THE UNIVERSITY OF CHICAGO

MEMBRANE SELECTIVITY OF PHOSPHATIDYLSERINE RECOGNIZING PROTEINS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY

ZHILIANG GONG

CHICAGO, ILLINOIS

JUNE 2017

To my wife, parents, and sister.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

While the thesis might fill the score sheet for my past five and a half years, the people behind the work filled my life with joy and meaning.

First and foremost, I would like to thank my mentor, Ka Yee Lee, for her guidance, support, encouragement, and tolerance. She has exemplified the ideal mentor and a great scientist. Her deep knowledge in the lipid membrane and related experimental and computational techniques provided me the bread and butter for conducting the studies. Her unwaning support and encouragement allowed me the freedom to pursue directions that interested me the most. She has tolerated my mistakes after mistakes, yet ever so encouraging. The more time goes by, the more I realized how lucky I was to belong to a group led by Ka Yee.

I owe my deep gratitude to past and current the Lee Lab members, who have made the lab my second family. I have been lucky to have some great close collaborators in the lab. Greg Tietjen helped me start with the Tim proteins, and taught me a lot on x-ray scattering on several scattering trips. Together we made important discoveries on Tim4 binding. Dan Kerr has helped a great amount with data analysis and molecular dynamics, taught me lots of statistics, and has been a constant pal in x-ray scattering trips. Luke Hwang has offered great help in my x-ray scattering trips, and great encouragement whenever I felt down. Ben Slaw has helped greatly with his acute observation on his first x-ray scattering trip with me, and provided much help with his careful work on the trough. Kathleen Cao taught me everything I need to know about the Langmuir trough, and rushed to help multiple times for experiments on the trough. Michael Henderson has exemplified a first-class scientist in the making. His meticulous planning and

dedicated execution have been an inspiration for me. His help with several x-ray scattering trips enabled us to make significantly better use of the beam time. Undergraduate Tiffany Suwatthee has proven a great help in the x-ray scattering and fluorescence spectroscopy experiments. I believe she would succeed no matter what her future course is, given her fearless appetite for knowledge, quick grasp of techniques, and high level attention to details. I appreciate the opportunity to have mentored two brilliant summer high school students, Nathaniel Posner and Ilana Emanuel. I want to thank all other Lee Lab members for making the Lee Lab such a great family: Peter Chung, Alessandra Leong, Andrew Molina, Charles Heffern, Jiayu Wang, Niels Holten-Andersen, Jaemin Chin, etc.

I also owe my deep gratitude to collaborators outside of the Lee Lab. A large portion of my data were gathered at APS Sector 15, which would not have been possible without the help of Binhua Lin and Wei Bu. Binhua has kept the facility in great shape so that our trips are always smooth. She has deep understanding of our projects and their significance. It was partially through her advocacy that these projects had the chance of getting precious bearing significant fruits. Wei has been my mentor in x-ray fluorescence experiments. His depth of knowledge in the facility has been a necessity for our experiments. His willingness to extend his hours to help, and the countless hours and nights he spent with us on our projects deeply touched all of us. The fluorescence spectroscopy experiments were mainly conducted in the Biophysics Core Facility, managed very well by Elena Solomaha. I want to thank her for her dedication in keeping the facility running well. On the occasions that the fluorimeter had trouble, she always rushed to fix them so that my projects would suffer the least delay.

I want to thank Prof. Erin Adams for her support in providing the protein samples and her belief in our biophysical studies. Thank Adrienne Luoma and Charles Dulberger from the Adams Lab for preparing protein batches for me. All of our protein studies could not have been possible without the continuing support from Erin and her lab. I want to thank Prof. Ted Steck for our intriguing discussions on the thermodynamics of protein binding. His breadth and depth of knowledge gave me a constant source of inspiration. I also want to thank Prof. Ho Sup Yoon and Dr. Sreekanth Rajan at the Nanyang Technology University for providing with us the cDNA of the C2 domain of MFG-E8.

Thank my committee, Professors Bozhi Tian, Benoit Roux, and Erin Adams. Erin's support for our projects has been indispensable. Bozhi has been my personal role model, as a brilliant scientist and great futurist thinker. Benoit taught our lab a lot on the art and science of molecular dynamics.

For the past five and a half years, all the years before that, and all the years to come, I thank my family. Thank my wife, Shanshan, for deciding to come to this foreign nation with me, for her tireless support, and for making life always a joy. Thank my parents for their hard laboring to bring me up. Thank my sister, Liping, for sacrificing high school so that I can receive a good education, and for writing me a check so that I have the funds to apply for graduate schools. Thank Grandma and Grandpa for their love and support. I'm deeply sorry for not making it to your funerals, Aunt Yuanxiang Fu, Uncle Chunhua Fu, and Uncle Youbao He. I miss you so much.

CHAPTER 1

INTRODUCTION

1.1 Understanding the physicochemical sensitivity of phosphatidylserine recognizing proteins

The work presented in this thesis is an effort to gain a more holistic understanding of physicochemical regulators that drive the membrane selectivity of proteins that recognize phosphatidylserine (PS). To date, more than forty PS-binding proteins have been identified [1, 2, 3]; however, what has been largely lacking is the general approach for describing and understanding PS/protein interaction in the membrane context. While the importance of various physicochemical properties of lipid membranes in protein/lipid interactions has been more progressively recognized since the proposal of the fluid mosaic model in the early 1970s [4], most studies in the field have focused on one or two aspects of the lipid membrane, unable to account for the important synergistic effects arising from multiple factors. Strides in resolving the structure of membrane binding proteins in solution or in crystal have enabled molecular-level description of protein-lipid interactions; however, these studies tend to focus on local chemical identities at the expense of general membrane properties and their effects on protein structural arrangement.

The work presented here clearly demonstrate the necessity of considering both the physical and chemical properties of the membrane in understanding lipid-protein interactions. In doing so