 proteins were concentrated to 1-2 mg/ml and stored at -80°C for use in mass spectrometry studies. The final protein concentration was estimated by measuring the absorbance at 280 using an extinction coefficient of 44,900 $M^{-1}cm^{-1}$.

Purification of PMP22 into SCOR Bicelles: Cells containing overexpressed PMP22 were resuspended in lysis buffer (75 mM Tris-HCl (pH 7.5), 300 mM NaCl, 0.2 mM EDTA) at 20 mL buffer per gram of cell pellet. Lysozyme (0.2 mg/mL), DNase (0.02 mg/mL), RNase (0.02 mg/mL), phenylmethanesulfonyl fluoride (PMSF, 1 mM), and magnesium acetate (5 mM) were added and the mixture was tumbled for at least 30 minutes at 4 °C, followed by probe-sonication on ice for 10 minutes. To dissolve the protein from the membrane and inclusion bodies, lauryl

betaine (empigen detergent, BOC Sciences) to a final concentration of 3%, as well as glycerol (15%) and tris(2-carboxyethyl)phosphine (TCEP, 1 mM), were added, and the mixture was rotated at 4□C for at least an hour. The insoluble cell debris were then removed by centrifugation of the lysate at 20,000 rpm and 4□C for 30 minutes. The supernatant was collected and mixed with 0.5 mL of pre-equilibrated HisPurTM Ni-NTA Superflow Agarose resin (Thermo Scientific) per gram of cell pellet. The resulting mixture was rotated at 4 °C for ~14 hours.

The resin was then packed into a chromatography column and washed first with 5 column volumes (CV) of purification buffer (40 mM HEPES (pH 7.5), 300 mM NaCl) containing 15% glycerol, 3% empigen, and 1 mM TCEP, followed by purification buffer containing 40 mM imidazole, 15% glycerol, 1.5% empigen, and 1 mM TCEP, until the measured absorbance at 280 nm reaches the baseline, indicating elution of weakly bound impurities (typically ~15 CV). The empigen detergent was then exchanged with n-tetradecylphosphocholine (TDPC, Anatrace) by washing the resin with 12 CV of 25 mM sodium phosphate buffer (pH 7.2) containing 0.2% TDPC and 1 mM TCEP. The PMP22 fusion protein was then eluted from the resin using 50 mM Tris buffer (pH 8) containing 500 mM imidazole, 0.2% TDPC and 1 mM TCEP.

The eluted protein was buffer exchanged twice with 20 mM sodium dihydrogen phosphate buffer (pH 7.5) containing 1 mM TCEP to get rid of the imidazole. Each buffer exchange cycle consisted of concentrating the solution to ~1 mL by centrifugation at 3,000 x g using a 10K MWCO Amicon Ultra-15 centrifugal filter device (Millipore) and diluting back to ~15 mL. The final solution was then concentrated to ~5 mL before loading into a pre-equilibrated HiTrap SP FF cation exchange column (GE Healthcare) for cation exchange chromatography. The cation exchange chromatography serves two purposes: to remove a common E. coli stress protein YodA that cooverexpressed and coeluted with the PMP22 fusion protein, and to replace the membrane

mimetic from TDPC to SCOR bicelles. After loading the sample, the column was then washed with at least 30 CV of 20 mM sodium dihydrogen phosphate buffer (pH 7.5) containing 1 mM DTT, 0.2% SCOR bicelle (q = 0.33), and 0.3 mM β-n-dodecylmelibioside (DDMB, Anatrace) to exchange the membrane mimetic. The PMP22 fusion protein in SCOR bicelles was eluted by running a gradient where the concentration of NaCl is increased from 0 to 1 M over the course of 14 CV.

Cleavage of the fusion partner of PMP22 was carried out by adding 1000 units of thrombin (Recothrom) to ~0.8 mg/mL of purified PMP22 fusion protein adjusted to pH 8.0 by addition of 20 mM Tris buffer (pH 8.0) prior to cleavage. The solution was rotated gently at room temperature overnight. The His tag-containing fusion partner, as well as uncleaved PMP22 fusion protein, were separated from the cleaved PMP22 by performing another Ni affinity chromatography. Pre-equilibrated Ni-NTA resin (at 1 mL per 10 mg fusion protein) and 20 mM imidazole were added to the cleavage reaction mixture before rotating for at least an hour at 4□C. The resin was packed into a chromatography column and the flow through containing most of the cleaved PMP22 was collected. Successive washes with 20 mM Tris buffer (pH 8.0) containing 0.2% SCOR bicelle (q = 0.33), 0.3 mM DDMB, 1 mM TCEP, and increasing concentrations of imidazole (20 mM, 30 mM, 40 mM, 250 mM) were carried out. The solution containing the cleaved PMP22 was concentrated before loading into a pre-equilibrated size exclusion column (GE Superdex 200 Increase 10/300 GL). Size exclusion chromatography was performed using 10 mM acetate buffer (pH 5.0) containing 100 mM NaCl, 0.2 % SCOR bicelle (q = 0.33), 1 mM EDTA, 5 mM TCEP, and 0.3 mM DDMB to remove empty bicelles and to exchange the buffer. Fractions containing PMP22 were combined and concentrated using 10K MWCO Amicon concentrator for use in native ion mobility-mass spectrometry experiments. The final concentration of PMP22 in SCOR bicelles

was determined by measuring the absorbance at 280 nm using an extinction coefficient of 41,940 M-1 cm-1.

IV. Chapter 4 Supporting Information

Masses of KCNQ1 VSD (Da)		
	Expected	Observed
	Monomer	
KCNQ1 VSD WT	18158	18208
KCNQ1 VSD R231C	18086	18152
KCNQ1 VSD E115G	18105	18131
KCNQ1 VSD H126L	18134	18181

Figure IV-1 . Measured masses of KCNQ1 VSD variants. Expected and observed masses for KCNQ1 VSD WT, R231C, E115G, and H126L. As peak splitting related to single amino acid truncations was observed (~175 Da), these masses are calculated from the peak without truncation, which is also the most intense peak.

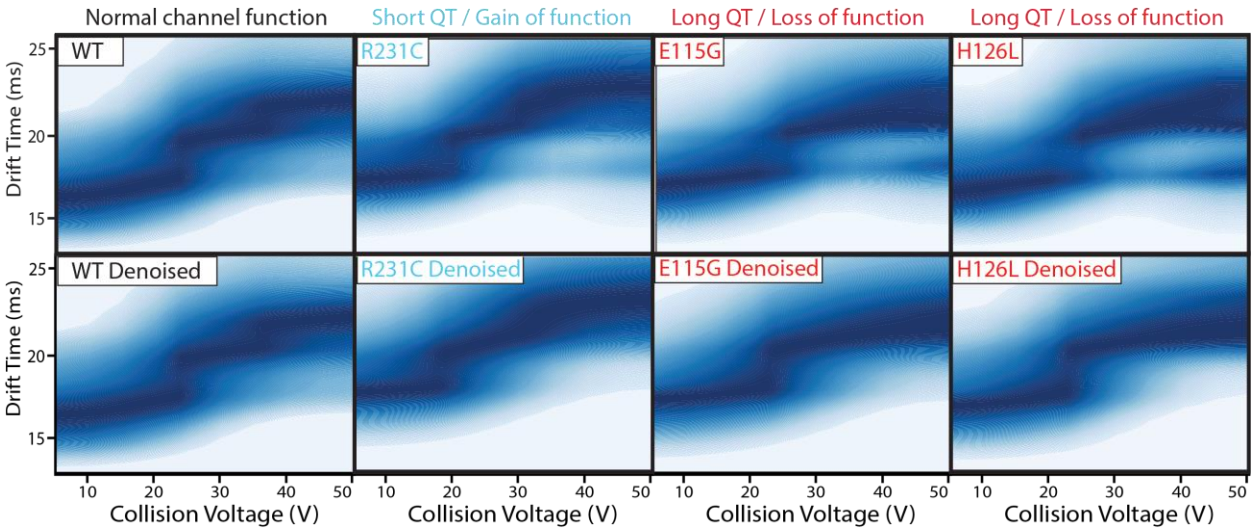

Figure IV-2 . Denoising of KCNQ1 fingerprints. Noise was noted in the low mass to charge and high voltage space, especially for mutant KCNQ1 forms which started at relatively lower concentrations before buffer and detergent exchange (top, averaged fingerprints N = 3). This noise was manually removed from each replicate's text file to yield de-noised fingerprints, for which averaged examples are shown (bottom, N=3).

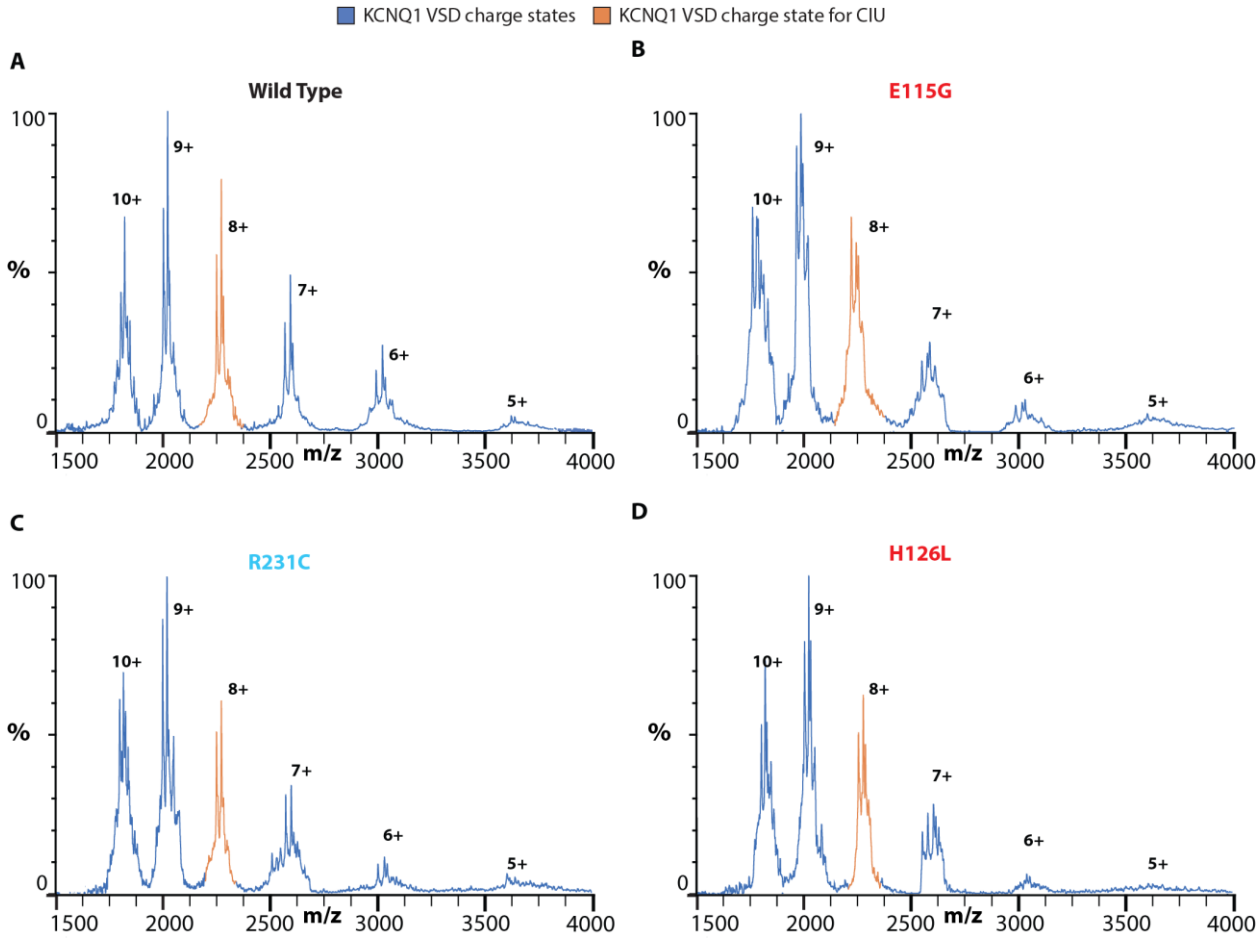

Figure IV-3 . Mass spectra of KCNQ1 variants. A-D. Mass spectra for KCNQ1 VSD WT, R231C, E115G, and H126L show the same charge state distribution of 5-10+ monomeric protein. Peak splitting of each charge state indicates a single amino acid truncation (~180 Da) for WT KCNQ1 VSD and possible additional amino acid truncations in the mutant species. CIU fingerprints generated with small extraction windows only encompassing individual truncations exhibited no significant differences, so the entire peak width was extracted for the CIU fingerprints shown in this work. These spectra are post processed through background subtraction to add in visualization.

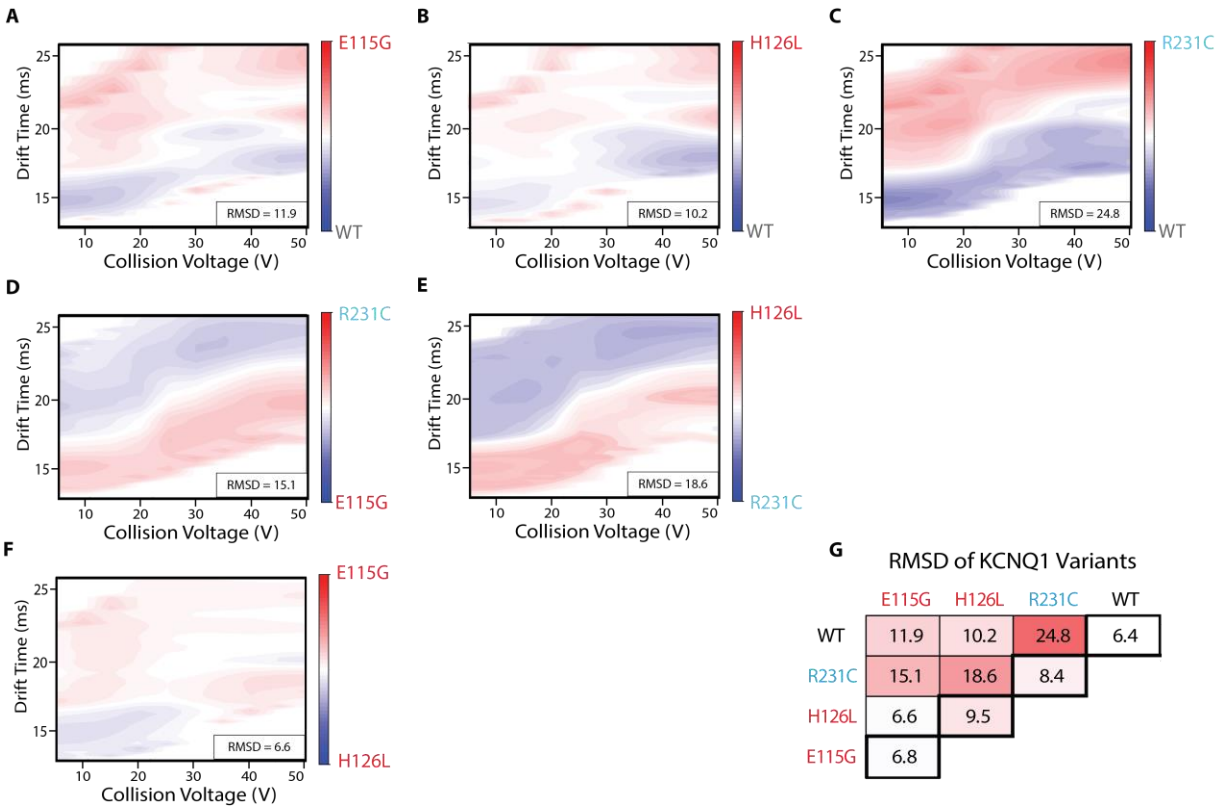

Figure IV-4 . Comprehensive difference plots for all KCNQ1 variants. **A-F**. De-noised fingerprints were averaged, N = 3, for each KCNQ1 VSD form and comprehensive difference analysis was performed using CIUSuite 2.1 to yield an RMSD for each pairwise comparison. **G**. Summary of all RMSDs and baseline replicate RMSDs (thick border).

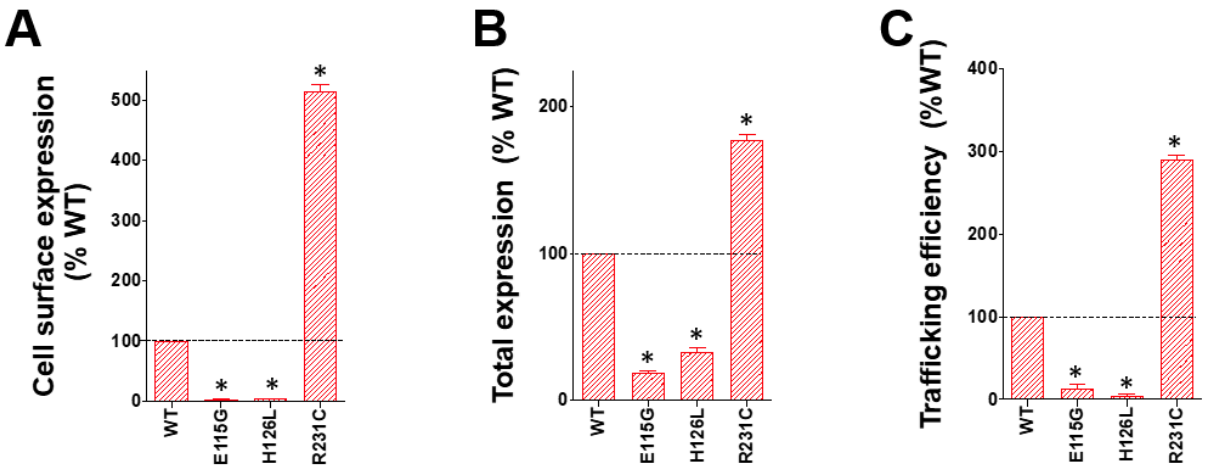

Figure IV-5 Trafficking and expression of KCNQ1 variants. Data collected by Hui Huang, Sanders lab at Vanderbilt University. The cell surface expression (A), total expression (B), and trafficking efficiency (C) of KCNQ1 mutants. HEK293 cells transiently expressing WT or mutant KCNQ1 were first treated with PE-conjugated antibody to stain surface KCNQ1 and then, following fixing and permeabilization, with Alexa Fluor 647-conjugated antibody to stain intracellular KCNQ1. Values are given as % of WT after correction for nonspecific staining. Mutant trafficking efficiency %WT=[mutant surface/mutant total]/[WT surface/WT total]*100. Results are expressed as mean ± S.E.M. of at least three independent experiments. *Significantly different from WT KCNQ1, $P<0.05$.

V. Chapter 5 Supporting Information

Measured Masses				
	Expected	Observed	Expected	Observed
	Monomer		Dimer	
A138F TSPO (Micelle)	17976	18005	35952	36033
A138F TSPO (Nanodisc)	17976	17994	35952	36008
CYP 3A4 - holo (Micelle)	57561	57783	-	-
CYP 3A4 - holo (Nanodisc)	57561	57588	-	-
PMP22 L16P (Micelle)	19260	19265	38520	38800
PMP22 L16P (Bicelle)	19260	19480	38520	38998
PMP22 WT with tag (Micelle)	30353	30214	-	-
PMP22 WT with tag (Bicelle)	30353	30237	-	-
PMP22 WT with tag (Nanodisc)	30353	30304	-	-

Figure V-1 . Expected and observed masses for all protein forms studied in Chapter 5.

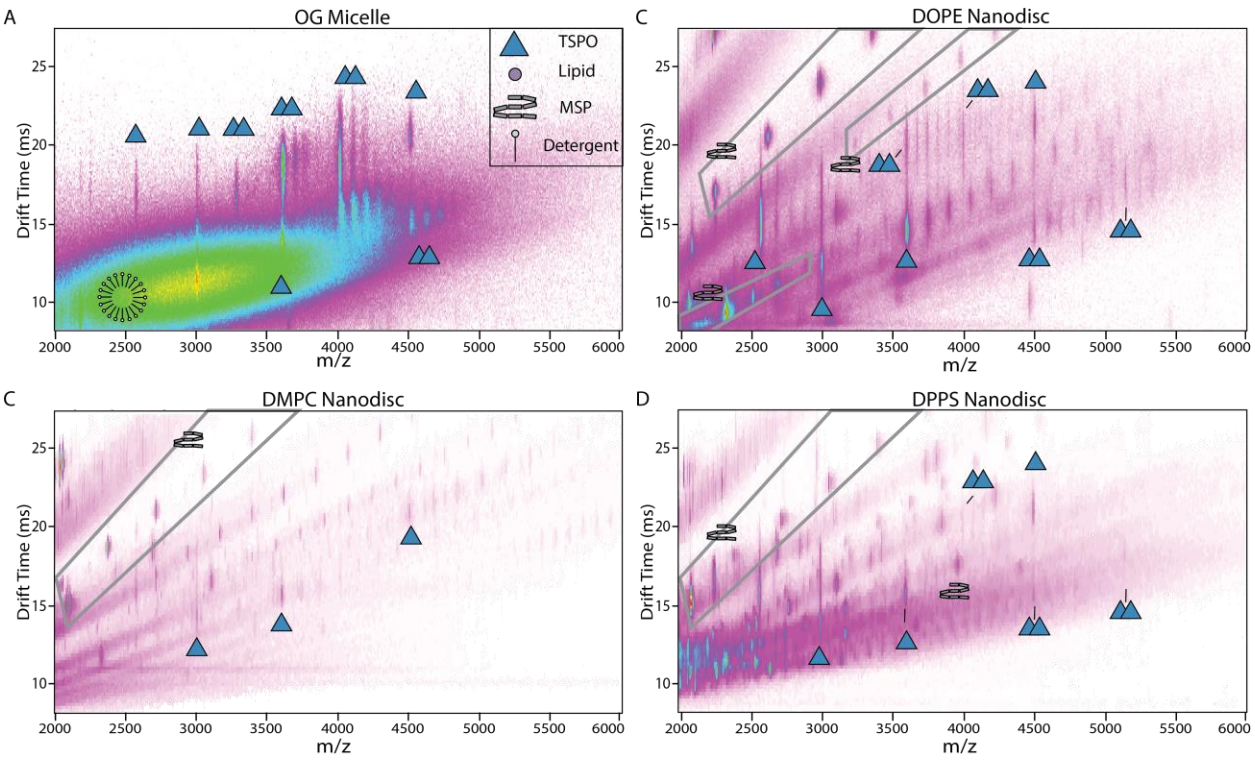

Figure V-2 IM-MS datasets of TSPO liberated from various solubilization conditions. A. Example data of TSPO liberated from OG micelles show intense signals related to dimeric complex. B,C,D. TSPO prepared in MSP1D1 nanodiscs with different lipid compositions of DMPC, DOPE, and DPPS. DMPC and DOPE IM-MS are shown at 80V trap collision voltage and DPPS at 200V trap collision voltage as no protein could be observed at lower energies for that lipid composition. Only DOPE nanodiscs exhibit significant amounts of dimeric complex.

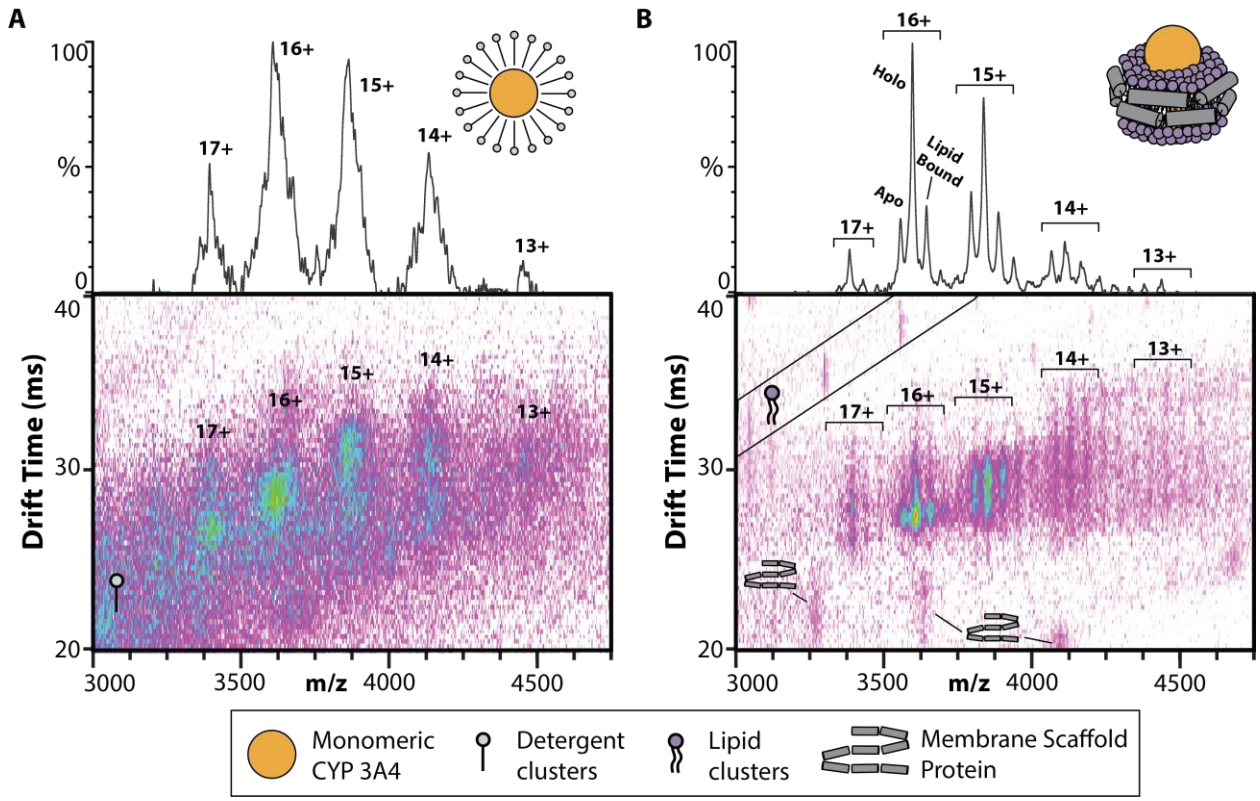

Figure V-3 Mass spectra and IM-MS datasets of CYP 3A4 liberated from micelles and nanodiscs. A. CYP 3A4 can be observed with the charge state distribution of 13-17+ when liberated from OG micelles. These peaks are broad enough that apo and holo protein forms cannot be resolved, however the centroid of the peak aligns well with weight for the holo CYP 3A4. The peak broadness could be caused by detergent adduction. B. When liberated from nanodiscs, the apo, holo, and lipid bound peaks of CYP 3A4 can be resolved for the charge states 13-17+.

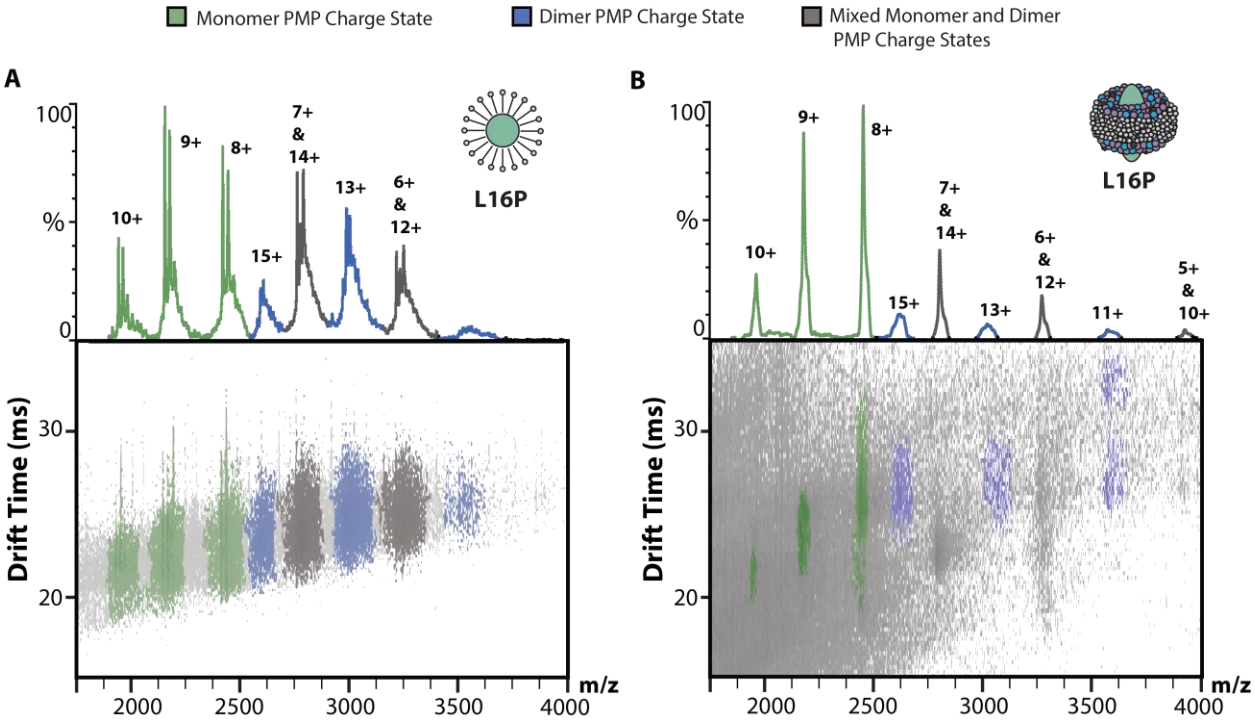

Figure V-4 . Mass spectra and IM-MS datasets of L16P PMP22 liberated from micelles and SCOR bicelles. A, B. When released from C12E8 micelles or SCOR bicelles, monomeric (green), dimeric (blue), and overlapping monomeric and dimeric (gray) charge states of L16P PMP22 can be detected. The 13+ dimeric charge state was chosen for CIU analysis.

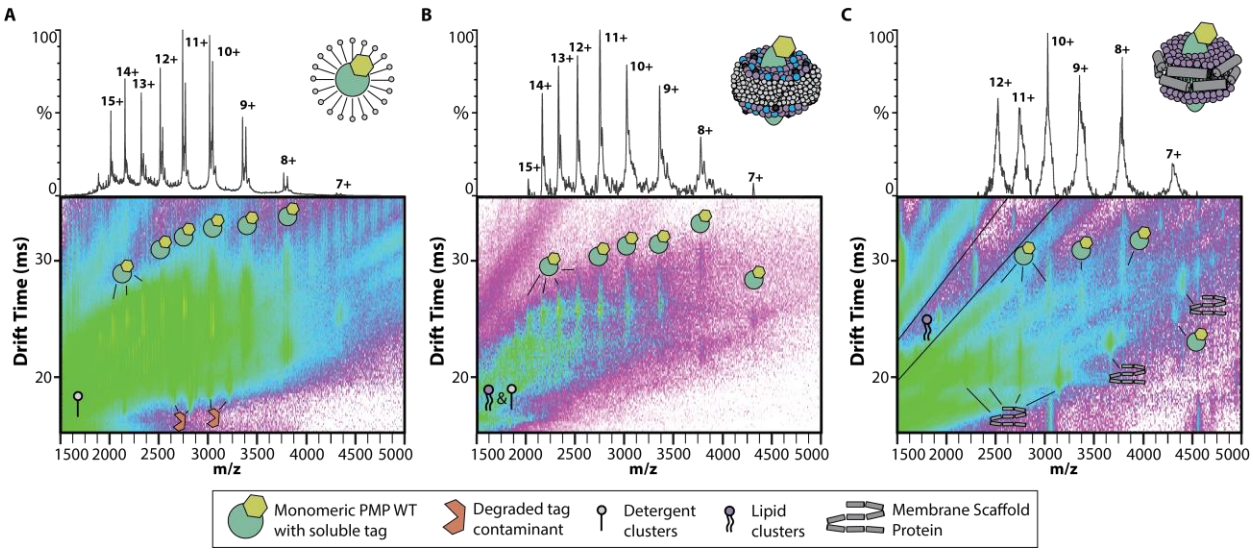

Figure V-5 Mass spectra and IM-MS datasets of WT_{tag} PMP22 liberated from micelles, SCOR bicelles, and POPC nanodiscs. A. When released from C12E8 micelles monomeric WT_{tag} PMP22 can be observed from 7-15+, along with detergent related noise and a smaller contaminant protein. This contaminant could be related to degraded forms of the 11 kDa soluble tag, but we do not observe it to interact with the protein and the signals do not interfere with the 9+ charge state CIU. ted. The 9+ monomeric charge state was chosen for CIU analysis. B. When released from SCOR bicelles monomeric WT_{tag} PMP22 can be observed from 7-15+, along with detergent and lipid related noise signals. C. When released from POPC nanodiscs, monomeric WT_{tag} PMP22 can be observed from 7-12+, however any higher charge state overlaps significantly with MSP1D1 signals and cannot be distinguished. Some charge states can be separated in IM space from MSP1D1 although they possess significant mass to charge overlap. The 9+ charge state is the most distant in mass to charge space from MSP1D1 signals and was chosen for CIU analysis.

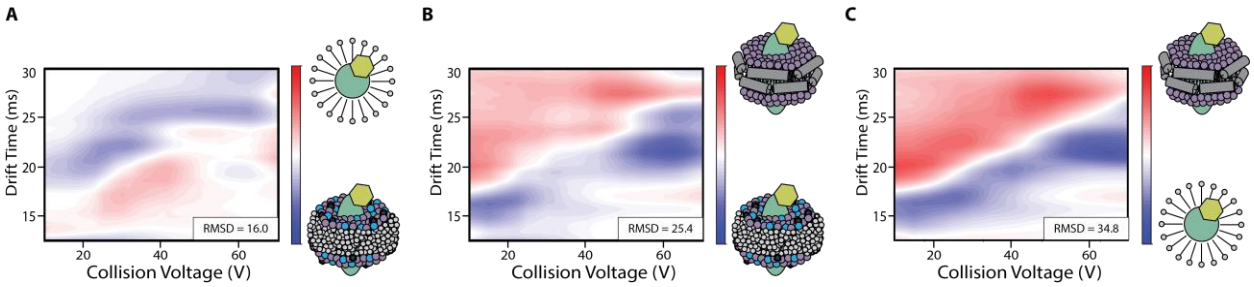

Figure V-6 Comprehensive difference analysis of WT_{tag} PMP22 9+ CIU fingerprints from C12E8 micelles, SCOR bicelles, and POPC nanodiscs. Baselines replicate RMSDs are as follows: micelle 2%, bicelle 6%, nanodisc 16%. A. Bicelle subtracted from micelle, RMSD = 16.0%. B. Bicelle subtracted from nanodisc, RMSD = 25.4%. C. Micelle subtracted from nanodisc, RMSD = 34.8%.

VI. Chapter 5 Supporting Methods

Nanodisc incorporation protocol. Except for CYP 3A4 which was incorporated into POPC MSPE3D1(-) nanodiscs using a microfluidic device, all proteins were incorporated into nanodiscs using the protocol outlined by Sligar et al 2010.[233] Successful incorporation conditions for TSPO and WT_{tag} are as follows: For TSPO, the DOPE:MSP1D1(+histag):TSPO ratio was 65:4:1, which was incubated at DOPE T_m without TSPO for 30 minutes. The TSPO was added and incubated for additional 30 minutes before being added to detergent removal resin (Biobeads) and incubated for 4 hours. Samples were prepared for nMS using 6 kDa or 40 kDa buffer exchange spin tubes. For WT_{tag} PMP22: the POPC:MSP1D1(-): WT_{tag} PMP22 ratio was 65:4:1 and the mixture was incubated on ice without WT_{tag} PMP22 for 30 minutes. Then WT_{tag} PMP22 and extra secondary detergent (DDM) were added to the mixture and incubated for additional 60 minutes. This was added to detergent removal resin (Amberlite XAD) and incubated at 6° C while shaking overnight. To remove empty nanodiscs, samples were then incubated with an equilibrated nickel affinity resin for 30 minutes and eluted with 1 mM imidazole. Size exclusion chromatography was performed on the eluent using a Superdex increase 10/300 gL prepacked Tricorn column equilibrated in 200 mM ammonium acetate pH 7.4. Samples were concentrated using 10 kDa concentrators prior to nMS analysis.

SCOR bicelles incorporation protocol. Cells containing overexpressed PMP22 were resuspended in lysis buffer (75 mM Tris-HCl (pH 7.5), 300 mM NaCl, 0.2 mM EDTA) at 20 mL buffer per gram of cell pellet. Lysozyme (0.2 mg/mL), DNase (0.02 mg/mL), RNase (0.02 mg/mL), phenylmethanesulfonyl fluoride (PMSF, 1 mM), and magnesium acetate (5 mM) were added and the mixture was tumbled for at least 30 minutes at 4°C, followed by probe-sonication on ice for 10 minutes. To dissolve the protein from the membrane and inclusion bodies, lauryl

betaine (empigen detergent, BOC Sciences) to a final concentration of 3%, as well as glycerol (15%) and tris(2-carboxyethyl)phosphine (TCEP, 1 mM), were added, and the mixture was rotated at 4□C for at least an hour. The insoluble cell debris were then removed by centrifugation of the lysate at 20,000 rpm and 4°C for 30 minutes. The supernatant was collected and mixed with 0.5 mL of pre-equilibrated HisPurTM Ni-NTA Superflow Agarose resin (Thermo Scientific) per gram of cell pellet. The resulting mixture was rotated at 4 °C for ~14 hours.

The resin was then packed into a chromatography column and washed first with 5 column volumes (CV) of purification buffer (40 mM HEPES (pH 7.5), 300 mM NaCl) containing 15% glycerol, 3% empigen, and 1 mM TCEP, followed by purification buffer containing 40 mM imidazole, 15% glycerol, 1.5% empigen, and 1 mM TCEP, until the measured absorbance at 280 nm reaches the baseline, indicating elution of weakly bound impurities (typically ~15 CV). The empigen detergent was then exchanged with n-tetradecylphosphocholine (TDPC, Anatrace) by washing the resin with 12 CV of 25 mM sodium phosphate buffer (pH 7.2) containing 0.2% TDPC and 1 mM TCEP. The PMP22 fusion protein was then eluted from the resin using 50 mM Tris buffer (pH 8) containing 500 mM imidazole, 0.2% TDPC and 1 mM TCEP.

The eluted protein was buffer exchanged twice with 20 mM sodium dihydrogen phosphate buffer (pH 7.5) containing 1 mM TCEP to remove the imidazole. Each buffer exchange cycle consisted of concentrating the solution to ~1 mL by centrifugation at 3,000 x g using a 10 K MWCO Amicon Ultra-15 centrifugal filter device (Millipore) and diluting back to ~15 mL. The final solution was then concentrated to ~5 mL before loading into a pre-equilibrated HiTrap SP FF cation exchange column (GE Healthcare) for cation exchange chromatography. The cation exchange chromatography serves two purposes: to remove a common E. coli stress protein YodA that cooverexpressed and coeluted with the PMP22 fusion protein, and to replace the membrane

mimetic from TDPC to SCOR bicelles. After loading the sample, the column was then washed with at least 30 CV of 20 mM sodium dihydrogen phosphate buffer (pH 7.5) containing 1 mM DTT, 0.2% SCOR bicelle (q = 0.33), and 0.3 mM β-n-dodecylmelibioside (DDMB, Anatrace) to exchange the membrane mimetic. The PMP22 fusion protein in SCOR bicelles was eluted by running a gradient where the concentration of NaCl is increased from 0 to 1 M over the course of 14 CV.

Cleavage of the fusion partner of PMP22 was carried out by adding 1000 units of thrombin (Recothrom) to ~0.8 mg/mL of purified PMP22 fusion protein adjusted to pH 8.0 by addition of 20 mM Tris buffer (pH 8.0) prior to cleavage. The solution was rotated gently at room temperature overnight. The His tag-containing fusion partner, as well as uncleaved PMP22 fusion protein, were separated from the cleaved PMP22 by performing another Ni affinity chromatography. Pre-equilibrated Ni-NTA resin (at 1 mL per 10 mg fusion protein) and 20 mM imidazole were added to the cleavage reaction mixture before rotating for at least an hour at 4 °C. The resin was packed into a chromatography column and the flow through containing most of the cleaved PMP22 was collected. Successive washes with 20 mM Tris buffer (pH 8.0) containing 0.2% SCOR bicelle (q = 0.33), 0.3 mM DDMB, 1 mM TCEP, and increasing concentrations of imidazole (20 mM, 30 mM, 40 mM, 250 mM) were carried out. The solution containing the cleaved PMP22 was concentrated before loading into a pre-equilibrated size exclusion column (GE Superdex 200 Increase 10/300 GL). Size exclusion chromatography was performed using 10 mM acetate buffer (pH 5.0) containing 100 mM NaCl, 0.2 % SCOR bicelle (q = 0.33), 1 mM EDTA, 5 mM TCEP, and 0.3 mM DDMB to remove empty bicelles and to exchange the buffer. Fractions containing PMP22 were combined and concentrated using 10K MWCO Amicon concentrator for use in native

ion mobility-mass spectrometry experiments. The final concentration of PMP22 in SCOR bicelles was determined by measuring References

References

(1) Muller, M. P.; Jiang, T.; Sun, C.; Lihan, M.; Pant, S.; Mahinthichaichan, P.; Trifan, A.; Tajkhorshid, E. Characterization of Lipid-Protein Interactions and Lipid-Mediated Modulation of Membrane Protein Function through Molecular Simulation. *Chem. Rev.* **2019**, *119* (9), 6086–6161. https://doi.org/10.1021/acs.chemrev.8b00608.

(2) Dowhan, W.; Vitrac, H.; Bogdanov, M. Lipid-Assisted Membrane Protein Folding and Topogenesis. *Protein J.* **2019**, *38*, 274–288. https://doi.org/10.1007/s10930-019-09826-7.

(3) Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P. A Comprehensive Map of Molecular Drug Targets. *Nat. Rev. Drug Discov.* **2016**, *16* (1), 19–34. https://doi.org/10.1038/nrd.2016.230.

(4) Rudnick, G.; Sandtner, W. Serotonin Transport in the 21st Century. *The Journal of general physiology*. NLM (Medline) November 4, 2019, pp 1248–1264. https://doi.org/10.1085/jgp.201812066.

(5) Taciak, P. P.; Lysenko, N.; Mazurek, A. P. Drugs Which Influence Serotonin Transporter and Serotonergic Receptors: Pharmacological and Clinical Properties in the Treatment of Depression. *Pharmacological Reports*. Elsevier B.V. February 1, 2018, pp 37–46. https://doi.org/10.1016/j.pharep.2017.07.011.

(6) Möller, I. R.; Slivacka, M.; Nielsen, A. K.; Rasmussen, S. G. F.; Gether, U.; Loland, C. J.; Rand, K. D. Conformational Dynamics of the Human Serotonin Transporter during Substrate and Drug Binding. *Nat. Commun.* **2019**, *10* (1), 1–13. https://doi.org/10.1038/s41467-019-09675-z.

(7) Coleman, J. A.; Green, E. M.; Gouaux, E. X-Ray Structures and Mechanism of the Human Serotonin Transporter. *Nature* **2016**, *532* (7599), 334–339. https://doi.org/10.1038/nature17629.

(8) Coleman, J. A.; Yang, D.; Zhao, Z.; Wen, P. C.; Yoshioka, C.; Tajkhorshid, E.; Gouaux, E. Serotonin Transporter–Ibogaine Complexes Illuminate Mechanisms of Inhibition and Transport. *Nature*. Nature Publishing Group April 24, 2019, pp 141–145. https://doi.org/10.1038/s41586-019-1135-1.

(9) Coleman, J. A.; Gouaux, E. Structural Basis for Recognition of Diverse Antidepressants by the Human Serotonin Transporter. *Nat. Struct. Mol. Biol.* **2018**, *25* (2), 170–175. https://doi.org/10.1038/s41594-018-0026-8.

(10) Cecchetti, C.; Pyle, E.; Byrne, B. Transporter Oligomerisation: Roles in Structure and Function. *Biochemical Society Transactions*. Portland Press Ltd February 28, 2018, pp 433–440. https://doi.org/10.1042/BST20180316.

(11) Notter, T.; Coughlin, J. M.; Sawa, A.; Meyer, U. Reconceptualization of Translocator Protein as a Biomarker of Neuroinflammation in Psychiatry. *Molecular Psychiatry*. Nature Publishing Group January 1, 2018, pp 36–47. https://doi.org/10.1038/mp.2017.232.

(12) Notter, T.; Coughlin, J. M.; Gschwind, T.; Weber-Stadlbauer, U.; Wang, Y.; Kassiou, M.;

Vernon, A. C.; Benke, D.; Pomper, M. G.; Sawa, A.; Meyer, U. Translational Evaluation of Translocator Protein as a Marker of Neuroinflammation in Schizophrenia. *Mol. Psychiatry* **2018**, *23* (2), 323–334. https://doi.org/10.1038/mp.2016.248.

(13) Janczar, K.; Su, Z.; Raccagni, I.; Anfosso, A.; Kelly, C.; Durrenberger, P. F.; Gerhard, A.; Roncaroli, F. The 18-KDa Mitochondrial Translocator Protein in Gliomas: From the Bench to Bedside. *Biochem. Soc. Trans.* **2015**, *43* (4), 579–585. https://doi.org/10.1042/BST20150064.

(14) Li, F.; Liu, J.; Zheng, Y.; Garavito, R. M.; Ferguson-Miller, S. Crystal Structures of Translocator Protein (TSPO) and Mutant Mimic of a Human Polymorphism. *Science (80-.).* **2015**, *347* (6221), 555–558. https://doi.org/10.1126/science.1260590.

(15) Morin, D.; Musman, J.; Pons, S.; Berdeaux, A.; Ghaleh, B. Mitochondrial Translocator Protein (TSPO): From Physiology to Cardioprotection. *Biochem. Pharmacol.* **2016**, *105*, 1–13. https://doi.org/10.1016/j.bcp.2015.12.003.

(16) Batoko, H.; Veljanovski, V.; Jurkiewicz, P. Enigmatic Translocator Protein (TSPO) and Cellular Stress Regulation. *Trends in Biochemical Sciences*. Elsevier Ltd September 1, 2015, pp 497–503. https://doi.org/10.1016/j.tibs.2015.07.001.

(17) Fusco, C.; Spagnoli, C.; Salerno, G. G.; Pavlidis, E.; Frattini, D.; Pisani, F. Hereditary Neuropathy with Liability to Pressure Palsy (HNPP): Report of a Family with a New Point Mutation in PMP22 Gene. *Ital. J. Pediatr.* **2017**, *43* (1), 97. https://doi.org/10.1186/s13052-017-0414-4.

(18) Matsunami, N.; Smith, B.; Ballard, L.; Lensch, M. W.; Robertson, M.; Albertsen, H.; Hanemann, C. O.; Müller, H. W.; Bird, T. D.; White, R.; Chance, P. F. Peripheral Myelin Protein–22 Gene Maps in the Duplication in Chromosome 17p11.2 Associated with Charcot–Marie–Tooth 1A. *Nat. Genet.* **1992**, *1* (3), 176–179. https://doi.org/10.1038/ng0692-176.

(19) Roa, B. B.; Dyck, P. J.; Marks, H. G.; Chance, P. F.; Lupski, J. R. Dejerine−Sottas Syndrome Associated with Point Mutation in the Peripheral Myelin Protein 22 (PMP22) Gene. *Nat. Genet.* **1993**, *5* (3), 269–273. https://doi.org/10.1038/ng1193-269.

(20) Schlebach, J. P.; Narayan, M.; Alford, C.; Mittendorf, K. F.; Carter, B. D.; Li, J.; Sanders, C. R. Conformational Stability and Pathogenic Misfolding of the Integral Membrane Protein PMP22. *J. Am. Chem. Soc.* **2015**, *137* (27), 8758–8768. https://doi.org/10.1021/jacs.5b03743.

(21) Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Section 3.4, Membrane Proteins. In *Molecular Cell Biology. 4th Edition*; W. H. Freeman: New York, 2000.

(22) Von Heijne, G. Membrane-Protein Topology. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group December 8, 2006, pp 909–918. https://doi.org/10.1038/nrm2063.

(23) Baker, J. A.; Wong, W. C.; Eisenhaber, B.; Warwicker, J.; Eisenhaber, F. Charged Residues next to Transmembrane Regions Revisited: "Positive-inside Rule" Is Complemented by the "Negative inside Depletion/Outside Enrichment Rule." *BMC Biol.* **2017**, *15* (1). https://doi.org/10.1186/s12915-017-0404-4.

(24) Von Heijne, G. *Membrane Protein Structure Prediction Hydrophobicity Analysis and the Positive-inside Rule*; 1992; Vol. 225.

(25) Almén, M. S.; Nordström, K. J. V.; Fredriksson, R.; Schiöth, H. B. Mapping the Human Membrane Proteome: A Majority of the Human Membrane Proteins Can Be Classified

According to Function and Evolutionary Origin. *BMC Biol.* **2009**, *7*, 50. https://doi.org/10.1186/1741-7007-7-50.
(26) Berg, J. M.; Tymoczko, J. L.; Gatto, G. J.; Lubert, S. *Biochemistry*, 8th ed.; W. H. Freeman, 2015.
(27) Matthies, D.; Bae, C.; Toombes, G. E. S.; Fox, T.; Bartesaghi, A.; Subramaniam, S.; Swartz, K. J. Single-Particle Cryo-EM Structure of a Voltage-Activated Potassium Channel in Lipid Nanodiscs. *Elife* **2018**, *7*. https://doi.org/10.7554/eLife.37558.
(28) Huang, C.-Y.; Olieric, V.; Ma, P.; Howe, N.; Vogeley, L.; Liu, X.; Warshamanage, R.; Weinert, T.; Panepucci, E.; Kobilka, B.; Diederichs, K.; Wang, M.; Caffrey, M. In Meso in Situ Serial X-Ray Crystallography of Soluble and Membrane Proteins at Cryogenic Temperatures. *Res. Pap. Acta Cryst* **2016**, *72*, 93–112. https://doi.org/10.1107/S2059798315021683.
(29) Monk, B. C.; Tomasiak, T. M.; Keniya, M. V; Huschmann, F. U.; Tyndall, J. D. A.; O'connell Iii, J. D.; Cannon, R. D.; Mcdonald, J. G.; Rodriguez, A.; Finer-Moore, J. S.; Stroud, R. M. Architecture of a Single Membrane Spanning Cytochrome P450 Suggests Constraints That Orient the Catalytic Domain Relative to a Bilayer. https://doi.org/10.1073/pnas.1324245111.
(30) Šrejber, M.; Navrátilová, V.; Paloncýová, M.; Bazgier, V.; Berka, K.; Anzenbacher, P.; Otyepka, M. Membrane-Attached Mammalian Cytochromes P450: An Overview of the Membrane's Effects on Structure, Drug Binding, and Interactions with Redox Partners. *Journal of Inorganic Biochemistry*. Elsevier Inc. June 1, 2018, pp 117–136. https://doi.org/10.1016/j.jinorgbio.2018.03.002.
(31) Wright, W. C.; Chenge, J.; Chen, T. Structural Perspectives of the CYP3A Family and Their Small Molecule Modulators in Drug Metabolism. *Liver Research*. KeAi Communications Co. December 1, 2019, pp 132–142. https://doi.org/10.1016/j.livres.2019.08.001.
(32) Shimizu, K.; Cao, W.; Saad, G.; Shoji, M.; Terada, T. Comparative Analysis of Membrane Protein Structure Databases. *Biochimica et Biophysica Acta - Biomembranes*. Elsevier B.V. May 1, 2018, pp 1077–1091. https://doi.org/10.1016/j.bbamem.2018.01.005.
(33) Moraes, I.; Evans, G.; Sanchez-Weatherby, J.; Newstead, S.; Stewart, P. D. S. Membrane Protein Structure Determination - The next Generation. *Biochimica et Biophysica Acta - Biomembranes*. Elsevier January 1, 2014, pp 78–87. https://doi.org/10.1016/j.bbamem.2013.07.010.
(34) Lacapère, J. J.; Pebay-Peyroula, E.; Neumann, J. M.; Etchebest, C. Determining Membrane Protein Structures: Still a Challenge! *Trends Biochem. Sci.* **2007**, *32* (6), 259–270. https://doi.org/10.1016/j.tibs.2007.04.001.
(35) Weekes, M. P.; Antrobus, R.; Lill, J. R.; Duncan, L. M.; Hör, S.; Lehner, P. J. *Comparative Analysis of Techniques to Purify Plasma Membrane Proteins*; 2010; Vol. 21.
(36) Rawlings, A. E. Membrane Proteins: Always an Insoluble Problem? *Biochem. Soc. Trans.* **2016**, *44* (3), 790–795. https://doi.org/10.1042/BST20160025.
(37) Yin, H.; Flynn, A. D. Drugging Membrane Protein Interactions. **2016**. https://doi.org/10.1146/annurev-bioeng-092115-025322.
(38) Cutolo, P.; Basdevant, N.; Bernadat, G.; Bachelerie, F.; Ha-Duong, T. Interaction of Chemokine Receptor CXCR4 in Monomeric and Dimeric State with Its Endogenous Ligand CXCL12: Coarse-Grained Simulations Identify Differences. *J. Biomol. Struct.*

Dyn. **2017**, *35* (2), 399–412. https://doi.org/10.1080/07391102.2016.1145142.
(39) Muller, C.; Morales, P.; Reggio, P. H. Cannabinoid Ligands Targeting TRP Channels. *Front. Mol. Neurosci.* **2019**, *11*, 487. https://doi.org/10.3389/fnmol.2018.00487.
(40) Gupta, K.; Donlan, J. A. C.; Hopper, J. T. S.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The Role of Interfacial Lipids in Stabilizing Membrane Protein Oligomers. *Nature* **2017**, *541* (7637), 421–424. https://doi.org/10.1038/nature20820.
(41) Muller, M. P.; Jiang, T.; Sun, C.; Lihan, M.; Pant, S.; Mahinthichaichan, P.; Trifan, A.; Tajkhorshid, E. Characterization of Lipid-Protein Interactions and Lipid-Mediated Modulation of Membrane Protein Function through Molecular Simulation. *Chemical Reviews*. American Chemical Society May 8, 2019, pp 6086–6161. https://doi.org/10.1021/acs.chemrev.8b00608.
(42) Haffke, M.; Duckely, M.; Bergsdorf, C.; Jaakola, V. P.; Shrestha, B. Development of a Biochemical and Biophysical Suite for Integral Membrane Protein Targets: A Review. *Protein Expr. Purif.* **2020**, *167*, 105545. https://doi.org/10.1016/j.pep.2019.105545.
(43) Cheng, X.; Kim, J. K.; Kim, Y.; Bowie, J. U.; Im, W. Molecular Dynamics Simulation Strategies for Protein-Micelle Complexes. *Biochim. Biophys. Acta - Biomembr.* **2016**, *1858* (7), 1566–1572. https://doi.org/10.1016/j.bbamem.2015.12.012.
(44) Warschawski, D. E.; Arnold, A. A.; Beaugrand, M.; Gravel, A.; Chartrand, É.; Marcotte, I. Choosing Membrane Mimetics for NMR Structural Studies of Transmembrane Proteins. *Biochimica et Biophysica Acta - Biomembranes*. Elsevier August 1, 2011, pp 1957–1974. https://doi.org/10.1016/j.bbamem.2011.03.016.
(45) Cross, T. A.; Sharma, M.; Yi, M.; Zhou, H. X. Influence of Solubilizing Environments on Membrane Protein Structures. *Trends in Biochemical Sciences*. Elsevier Current Trends February 1, 2011, pp 117–125. https://doi.org/10.1016/j.tibs.2010.07.005.
(46) Lyukmanova, E. N.; Shenkarev, Z. O.; Khabibullina, N. F.; Kopeina, G. S.; Shulepko, M. A.; Paramonov, A. S.; Mineev, K. S.; Tikhonov, R. V.; Shingarova, L. N.; Petrovskaya, L. E.; Dolgikh, D. A.; Arseniev, A. S.; Kirpichnikov, M. P. Lipid-Protein Nanodiscs for Cell-Free Production of Integral Membrane Proteins in a Soluble and Folded State: Comparison with Detergent Micelles, Bicelles and Liposomes. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818* (3), 349–358. https://doi.org/10.1016/j.bbamem.2011.10.020.
(47) Dü, U. H. N.; Gildenberg, M.; Ramamoorthy, A. The Magic of Bicelles Lights Up Membrane Protein Structure. **2012**. https://doi.org/10.1021/cr300061w.
(48) Bayburt, T. H.; Sligar, S. G. Membrane Protein Assembly into Nanodiscs. *FEBS Lett.* **2010**, *584* (9), 1721–1727. https://doi.org/10.1016/j.febslet.2009.10.024.
(49) Arnold, A.; Labrot, T.; Oda, R.; Dufourc, E. J. Cation Modulation of Bicelle Size and Magnetic Alignment as Revealed by Solid-State NMR and Electron Microscopy. *Biophys. J.* **2002**, *83* (5), 2667–2680. https://doi.org/10.1016/S0006-3495(02)75276-4.
(50) Whiles, J. A.; Glover, K. J.; Vold, R. R.; Komives, E. A. Methods for Studying Transmembrane Peptides in Bicelles: Consequences of Hydrophobic Mismatch and Peptide Sequence. *J. Magn. Reson.* **2002**, *158* (1–2), 149–156. https://doi.org/10.1016/S1090-7807(02)00068-X.
(51) Barnaba, C.; Ramamoorthy, A. Picturing the Membrane-Assisted Choreography of Cytochrome P450 with Lipid Nanodiscs. *ChemPhysChem* **2018**, *19* (20), 2603–2613. https://doi.org/10.1002/cphc.201800444.
(52) Henrich, E.; Peetz, O.; Hein, C.; Laguerre, A.; Hoffmann, B.; Hoffmann, J.; Dötsch, V.;

Bernhard, F.; Morgner, N. Analyzing Native Membrane Protein Assembly in Nanodiscs by Combined Non-Covalent Mass Spectrometry and Synthetic Biology. *Elife* **2017**, *6*, 1–19. https://doi.org/10.7554/eLife.20954.

(53) Wallin, E.; Heijne, G. Von. Genome-Wide Analysis of Integral Membrane Proteins from Eubacterial, Archaean, and Eukaryotic Organisms. *Protein Sci.* **1998**, *7* (4), 1029–1038. https://doi.org/10.1002/pro.5560070420.

(54) Konijnenberg, A.; Van Dyck, J. F.; Kailing, L. L.; Sobott, F. Extending Native Mass Spectrometry Approaches to Integral Membrane Proteins. *Biol. Chem.* **2015**, *396* (9–10), 991–1002. https://doi.org/10.1515/hsz-2015-0136.

(55) Sutto, L.; Marsili, S.; Valencia, A.; Gervasio, F. L. From Residue Coevolution to Protein Conformational Ensembles and Functional Dynamics. *Biophys. Comput. Biol. Downloaded Univ. MICHIGAN April* **2015**, *112* (44), 2020. https://doi.org/10.1073/pnas.1508584112.

(56) White, S. H. Biophysical Dissection of Membrane Proteins. *Nature*. Nature Publishing Group May 21, 2009, pp 344–346. https://doi.org/10.1038/nature08142.

(57) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. **2006**, *22* (5), 623–625. https://doi.org/10.1093/bioinformatics/btk023.

(58) Kozma, D.; Simon, I.; Tusnády, G. E. PDBTM: Protein Data Bank of Transmembrane Proteins after 8 Years. *Nucleic Acids Res.* **2013**, *41* (D1). https://doi.org/10.1093/nar/gks1169.

(59) Vit, O.; Petrak, J. Integral Membrane Proteins in Proteomics. How to Break Open the Black Box? *Journal of Proteomics*. Elsevier B.V. February 5, 2017, pp 8–20. https://doi.org/10.1016/j.jprot.2016.08.006.

(60) Brunner, J. D.; Schenck, S. Production and Application of Nanobodies for Membrane Protein Structural Biology. In *Expression, Purification, and Structural Biology of Membrane Proteins*; Perez, C., Maier, T., Eds.; Springer US: New York, NY, 2020; pp 167–184. https://doi.org/10.1007/978-1-0716-0373-4_12.

(61) Caffrey, M. A Comprehensive Review of the Lipid Cubic Phase or in Meso Method for Crystallizing Membrane and Soluble Proteins and Complexes. *Acta Cryst* **2015**, *71*, 3–18. https://doi.org/10.1107/S2053230X14026843.

(62) Rasmussen, S. G. F.; Devree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal Structure of the β 2 Adrenergic Receptor-Gs Protein Complex. *Nature* **2011**, *477* (7366), 549–557. https://doi.org/10.1038/nature10361.

(63) Asarnow, D.; Palovcak, E.; Gao, Y.; Julius, D.; Cheng, Y. Ion Channel in Lipid Nanodisc By Single Particle Cryo-EM - Pushing The Technology Limit. *Microsc. Microanal.* **2017**, *23* (S1), 822–823. https://doi.org/10.1017/s1431927617004779.

(64) Thonghin, N.; Kargas, V.; Clews, J.; Ford, R. C. Cryo-Electron Microscopy of Membrane Proteins. *Methods*. Academic Press Inc. September 1, 2018, pp 176–186. https://doi.org/10.1016/j.ymeth.2018.04.018.

(65) Cheng, Y. Membrane Protein Structural Biology in the Era of Single Particle Cryo-EM. *Current Opinion in Structural Biology*. Elsevier Ltd October 1, 2018, pp 58–63. https://doi.org/10.1016/j.sbi.2018.08.008.

(66) Liang, B.; Tamm, L. K. NMR as a Tool to Investigate the Structure, Dynamics and Function of Membrane Proteins. *Nature Structural and Molecular Biology*. Nature Publishing Group June 7, 2016, pp 468–474. https://doi.org/10.1038/nsmb.3226.
(67) Bibow, S. Exploring Lipid and Membrane Protein Dynamics Using Lipid-Bilayer Nanodiscs and Solution-State NMR Spectroscopy. In *Methods in Molecular Biology*; Humana Press Inc., 2020; Vol. 2127, pp 397–419. https://doi.org/10.1007/978-1-0716-0373-4_25.
(68) Asarnow, D.; Palovcak, E.; Gao, Y.; Julius, D.; Cheng, Y. Ion Channel in Lipid Nanodisc By Single Particle Cryo-EM-Pushing The Technology Limit. *Microsc. Microanal* **2020**, *23* (1), 2017. https://doi.org/10.1017/S1431927617004779.
(69) Frey, L.; Lakomek, N.-A.; Riek, R.; Bibow, S. Micelles, Bicelles, and Nanodiscs: Comparing the Impact of Membrane Mimetics on Membrane Protein Backbone Dynamics. *Angew. Chemie Int. Ed.* **2017**, *56* (1), 380–383. https://doi.org/10.1002/anie.201608246.
(70) Patching, S. G. Surface Plasmon Resonance Spectroscopy for Characterisation of Membrane Protein-Ligand Interactions and Its Potential for Drug Discovery. *Biochimica et Biophysica Acta - Biomembranes*. 2014, pp 43–55. https://doi.org/10.1016/j.bbamem.2013.04.028.
(71) Miles, A. J.; Wallace, B. A. Circular Dichroism Spectroscopy of Membrane Proteins. *Chem. Soc. Rev* **2016**, *45*, 4859. https://doi.org/10.1039/c5cs00084j.
(72) Krainer, G.; Keller, S.; Schlierf, M. Structural Dynamics of Membrane-Protein Folding from Single-Molecule FRET. *Current Opinion in Structural Biology*. Elsevier Ltd October 1, 2019, pp 124–137. https://doi.org/10.1016/j.sbi.2019.05.025.
(73) Rabuck, J. N.; Hyung, S. J.; Ko, K. S.; Fox, C. C.; Soellner, M. B.; Ruotolo, B. T. Activation State-Selective Kinase Inhibitor Assay Based on Ion Mobility-Mass Spectrometry. *Anal. Chem.* **2013**, *85* (15), 6995–7002. https://doi.org/10.1021/ac4012655.
(74) Reddy Bolla, J.; Agasid, M. T.; Mehmood, S.; Robinson, C. V; ARjatscls, R. Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. *Annu. Rev. Biochem.* **2019**, *88*, 85–111. https://doi.org/10.1146/annurev-biochem-013118.
(75) Sipe, S. N.; Patrick, J. W.; Laganowsky, A.; Brodbelt, J. S. Enhanced Characterization of Membrane Protein Complexes by Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2019**. https://doi.org/10.1021/acs.analchem.9b03689.
(76) Cong, X.; Liu, Y.; Liu, W.; Liang, X.; Russell, D. H.; Laganowsky, A. Determining Membrane Protein-Lipid Binding Thermodynamics Using Native Mass Spectrometry. *J. Am. Chem. Soc.* **2016**, *138* (13), 4346–4349. https://doi.org/10.1021/jacs.6b01771.
(77) Helbig, A. O.; Heck, A. J. R.; Slijper, M. Exploring the Membrane Proteome-Challenges and Analytical Strategies. *Journal of Proteomics*. Elsevier March 10, 2010, pp 868–878. https://doi.org/10.1016/j.jprot.2010.01.005.
(78) Borysik, A. J.; Hewitt, D. J.; Robinson, C. V. Detergent Release Prolongs the Lifetime of Native-like Membrane Protein Conformations in the Gas-Phase. *J. Am. Chem. Soc.* **2013**, *135* (16), 6078–6083. https://doi.org/10.1021/ja401736v.
(79) Souda, P.; Ryan, C. M.; Cramer, W. A.; Whitelegge, J. Profiling of Integral Membrane Proteins and Their Post Translational Modifications Using High-Resolution Mass Spectrometry. *Methods* **2011**, *55* (4), 330–336. https://doi.org/10.1016/j.ymeth.2011.09.019.
(80) Catherman, A. D.; Li, M.; Tran, J. C.; Durbin, K. R.; Compton, P. D.; Early, B. P.;

Thomas, P. M.; Kelleher, N. L. Top Down Proteomics of Human Membrane Proteins from Enriched Mitochondrial Fractions. **2013**. https://doi.org/10.1021/ac3031527.
(81) Moore, S. M.; Hess, S. M.; Jorgenson, J. W. Extraction, Enrichment, Solubilization, and Digestion Techniques for Membrane Proteomics. **2016**. https://doi.org/10.1021/acs.jproteome.5b01122.
(82) Wiśniewski, J. R. Filter-Aided Sample Preparation: The Versatile and Efficient Method for Proteomic Analysis. In *Methods in Enzymology*; Academic Press Inc., 2017; Vol. 585, pp 15–27. https://doi.org/10.1016/bs.mie.2016.09.013.
(83) Yu, Y.-Q.; Gilar, M.; Lee, P. J.; Bouvier, E. S. P.; Gebler, J. C. Enzyme-Friendly, Mass Spectrometry-Compatible Surfactant for In-Solution Enzymatic Digestion of Proteins. **2003**. https://doi.org/10.1021/ac0346196.
(84) Konijnenberg, A.; Bannwarth, L.; Yilmaz, D.; Koçer, A.; Venien-Bryan, C.; Sobott, F. Top-down Mass Spectrometry of Intact Membrane Protein Complexes Reveals Oligomeric State and Sequence Information in a Single Experiment. *Protein Sci.* **2015**, *24* (8), 1292–1300. https://doi.org/10.1002/pro.2703.
(85) Zhan, L. P.; Liu, C. Z.; Nie, Z. X. Mass Spectrometry of Membrane Proteins. In *Membrane Biophysics: New Insights and Methods*; Springer Singapore, 2017; pp 285–317. https://doi.org/10.1007/978-981-10-6823-2_10.
(86) Yan, Y.; Chen, G.; Wei, H.; Y-C Huang, R.; Mo, J.; Rempel, D. L.; Tymiak, A. A.; Gross, M. L. Fast Photochemical Oxidation of Proteins (FPOP) Maps the Epitope of EGFR Binding to Adnectin. *J. Am. Soc. Mass Spectrom* **2014**, *25*, 2084–2092. https://doi.org/10.1007/s13361-014-0993-x.
(87) Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E. FPOP-LC-MS/MS Suggests Differences in Interaction Sites of Amphipols and Detergents with Outer Membrane Proteins. *J. Am. Soc. Mass Spectrom 28*, 50–55. https://doi.org/10.1007/s13361-016-1421-1.
(88) Duncan, A. L.; Song, W.; Sansom, M. S. P. Lipid-Dependent Regulation of Ion Channels and G Protein–Coupled Receptors: Insights from Structures and Simulations. *Annu. Rev. Pharmacol. Toxicol.* **2020**, *60* (1), 31–50. https://doi.org/10.1146/annurev-pharmtox-010919-023411.
(89) Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. LIPID MAPS Online Tools for Lipid Research. *Nucleic Acids Res.* **2007**, *35* (SUPPL.2). https://doi.org/10.1093/nar/gkm324.
(90) Hu, C.; Duan, Q.; Han, X. Strategies to Improve/Eliminate the Limitations in Shotgun Lipidomics. *Proteomics* **2019**, 1900070. https://doi.org/10.1002/pmic.201900070.
(91) Kind, T.; Liu, K.-H.; Lee, D. Y.; Defelice, B.; Meissen, J. K.; Fiehn, O. LipidBlast-in-Silico Tandem Mass Spectrometry Database for Lipid Identification HHS Public Access. *Nat Methods* **2013**, *10* (8), 755–758. https://doi.org/10.1038/nmeth.2551.
(92) Hartler, J.; Triebl, A.; Ziegl, A.; Trötzmüller, M.; Rechberger, G. N.; Zeleznik, O. A.; Zierler, K. A.; Torta, F.; Cazenave-Gassiot, A.; Wenk, M. R.; Fauland, A.; Wheelock, C. E.; Armando, A. M.; Quehenberger, O.; Zhang, Q.; Wakelam, M. J. O.; Haemmerle, G.; Spener, F.; Köfeler, H. C.; Thallinger, G. G. Deciphering Lipid Structures Based on Platform-Independent Decision Rules. *Nat. Methods* **2017**, *14* (12), 1171–1174. https://doi.org/10.1038/nmeth.4470.
(93) Gupta, K.; Li, J.; Liko, I.; Gault, J.; Bechara, C.; Wu, D.; Hopper, J. T. S.; Giles, K.; Benesch, J. L. P.; Robinson, C. V. Identifying Key Membrane Protein Lipid Interactions Using Mass Spectrometry. *Nat. Protoc.* **2018**, *13* (5), 1106–1120.

https://doi.org/10.1038/nprot.2018.014.
(94) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane Proteins Bind Lipids Selectively to Modulate Their Structure and Function. *Nature* **2014**, *510* (7503), 172–175. https://doi.org/10.1038/nature13419.
(95) Laganowsky, A.; Reading, E.; Hopper, J. T. S.; Robinson, C. V. Mass Spectrometry of Intact Membrane Protein Complexes. *Nat. Protoc.* **2013**, *8* (4), 639–651. https://doi.org/10.1038/nprot.2013.024.
(96) Liko, I.; Degiacomi, M. T.; Lee, S.; Newport, T. D.; Gault, J.; Reading, E.; Hopper, J. T. S.; Housden, N. G.; White, P.; Colledge, M.; Sula, A.; Wallace, B. A.; Kleanthous, C.; Stansfeld, P. J.; Bayley, H.; Benesch, J. L. P.; Allison, T. M.; Robinson, C. V. Lipid Binding Attenuates Channel Closure of the Outer Membrane Protein OmpF. *Proc. Natl. Acad. Sci.* **2018**, *115* (26), 6691–6696. https://doi.org/10.1073/pnas.1721152115.
(97) Gault, J.; Donlan, J. A. C.; Liko, I.; Hopper, J. T. S.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.; Belov, M.; Damoc, E.; Makarov, A.; Robinson, C. V. High-Resolution Mass Spectrometry of Small Molecules Bound to Membrane Proteins. *Nat. Methods* **2016**, *13* (4), 333–336. https://doi.org/10.1038/nmeth.3771.
(98) Keener, J. E.; Zambrano, D. E.; Zhang, G.; Zak, C. K.; Reid, D. J.; Deodhar, B. S.; Pemberton, J. E.; Prell, J. S.; Marty, M. T. Chemical Additives Enable Native Mass Spectrometry Measurement of Membrane Protein Oligomeric State within Intact Nanodiscs. **2018**. https://doi.org/10.1021/jacs.8b11529.
(99) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V. Ion Mobility-Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3* (7), 1139–1152. https://doi.org/10.1038/nprot.2008.78.
(100) Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. Collision Induced Unfolding of Isolated Proteins in the Gas Phase: Past, Present, and Future. *Current Opinion in Chemical Biology*. Elsevier Ltd February 1, 2018, pp 93–100. https://doi.org/10.1016/j.cbpa.2017.11.010.
(101) Polasky, D. A.; Dixit, S. M.; Fantin, S. M.; Ruotolo, B. T. CIUSuite 2 : Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data CIUSuite 2 : Next-Generation Software for the Analysis of Gas-. *Anal. Chem.* **2019**, Just Accepted Manuscript. https://doi.org/10.1021/acs.analchem.8b05762.
(102) Liu, Y.; Cong, X.; Liu, W.; Laganowsky, A. Characterization of Membrane Protein–Lipid Interactions by Mass Spectrometry Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 579–586. https://doi.org/10.1007/s13361-016-1555-1.
(103) Borysik, A. J.; Hewitt, D. J.; Robinson, C. V. Detergent Release Prolongs the Lifetime of Native-like Membrane Protein Conformations in the Gas-Phase. *J. Am. Chem. Soc* **2013**, *135*, 37. https://doi.org/10.1021/ja401736v.
(104) Hopper, J. T. S.; Yu, Y. T. C.; Li, D.; Raymond, A.; Bostock, M.; Liko, I.; Mikhailov, V.; Laganowsky, A.; Benesch, J. L. P.; Caffrey, M.; Nietlispach, D.; Robinson, C. V. Detergent-Free Mass Spectrometry of Membrane Protein Complexes. *Nat. Methods* **2013**, *10* (12), 1206–1208. https://doi.org/10.1038/nmeth.2691.
(105) Zhou, M.; Morgner, N.; Barrera, N. P.; Politis, A.; Isaacson, S. C.; Matak-Vinkovic, D.; Murata, T.; Bernal, R. A.; Stock, D.; Robinson, C. V. Mass Spectrometry of Intact V-Type ATPases Reveals Bound Lipids and the Effects of Nucleotide Binding. *Science (80-.).*

2011, *334* (6054), 380–385. https://doi.org/10.1126/science.1210148.
(106) Landreh, M.; Marklund, E. G.; Uzdavinys, P.; Degiacomi, M. T.; Coincon, M.; Gault, J.; Gupta, K.; Liko, I.; Benesch, J. L. P.; Drew, D.; Robinson, C. V. Integrating Mass Spectrometry with MD Simulations Reveals the Role of Lipids in Na+/H+ Antiporters. *Nat. Commun.* **2017**, *8*, 1–9. https://doi.org/10.1038/ncomms13993.
(107) Almén, M. S.; Nordström, K. J. V.; Fredriksson, R.; Schiöth, H. B. Mapping the Human Membrane Proteome: A Majority of the Human Membrane Proteins Can Be Classified According to Function and Evolutionary Origin. *BMC Biol.* **2009**, *7*, 50. https://doi.org/10.1186/1741-7007-7-50.
(108) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How Many Drug Targets Are There? *Nat. Rev. Drug Discov.* **2006**, *5* (12), 993–996. https://doi.org/10.1038/nrd2199.
(109) Arinaminpathy, Y.; Khurana, E.; Engelman, D. M.; Gerstein, M. B. Computational Analysis of Membrane Proteins: The Largest Class of Drug Targets. *Drug Discov. Today* **2009**, *14* (23–24), 1130–1135. https://doi.org/10.1016/j.drudis.2009.08.006.
(110) Speers, A. E.; Wu, C. C. Proteomics of Integral Membrane Proteins - Theory and Application. *Chem. Rev.* **2007**, *107* (8), 3687–3714. https://doi.org/10.1021/cr068286z.
(111) Raschle, T.; Hiller, S.; Yu, T. Y.; Rice, A. J.; Walz, T.; Wagner, G. Structural and Functional Characterization of the Integral Membrane Protein VDAC-1 in Lipid Bilayer Nanodiscs. *J. Am. Chem. Soc.* **2009**, *131* (49), 17777–17779. https://doi.org/10.1021/ja907918r.
(112) Bill, R. M.; Henderson, P. J. F.; Iwata, S.; Kunji, E. R. S.; Michel, H.; Neutze, R.; Newstead, S.; Poolman, B.; Tate, C. G.; Vogel, H. Overcoming Barriers to Membrane Protein Structure Determination. *Nat. Biotechnol.* **2011**, *29* (4), 335–340. https://doi.org/10.1038/nbt.1833.
(113) Srivastava, A. P.; Luo, M.; Zhou, W.; Symersky, J.; Bai, D.; Chambers, M. G.; Faraldo-Gómez, J. D.; Liao, M.; Mueller, D. M. High-Resolution Cryo-EM Analysis of the Yeast ATP Synthase in a Lipid Membrane. *Science (80-.).* **2018**, *360* (6389). https://doi.org/10.1126/science.aas9699.
(114) Gonen, T.; Cheng, Y.; Sliz, P.; Hiroaki, Y.; Fujiyoshi, Y.; Harrison, S. C.; Walz, T. Erratum: Lipid-Protein Interactions in Double-Layered Two-Dimensional AQPO Crystals (Nature (2005) 438 (633-638)). *Nature* **2006**, *441* (7090), 248. https://doi.org/10.1038/nature04775.
(115) Kuhlen, L.; Abrusci, P.; Johnson, S.; Gault, J.; Deme, J.; Caesar, J.; Dietsche, T.; Mebrhatu, M. T.; Ganief, T.; Macek, B.; Wagner, S.; Robinson, C. V.; Lea, S. M. Structure of the Core of the Type III Secretion System Export Apparatus. *Nat. Struct. Mol. Biol.* **2018**, 1–8. https://doi.org/10.1038/s41594-018-0086-9.
(116) Harvey, S. R.; Liu, Y.; Liu, W.; Wysocki, V. H.; Laganowsky, A. Surface Induced Dissociation as a Tool to Study Membrane Protein Complexes. *Chem. Commun.* **2017**, *53* (21), 3106–3109. https://doi.org/10.1039/c6cc09606a.
(117) Landreh, M.; Marty, M. T.; Gault, J.; Robinson, C. V. A Sliding Selectivity Scale for Lipid Binding to Membrane Proteins. *Curr. Opin. Struct. Biol.* **2016**, *39*, 54–60. https://doi.org/10.1016/j.sbi.2016.04.005.
(118) Marty, M. T.; Hoi, K. K.; Gault, J.; Robinson, C. V. Probing the Lipid Annular Belt by Gas-Phase Dissociation of Membrane Proteins in Nanodiscs. *Angew. Chemie - Int. Ed.* **2016**, *55* (2), 550–554. https://doi.org/10.1002/anie.201508289.
(119) Calabrese, A. N.; Watkinson, T. G.; Henderson, P. J. F.; Radford, S. E.; Ashcroft, A. E.

Amphipols Outperform Dodecylmaltoside Micelles in Stabilizing Membrane Protein Structure in the Gas Phase. *Anal. Chem.* **2015**, *87* (2), 1118–1126. https://doi.org/10.1021/ac5037022.
(120) Marty, M. T.; Hoi, K. K.; Robinson, C. V. Interfacing Membrane Mimetics with Mass Spectrometry. **2016**. https://doi.org/10.1021/acs.accounts.6b00379.
(121) Marty, M. T.; Zhang, H.; Cui, W.; Blankenship, R. E.; Gross, M. L.; Sligar, S. G. Native Mass Spectrometry Characterization of Intact Nanodisc Lipoprotein Complexes. *Anal. Chem.* **2012**, *84* (21), 8957–8960. https://doi.org/10.1021/ac302663f.
(122) Konijnenberg, A.; Yilmaz, D.; Ingólfsson, H. I.; Dimitrova, A.; Marrink, S. J.; Li, Z.; Vénien-Bryan, C.; Sobott, F.; Koçer, A. Global Structural Changes of an Ion Channel during Its Gating Are Followed by Ion Mobility Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2014**, *111* (48), 17170–17175. https://doi.org/10.1073/pnas.1413118111.
(123) Liu, Y.; Cong, X.; Liu, W.; Laganowsky, A. Characterization of Membrane Protein–Lipid Interactions by Mass Spectrometry Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 579–586. https://doi.org/10.1007/s13361-016-1555-1.
(124) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane Proteins Bind Lipids Selectively to Modulate Their Structure and Function. *Nature* **2014**, *510* (7503), 172–175. https://doi.org/10.1038/nature13419.
(125) Girard, C.; Liu, S.; Adams, D.; Lacroix, C.; Sinéus, M.; Boucher, C.; Papadopoulos, V.; Rupprecht, R.; Schumacher, M.; Groyer, G. Axonal Regeneration and Neuroinflammation: Roles for the Translocator Protein 18kDa. *J. Neuroendocrinol.* **2012**, *24* (1), 71–81. https://doi.org/10.1111/j.1365-2826.2011.02215.x.
(126) Papadopoulos, V.1. Papadopoulos, V., Fan, J. & Zirkin, B. Translocator protein (18 kDa): an update on its function in steroidogenesis. J. Neuroendocrinol. 30, 1–9 (2018).; Fan, J.; Zirkin, B. Translocator Protein (18 KDa): An Update on Its Function in Steroidogenesis. *J. Neuroendocrinol.* **2018**, *30* (2), 1–9. https://doi.org/10.1111/jne.12500.
(127) Li, F.; Liu, J.; Valls, L.; Hiser, C.; Ferguson-Miller, S. Identification of a Key Cholesterol Binding Enhancement Motif in Translocator Protein 18 KDa. *Biochemistry* **2015**, *54* (7), 1441–1443. https://doi.org/10.1021/bi5015453.
(128) Li, F.; Xia, Y.; Meiler, J.; Ferguson-Miller, S. Characterization and Modeling of the Oligomeric State and Ligand Binding Behavior of Purified Translocator Protein 18 KDa from Rhodobacter Sphaeroides. *Biochemistry* **2013**, *52* (34), 5884–5899. https://doi.org/10.1021/bi400431t.
(129) Li, F.; Liu, J.; Liu, N.; Kuhn, L. A.; Garavito, R. M.; Ferguson-Miller, S. Translocator Protein 18 KDa (TSPO): An Old Protein with New Functions? *Biochemistry* **2016**, *55* (20), 2821–2831. https://doi.org/10.1021/acs.biochem.6b00142.
(130) Allison, T. M.; Landreh, M.; Benesch, J. L. P.; Robinson, C. V. Low Charge and Reduced Mobility of Membrane Protein Complexes Has Implications for Calibration of Collision Cross Section Measurements. *Anal. Chem.* **2016**, *88* (11), 5879–5884. https://doi.org/10.1021/acs.analchem.6b00691.
(131) Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. P. Collision Cross Sections for Structural Proteomics. *Structure* **2015**, *23* (4), 791–799. https://doi.org/10.1016/J.STR.2015.02.010.
(132) Haynes, S. E.; Polasky, D. A.; Dixit, S. M.; Majmudar, J. D.; Neeson, K.; Ruotolo, B. T.; Martin, B. R. Variable-Velocity Traveling-Wave Ion Mobility Separation Enhancing Peak

Capacity for Data-Independent Acquisition Proteomics. *Anal. Chem.* **2017**, *89* (11), 5669–5672. https://doi.org/10.1021/acs.analchem.7b00112.
(133) Polasky, D. A.; Dixit, S. M.; Fantin, S. M.; Ruotolo, B. T. CIUSuite 2: Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data. *Anal. Chem.* **2019**, *91* (4), 3147–3155. https://doi.org/10.1021/acs.analchem.8b05762.
(134) Vallejo, D. D.; Polasky, D. A.; Kurulugama, R. T.; Eschweiler, J. D.; Fjeldsted, J. C.; Ruotolo, B. T. A Modified Drift Tube Ion Mobility-Mass Spectrometer for Charge-Multiplexed Collision-Induced Unfolding. **2019**. https://doi.org/10.1021/acs.analchem.9b00427.
(135) Mammone, A.; Turchi, M.; Cristianini, N. Support Vector Machines. *Wiley Interdiscip. Rev. Comput. Stat.* **2009**, *1* (3), 283–289. https://doi.org/10.1002/wics.49.
(136) Allison, T. M.; Reading, E.; Liko, I.; Baldwin, A. J.; Laganowsky, A.; Robinson, C. V. Quantifying the Stabilizing Effects of Protein-Ligand Interactions in the Gas Phase. *Nat. Commun.* **2015**, *6*, 1–10. https://doi.org/10.1038/ncomms9551.
(137) Landreh, M.; Costeira-Paulo, J.; Gault, J.; Marklund, E. G.; Robinson, C. V. Effects of Detergent Micelles on Lipid Binding to Proteins in Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2017**, *89* (14), 7425–7430. https://doi.org/10.1021/acs.analchem.7b00922.
(138) Borysik, A. J.; Robinson, C. V. The "sticky Business" of Cleaning Gas-Phase Membrane Proteins: A Detergent Oriented Perspective. *Phys. Chem. Chem. Phys.* **2012**, *14* (42), 14439–14449. https://doi.org/10.1039/c2cp41687e.
(139) Christopoulos, A. Allosteric Binding Sites on Cell-Surface Receptors: Novel Targets for Drug Discovery. *Nat. Rev. Drug Discov.* **2002**, *1* (3), 198–210. https://doi.org/10.1038/nrd746.
(140) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat. Rev. Drug Discov.* **2017**, *16* (12), 829–842. https://doi.org/10.1038/nrd.2017.178.
(141) Phillips, R.; Ursell, T.; Wiggins, P.; Sens, P. Emerging Roles for Lipids in Shaping Membrane-Protein Function. *Nature* **2009**, *459* (7245), 379–385. https://doi.org/10.1038/nature08147.
(142) Ng, D. P.; Poulsen, B. E.; Deber, C. M. Membrane Protein Misassembly in Disease. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818* (4), 1115–1122. https://doi.org/10.1016/j.bbamem.2011.07.046.
(143) Farinha, C. M.; Canato, S. From the Endoplasmic Reticulum to the Plasma Membrane: Mechanisms of CFTR Folding and Trafficking. *Cellular and Molecular Life Sciences*. Birkhauser Verlag AG January 1, 2017, pp 39–55. https://doi.org/10.1007/s00018-016-2387-7.
(144) Moeller, H. B.; Rittig, S.; Fenton, R. A. Nephrogenic Diabetes Insipidus: Essential Insights into the Molecular Background and Potential Therapies for Treatment. *Endocrine Reviews*. Oxford Academic April 1, 2013, pp 278–301. https://doi.org/10.1210/er.2012-1044.
(145) Marinko, J. T.; Huang, H.; Penn, W. D.; Capra, J. A.; Schlebach, J. P.; Sanders, C. R. Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis. **2019**. https://doi.org/10.1021/acs.chemrev.8b00532.
(146) Huang, H.; Kuenze, G.; Smith, J. A.; Taylor, K. C.; Duran, A. M.; Hadziselimovic, A.;

Meiler, J.; Vanoye, C. G.; George, A. L.; Sanders, C. R. Mechanisms of KCNQ1 Channel Dysfunction in Long QT Syndrome Involving Voltage Sensor Domain Mutations. *Sci. Adv.* **2018**, *4* (3), eaar2631. https://doi.org/10.1126/sciadv.aar2631.

(147) Li, J.; Parker, B.; Martyn, C.; Natarajan, C.; Guo, J. The PMP22 Gene and Its Related Diseases. *Molecular neurobiology*. 2013, pp 673–698. https://doi.org/10.1007/s12035-012-8370-x.

(148) Pantera, H.; Shy, M. E.; Svaren, J. Regulating PMP22 Expression as a Dosage Sensitive Neuropathy Gene. *Brain Research*. Elsevier B.V. January 1, 2020, p 146491. https://doi.org/10.1016/j.brainres.2019.146491.

(149) Jetten, A. M.; Suteri, U. The Peripheral Myelin Protein 22 and Epithelial Membrane Protein Family. *Progress in Nucleic Acid Research and Molecular Biology*. January 1, 2000, pp 97–129. https://doi.org/10.1016/s0079-6603(00)64003-5.

(150) Vanoye, C. G.; Sakakura, M.; Follis, R. M.; Trevisan, A. J.; Narayan, M.; Li, J.; Sanders, C. R.; Carter, B. D. Peripheral Myelin Protein 22 Modulates Store-Operated Calcium Channel Activity, Providing Insights into Charcot-Marie-Tooth Disease Etiology. *J. Biol. Chem.* **2019**, *294* (32), 12054–12065. https://doi.org/10.1074/jbc.RA118.006248.

(151) Fabbretti, E.; Edomi, P.; Brancolini, C.; Schneider, C. Apoptotic Phenotype Induced by Overexpression of Wild-Type Gas3/PMP22: Its Relation to the Demyelinating Peripheral Neuropathy CMT1A. *Genes Dev.* **1995**, *9* (15), 1846–1856. https://doi.org/10.1101/gad.9.15.1846.

(152) Lee, S.; Amici, S.; Tavori, H.; Zeng, W. M.; Freeland, S.; Fazio, S.; Notterpek, L. PMP22 Is Critical for Actin-Mediated Cellular Functions and for Establishing Lipid Rafts. *J. Neurosci.* **2014**, *34* (48), 16140–16152. https://doi.org/10.1523/JNEUROSCI.1908-14.2014.

(153) Roux, K. J.; Amici, S. A.; Fletcher, B. S.; Notterpek, L. Modulation of Epithelial Morphology, Monolayer Permeability, and Cell Migration by Growth Arrest Specific 3/Peripheral Myelin Protein 22. *Mol. Biol. Cell* **2005**, *16* (3), 1142–1151. https://doi.org/10.1091/mbc.E04-07-0551.

(154) D'Urso, D.; Ehrhardt, P.; Müller, H. W. Peripheral Myelin Protein 22 and Protein Zero: A Novel Association in Peripheral Nervous System Myelin. *J. Neurosci.* **1999**, *19* (9), 3396–3403. https://doi.org/10.1523/JNEUROSCI.19-09-03396.1999.

(155) Zhou, Y.; Miles, J. R.; Tavori, H.; Lin, M.; Khoshbouei, H.; Borchelt, D. R.; Bazick, H.; Landreth, G. E.; Lee, S.; Fazio, S.; Notterpek, L. PMP22 Regulates Cholesterol Trafficking and ABCA1-Mediated Cholesterol Efflux. *J. Neurosci.* **2019**, *39* (27), 5404–5418. https://doi.org/10.1523/JNEUROSCI.2942-18.2019.

(156) Van Paassen, B. W.; Van Der Kooi, A. J.; Van Spaendonck-Zwarts, K. Y.; Verhamme, C.; Baas, F.; De Visser, M. PMP22 Related Neuropathies: Charcot-Marie-Tooth Disease Type 1A and Hereditary Neuropathy with Liability to Pressure Palsies. *Orphanet Journal of Rare Diseases*. BioMed Central Ltd. March 19, 2014, pp 1–15. https://doi.org/10.1186/1750-1172-9-38.

(157) Sanders, C. R.; Myers, J. K. Disease-Related Misassembly of Membrane Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 25–51. https://doi.org/10.1146/annurev.biophys.33.110502.140348.

(158) Pareek, S.; Notterpek, L.; Snipes, G. J.; Naef, R.; Sossin, W.; Laliberté, J.; Iacampo, S.; Suter, U.; Shooter, E. M.; Murphy, R. A. *Neurons Promote the Translocation of Peripheral Myelin Protein 22 into Myelin*; 1997.

(159) Barrera, N. P.; Robinson, C. V. Advances in the Mass Spectrometry of Membrane Proteins: From Individual Proteins to Intact Complexes. *Annu. Rev. Biochem.* **2011**, *80* (1), 247–271. https://doi.org/10.1146/annurev-biochem-062309-093307.
(160) Marty, M. T.; Hoi, K. K.; Robinson, C. V. Interfacing Membrane Mimetics with Mass Spectrometry. *Acc. Chem. Res.* **2016**, *49* (11), 2459–2467. https://doi.org/10.1021/acs.accounts.6b00379.
(161) Chorev, D. S.; Tang, H.; Rouse, S. L.; Bolla, J. R.; von Kügelgen, A.; Baker, L. A.; Wu, D.; Gault, J.; Grünewald, K.; Bharat, T. A. M.; Matthews, S. J.; Robinson, C. V. The Use of Sonicated Lipid Vesicles for Mass Spectrometry of Membrane Protein Complexes. *Nat. Protoc.* **2020**, 1–17. https://doi.org/10.1038/s41596-020-0303-y.
(162) Keener, J. E.; Zambrano, D. E.; Zhang, G.; Zak, C. K.; Reid, D. J.; Deodhar, B. S.; Pemberton, J. E.; Prell, J. S.; Marty, M. T. Chemical Additives Enable Native Mass Spectrometry Measurement of Membrane Protein Oligomeric State within Intact Nanodiscs. *J. Am. Chem. Soc.* **2019**, *141* (2), 1054–1061. https://doi.org/10.1021/jacs.8b11529.
(163) Lippens, J. L.; Nshanian, M.; Spahr, C.; Egea, P. F.; Loo, J. A.; Campuzano, I. D. G. Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry as a Platform for Characterizing Multimeric Membrane Protein Complexes. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (1), 183–193. https://doi.org/10.1007/s13361-017-1799-4.
(164) Rabuck-Gibbons, J. N.; Keating, J. E.; Ruotolo, B. T. Collision Induced Unfolding and Dissociation Differentiates ATP-Competitive from Allosteric Protein Tyrosine Kinase Inhibitors. *Int. J. Mass Spectrom.* **2018**, *427*, 151–156. https://doi.org/10.1016/j.ijms.2017.12.002.
(165) Hopper, J. T. S.; Oldham, N. J. Collision Induced Unfolding of Protein Ions in the Gas Phase Studied by Ion Mobility-Mass Spectrometry: The Effect of Ligand Binding on Conformational Stability. *J Am Soc Mass Spectrom* **2009**, *20*, 1851–1858. https://doi.org/10.1016/j.jasms.2009.06.010.
(166) Tian, Y.; Han, L.; Buckner, A. C.; Ruotolo, B. T. Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures. *Anal. Chem.* **2015**, *87* (22), 11509–11515. https://doi.org/10.1021/acs.analchem.5b03291.
(167) Hernandez-Alba, O.; Wagner-Rousset, E.; Beck, A.; Cianférani, S. C. Native Mass Spectrometry, Ion Mobility, and Collision-Induced Unfolding for Conformational Characterization of IgG4 Monoclonal Antibodies. **2018**. https://doi.org/10.1021/acs.analchem.8b00912.
(168) Fantin, S. M.; Parson, K. F.; Niu, S.; Liu, J.; Polasky, D. A.; Dixit, S. M.; Ferguson-Miller, S. M.; Ruotolo, B. T. Collision Induced Unfolding Classifies Ligands Bound to the Integral Membrane Translocator Protein. *Anal. Chem* **2019**, *91*, 44. https://doi.org/10.1021/acs.analchem.9b03208.
(169) Schlebach, J. P.; Peng, D.; Kroncke, B. M.; Mittendorf, K. F.; Narayan, M.; Carter, B. D.; Sanders, C. R. Reversible Folding of Human Peripheral Myelin Protein 22, a Tetraspan Membrane Protein. *Biochemistry* **2013**, *52* (19), 3229–3241. https://doi.org/10.1021/bi301635f.
(170) Mittendorf, K. F.; Kroncke, B. M.; Meiler, J.; Sanders, C. R. The Homology Model of PMP22 Suggests Mutations Resulting in Peripheral Neuropathy Disrupt Transmembrane Helix Packing. *Biochemistry* **2014**, *53* (39), 6139–6141.

https://doi.org/10.1021/bi500809t.
(171) Tobler, A. R.; Notterpek, L.; Naef, R.; Taylor, V.; Suter, U.; Shooter, E. M. Transport of Trembler-J Mutant Peripheral Myelin Protein 22 Is Blocked in the Intermediate Compartment and Affects the Transport of the Wild-Type Protein by Direct Interaction. *J. Neurosci.* **1999**, *19* (6), 2027–2036. https://doi.org/10.1523/jneurosci.19-06-02027.1999.
(172) Fortun, J.; Go, J. C.; Li, J.; Amici, S. A.; Dunn, W. A.; Notterpek, L. Alterations in Degradative Pathways and Protein Aggregation in a Neuropathy Model Based on PMP22 Overexpression. *Neurobiol. Dis.* **2006**, *22* (1), 153–164. https://doi.org/10.1016/j.nbd.2005.10.010.
(173) Liu, N.; Yamauchi, J.; Shooter, E. M. Recessive, but Not Dominant, Mutations in Peripheral Myelin Protein 22 Gene Show Unique Patterns of Aggregation and Intracellular Trafficking. *Neurobiol. Dis.* **2004**, *17* (2), 300–309. https://doi.org/10.1016/j.nbd.2004.07.010.
(174) Tobler, A.; Liu, N.; Mueller, L.; Shooter, E. Differential Aggregation Of The Trembler And Trembler J Mutants Of Peripheral Myelin Protein 22. *J. Peripher. Nerv. Syst.* **2002**, *7* (3), 206–207. https://doi.org/10.1046/j.1529-8027.2002.02026_7.x.
(175) Nodera, H.; Nishimura, M.; Logigian, E. L.; Herrmann, D. N.; Kaji, R. HNPP Due to a Novel Missense Mutation of the PMP22 Gene. *Neurology* **2003**, *60* (11), 1863–1864. https://doi.org/10.1212/01.WNL.0000066049.13848.F2.
(176) Kumar, C. V.; Swetha, R. G.; Anbarasu, A.; Ramaiah, S. Computational Analysis Reveals the Association of Threonine 118 Methionine Mutation in PMP22 Resulting in CMT-1A. *Adv. Bioinformatics* **2014**, *2014*. https://doi.org/10.1155/2014/502618.
(177) Bello, M.; Torres, M. J.; Méndez-Tenorio, A.; Correa-Basurto, J. Conformational Changes Associated with L16P and T118M Mutations in the Membrane-Embedded PMP22 Protein, Consequential in CMT-1A. *J. Biomol. Struct. Dyn.* **2017**, *35* (13), 2880–2894. https://doi.org/10.1080/07391102.2016.1234415.
(178) Mittendorf, K. F.; Kroncke, B. M.; Meiler, J.; Sanders, C. R. The Homology Model of PMP22 Suggests Mutations Resulting in Peripheral Neuropathy Disrupt Transmembrane Helix Packing. *Biochemistry* **2014**, *53* (39), 6139–6141. https://doi.org/10.1021/bi500809t.
(179) Sakakura, M.; Hadziselimovic, A.; Wang, Z.; Schey, K. L.; Sanders, C. R. Structural Basis for the Trembler-J Phenotype of Charcot-Marie-Tooth Disease. *Structure* **2011**, *19* (8), 1160–1169. https://doi.org/10.1016/j.str.2011.05.009.
(180) Sanders, C. R.; Ismail-Beigi, F.; McEnery, M. W. Mutations of Peripheral Myelin Protein 22 Result in Defective Trafficking through Mechanisms Which May Be Common to Diseases Involving Tetraspan Membrane Proteins. *Biochemistry*. American Chemical Society August 14, 2001, pp 9453–9459. https://doi.org/10.1021/bi010894f.
(181) Duarte, J. M.; Biyani, N.; Baskaran, K.; Capitani, G. An Analysis of Oligomerization Interfaces in Transmembrane Proteins. *BMC Struct. Biol.* **2013**, *13* (1). https://doi.org/10.1186/1472-6807-13-21.
(182) Wu, H.; Wang, C.; Gregory, K. J.; Han, G. W.; Cho, H. P.; Xia, Y.; Niswender, C. M.; Katritch, V.; Meiler, J.; Cherezov, V.; Conn, P. J.; Stevens, R. C. Structure of a Class C GPCR Metabotropic Glutamate Receptor 1 Bound to an Allosteric Modulator #. *Science (80-.).* **2014**, *344* (6179), 58–64. https://doi.org/10.1126/science.1249489.
(183) Arkhipov, A.; Shan, Y.; Das, R.; Endres, N. F.; Eastwood, M. P.; Wemmer, D. E.; Kuriyan, J.; Shaw, D. E. Architecture and Membrane Interactions of the EGF Receptor.

Cell **2013**, *152* (3), 557–569. https://doi.org/10.1016/j.cell.2012.12.030.
(184) Schlebach, J. P.; Barrett, P. J.; Day, C. A.; Kim, J. H.; Kenworthy, A. K.; Sanders, C. R. Topologically Diverse Human Membrane Proteins Partition to Liquid-Disordered Domains in Phase-Separated Lipid Vesicles. *Biochemistry* **2016**, *55* (7), 985–988. https://doi.org/10.1021/acs.biochem.5b01154.
(185) Lorent, J. H.; Levental, I. Structural Determinants of Protein Partitioning into Ordered Membrane Domains and Lipid Rafts. *Chem. Phys. Lipids* **2015**, *192*, 23–32. https://doi.org/10.1016/j.chemphyslip.2015.07.022.
(186) Gopalakrishnan, G.; Awasthi, A.; Belkaid, W.; De Faria, O.; Liazoghli, D.; Colman, D. R.; Dhaunchak, A. S. Lipidome and Proteome Map of Myelin Membranes. *J. Neurosci. Res.* **2013**, *91* (3), 321–334. https://doi.org/10.1002/jnr.23157.
(187) Marinko, J. T.; Li, G. C.; Kenworthy, A. K.; Sanders, C. R. Peripheral Myelin Protein 22 Preferentially Partitions into Ordered Phase Membrane Domains. https://doi.org/10.1101/2020.01.28.923771.
(188) Hutchison, J. M.; Kuo-Chih Shih, Kuo-Chih Scheidt, H.; Fantin, S. M.; Pantelopulos, G. A.; Harrington, H. R.; Mittendorf, K. F.; Qian, S.; Stien, R. A.; Collier, Scott E. Melissa G. Chambers, Melissa G. Huang, Linjun Ritesh Mittal, R.; Katsaras, J. 8; Voehler, Markus W. Ruotolo, Brandon T. Huster, D.; McFeeters, Robert L. Straub, J. E.; Nieh, M.-P.; Sanders, C. R. Bicelles Rich in Both Sphingolipids and Cholesterol. *J. Am. Chem. Soc.* **2020**, *TBD*.
(189) Fortun, J.; Verrier, J. D.; Go, J. C.; Madorsky, I.; Dunn, W. A.; Notterpek, L. The Formation of Peripheral Myelin Protein 22 Aggregates Is Hindered by the Enhancement of Autophagy and Expression of Cytoplasmic Chaperones. *Neurobiol. Dis.* **2007**, *25* (2), 252–265. https://doi.org/10.1016/j.nbd.2006.09.018.
(190) Escribá, P. V.; González-Ros, J. M.; Goñi, F. M.; Kinnunen, P. K. J.; Vigh, L.; Sánchez-Magraner, L.; Fernández, A. M.; Busquets, X.; Horváth, I.; Barceló-Coblijn, G. Membranes: A Meeting Point for Lipids, Proteins and Therapies: Translational Medicine. *J. Cell. Mol. Med.* **2008**, *12* (3), 829–875. https://doi.org/10.1111/j.1582-4934.2008.00281.x.
(191) Ackerman, M. J. Genetic Purgatory and the Cardiac Channelopathies: Exposing the Variants of Uncertain/Unknown Significance Issue. *Hear. Rhythm* **2015**, *12* (11), 2325–2331. https://doi.org/10.1016/j.hrthm.2015.07.002.
(192) Oulas, A.; Minadakis, G.; Zachariou, M.; Spyrou, G. M. Selecting Variants of Unknown Significance through Network-Based Gene-Association Significantly Improves Risk Prediction for Disease-Control Cohorts. *Sci. Rep.* **2019**, *9* (1), 1–15. https://doi.org/10.1038/s41598-019-39796-w.
(193) Zhang, T.; Moss, A.; Cong, P.; Pan, M.; Chang, B.; Zheng, L.; Fang, Q.; Zareba, W.; Robinson, J.; Lin, C.; Li, Z.; Wei, J.; Zeng, Q.; Qi, M. LQTS Gene LOVD Database. *Hum. Mutat.* **2010**, *31* (11), E1801–E1810. https://doi.org/10.1002/humu.21341.
(194) Gault, J.; Liko, I.; Landreh, M.; Shutin, D.; Bolla, J. R.; Jefferies, D.; Agasid, M.; Yen, H.-Y.; Ladds, M. J. G. W.; Lane, D. P.; Khalid, S.; Mullen, C.; Remes, P.; Huguet, R.; McAlister, G.; Goodwin, M.; Viner, R.; Syka, J.; Robinson, C. V. Combining 'Native' with 'Omics' Based Mass Spectrometry to Identify Endogenous Ligands Bound to Membrane Proteins. *Nat. Methods* **2020**, *17*, 505–508. https://doi.org/10.1038/s41592-020-0821-0.
(195) Polasky, D. A.; Dixit, S. M.; Vallejo, D. D.; Kulju, K. D.; Ruotolo, B. T. An Algorithm

for Building Multi-State Classifiers Based on Collision-Induced Unfolding Data. *Anal. Chem.* **2019**, *91* (16), 10407–10412. https://doi.org/10.1021/acs.analchem.9b02650.
(196) Rabuck-Gibbons, J. N.; Keating, J. E.; Ruotolo, B. T. Collision Induced Unfolding and Dissociation Differentiates ATP-Competitive from Allosteric Protein Tyrosine Kinase Inhibitors. *Int. J. Mass Spectrom.* **2018**, *427*, 151–156. https://doi.org/10.1016/j.ijms.2017.12.002.
(197) Abbott, G. W. KCNE1 and KCNE3: The Yin and Yang of Voltage-Gated K+ Channel Regulation. *Gene*. Elsevier B.V. January 15, 2016, pp 1–13. https://doi.org/10.1016/j.gene.2015.09.059.
(198) Cui, J. Voltage-Dependent Gating: Novel Insights from KCNQ1 Channels. *Biophysical Journal*. Biophysical Society January 5, 2016, pp 14–25. https://doi.org/10.1016/j.bpj.2015.11.023.
(199) Wu, J.; Ding, W. G.; Horie, M. Molecular Pathogenesis of Long QT Syndrome Type 1. *Journal of Arrhythmia*. Elsevier B.V. October 1, 2016, pp 381–388. https://doi.org/10.1016/j.joa.2015.12.006.
(200) Wu, W.; Sanguinetti, M. C. Molecular Basis of Cardiac Delayed Rectifier Potassium Channel Function and Pharmacology. *Cardiac Electrophysiology Clinics*. W.B. Saunders June 1, 2016, pp 275–284. https://doi.org/10.1016/j.ccep.2016.01.002.
(201) Campuzano, O.; Fernandez-Falgueras, A.; Lemus, X.; Sarquella-Brugada, G.; Cesar, S.; Coll, M.; Mates, J.; Arbelo, E.; Jordà, P.; Perez-Serra, A.; del Olmo, B.; Ferrer-Costa, C.; Iglesias, A.; Fiol, V.; Puigmulé, M.; Lopez, L.; Pico, F.; Brugada, J.; Brugada, R. Short QT Syndrome: A Comprehensive Genetic Interpretation and Clinical Translation of Rare Variants. *J. Clin. Med.* **2019**, *8* (7), 1035. https://doi.org/10.3390/jcm8071035.
(202) Zhang, H.; Kharche, S.; Holden, A. V.; Hancox, J. C. Repolarisation and Vulnerability to Re-Entry in the Human Heart with Short QT Syndrome Arising from KCNQ1 Mutation-A Simulation Study. *Progress in Biophysics and Molecular Biology*. Pergamon January 1, 2008, pp 112–131. https://doi.org/10.1016/j.pbiomolbio.2007.07.020.
(203) Schwartz, P. J.; Crotti, L.; Insolia, R. Long-QT Syndrome. *Circ. Arrhythmia Electrophysiol.* **2012**, *5* (4), 868–877. https://doi.org/10.1161/CIRCEP.111.962019.
(204) Gussak, I.; Brugada, P.; Brugada, J.; Wright, R. S.; Kopecky, S. L.; Chaitman, B. R.; Bjerregaard, P. Idiopathic Short QT Interval:A New Clinical Syndrome? *Cardiology* **2000**, *94* (2), 99–102. https://doi.org/10.1159/000047299.
(205) Henrion, U.; Zumhagen, S.; Steinke, K.; Strutz-Seebohm, N.; Stallmeyer, B.; Lang, F.; Schulze-Bahr, E.; Seebohm, G. Overlapping Cardiac Phenotype Associated with a Familial Mutation in the Voltage Sensor of the KCNQ1 Channel. *Cell. Physiol. Biochem.* **2012**, *29* (5–6), 809–818. https://doi.org/10.1159/000178470.
(206) Bartos, D. C.; Duchatelet, S.; Burgess, D. E.; Klug, D.; Denjoy, I.; Peat, R.; Lupoglazoff, J. M.; Fressart, V.; Berthet, M.; Ackerman, M. J.; January, C. T.; Guicheney, P.; Delisle, B. P. R231C Mutation in KCNQ1 Causes Long QT Syndrome Type 1 and Familial Atrial Fibrillation. *Hear. Rhythm* **2011**, *8* (1), 48–55. https://doi.org/10.1016/j.hrthm.2010.09.010.
(207) Tester, D. J.; Will, M. L.; Haglund, C. M.; Ackerman, M. J. Compendium of Cardiac Channel Mutations in 541 Consecutive Unrelated Patients Referred for Long QT Syndrome Genetic Testing. *Hear. Rhythm* **2005**, *2* (5), 507–517. https://doi.org/10.1016/j.hrthm.2005.01.020.
(208) Zhong, Y.; Han, L.; Ruotolo, B. T. Collisional and Coulombic Unfolding of Gas-Phase

Proteins: High Correlation to Their Domain Structures in Solution. *Angew. Chemie Int. Ed.* **2014**, *53* (35), 9209–9212. https://doi.org/10.1002/anie.201403784.
(209) Cho, W.; Stahelin, R. V. Membrane-Protein Interactions in Cell Signaling and Membrane Trafficking. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34* (1), 119–151. https://doi.org/10.1146/annurev.biophys.33.110502.133337.
(210) Zhou, H.-X.; Cross, T. A. Influences of Membrane Mimetic Environments on Membrane Protein Structures. *Annu. Rev. Biophys.* **2013**, *42* (1), 361–392. https://doi.org/10.1146/annurev-biophys-083012-130326.
(211) Barnaba, C.; Ramamoorthy, A. Picturing the Membrane-Assisted Choreography of Cytochrome P450 with Lipid Nanodiscs. *ChemPhysChem* **2018**, *19* (20), 2603–2613. https://doi.org/10.1002/cphc.201800444.
(212) Waberer, L.; Henrich, E.; Peetz, O.; Morgner, N.; Dötsch, V.; Bernhard, F.; Volknandt, W. The Synaptic Vesicle Protein SV31 Assembles into a Dimer and Transports Zn2+. *J. Neurochem.* **2017**, *140* (2), 280–293. https://doi.org/10.1111/jnc.13886.
(213) Hopper, J. T. S.; Yu, Y. T. C.; Li, D.; Raymond, A.; Bostock, M.; Liko, I.; Mikhailov, V.; Laganowsky, A.; Benesch, J. L. P.; Caffrey, M.; Nietlispach, D.; Robinson, C. V. Detergent-Free Mass Spectrometry of Membrane Protein Complexes. *Nat. Methods* **2013**, *10* (12), 1206–1208. https://doi.org/10.1038/nmeth.2691.
(214) Hellwig, N.; Peetz, O.; Ahdash, Z.; Tascó, I.; Booth, P. J.; Mikusevic, V.; Diskowski, M.; Politis, A.; Hellmich, Y.; Hä, I.; Reading, E.; Morgner, N. Native Mass Spectrometry Goes More Native: Investigation of Membrane Protein Complexes Directly from SMALPs †. *Chem. Commun* **2018**, *54*, 13702. https://doi.org/10.1039/c8cc06284f.
(215) Dror S. Chorev1; Lindsay A. Baker2; Di Wu1; Victoria Beilsten-Edmands1; Sarah L. Rouse3; Tzviya Zeev-Ben-Mordehai2; * C. J.; Firdaus Samsudin5; Christoph Gerle6,7, S. K.; Alastair G. Stewart8, 9; Stephen J. Matthews3; Kay Grünewald2,10, C. V. R.; †. Protein Assemblies Ejected Directly from Native Membranes Yield Complexes for Mass Spectrometry.
(216) Wang, S. C.; Politis, A.; Di Bartolo, N.; Bavro, V. N.; Tucker, S. J.; Booth, P. J.; Barrera, N. P.; Robinson, C. V. Ion Mobility Mass Spectrometry of Two Tetrameric Membrane Protein Complexes Reveals Compact Structures and Differences in Stability and Packing. *J. Am. Chem. Soc.* **2010**, *132* (44), 15468–15470. https://doi.org/10.1021/ja104312e.
(217) Hosea, N. A.; Miller, G. P.; Guengerich, F. P. Elucidation of Distinct Ligand Binding Sites for Cytochrome P450 3A4. *Biochemistry* **2000**, *39* (20), 5929–5939. https://doi.org/10.1021/bi992765t.
(218) Gillam, E. M. J.; Baba, T.; Kim, B. R.; Ohmori, S.; Guengerich, F. P. Expression of Modified Human Cytochrome P450 3A4 in Escherichia Coli and Purification and Reconstitution of the Enzyme. *Arch. Biochem. Biophys.* **1993**, *305* (1), 123–131. https://doi.org/10.1006/abbi.1993.1401.
(219) Wade, J. H.; Jones, J. D.; Lenov, I. L.; Riordan, C. M.; Sligar, S. G.; Bailey, R. C. Microfluidic Platform for Efficient Nanodisc Assembly, Membrane Protein Incorporation, and Purification. *Lab Chip* **2017**, *17* (17), 2951–2959. https://doi.org/10.1039/c7lc00601b.
(220) Schlebach, J. P.; Peng, D.; Kroncke, B. M.; Mittendorf, K. F.; Narayan, M.; Carter, B. D.; Sanders, C. R. Reversible Folding of Human Peripheral Myelin Protein 22, a Tetraspan Membrane Protein. *Biochemistry* **2013**, *52* (19), 3229–3241. https://doi.org/10.1021/bi301635f.
(221) Polasky, D. A.; Dixit, S. M.; Vallejo, D. D.; Kulju, K. D.; Ruotolo, B. T. An Algorithm

for Building Multi-State Classifiers Based on Collision-Induced Unfolding Data. **2019**. https://doi.org/10.1021/acs.analchem.9b02650.
(222) Marklund, E. G.; Degiacomi, M. T.; Baldwin, A. J.; Benesch, J. L. P.; Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V; Baldwin, A. J.; Benesch, J. L. P. Collision Cross Sections for Structural Proteomics Resource Collision Cross Sections for Structural Proteomics. *Struct. Des.* **2015**, *23* (4), 791–799. https://doi.org/10.1016/j.str.2015.02.010.
(223) Papadopoulos, V.; Fan, J.; Zirkin, B. Translocator Protein (18 KDa): An Update on Its Function in Steroidogenesis. *J. Neuroendocrinol.* **2018**, *30* (2), e12500. https://doi.org/10.1111/jne.12500.
(224) Li, F.; Liu, J.; Garavito, R. M.; Ferguson-Miller, S. Evolving Understanding of Translocator Protein 18 KDa (TSPO). *Pharmacol. Res.* **2015**, *99*, 404–409. https://doi.org/10.1016/j.phrs.2015.03.022.
(225) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2002**, *2* (8), 853–856. https://doi.org/10.1021/nl025623k.
(226) Sahin, C.; Reid, D. J.; Marty, M. T.; Landreh, M. Scratching the Surface: Native Mass Spectrometry of Peripheral Membrane Protein Complexes. **2020**. https://doi.org/10.1042/BST20190787.
(227) Allen, K. N.; Entova, S.; Ray, L. C.; Imperiali, B. Monotopic Membrane Proteins Join the Fold. *Trends in Biochemical Sciences*. Elsevier Ltd January 1, 2019, pp 7–20. https://doi.org/10.1016/j.tibs.2018.09.013.
(228) Shah, M. B.; Jang, H. H.; Wilderman, P. R.; Lee, D.; Li, S.; Zhang, Q.; Stout, C. D.; Halpert, J. R. Effect of Detergent Binding on Cytochrome P450 2B4 Structure as Analyzed by X-Ray Crystallography and Deuterium-Exchange Mass Spectrometry. *Biophys. Chem.* **2016**, *216*, 1–8. https://doi.org/10.1016/j.bpc.2016.05.007.
(229) Lössl, P.; Snijder, J.; Heck, A. J. R. Boundaries of Mass Resolution in Native Mass Spectrometry. *J. Am. Soc. Mass Spectrom* **2014**, *25*, 906–917. https://doi.org/10.1007/s13361-014-0874-3.
(230) Denisov, I. G.; Sligar, S. G. Cytochromes P450 in Nanodiscs. *Biochimica et Biophysica Acta - Proteins and Proteomics*. Elsevier January 1, 2011, pp 223–229. https://doi.org/10.1016/j.bbapap.2010.05.017.
(231) Grinkova, Y. V.; Denisov, I. G.; Sligar, S. G. Functional Reconstitution of Monomeric CYP3A4 with Multiple Cytochrome P450 Reductase Molecules in Nanodiscs. *Biochem. Biophys. Res. Commun.* **2010**, *398* (2), 194–198. https://doi.org/10.1016/j.bbrc.2010.06.058.
(232) Han, L.; Ruotolo, B. T. Traveling-Wave Ion Mobility-Mass Spectrometry Reveals Additional Mechanistic Details in the Stabilization of Protein Complex Ions through Tuned Salt Additives. https://doi.org/10.1007/s12127-013-0121-9.
(233) Bayburt, T. H.; Sligar, S. G. Membrane Protein Assembly into Nanodiscs. *FEBS Lett.* **2010**, *584*, 1721–1727. https://doi.org/10.1016/j.febslet.2009.10.024.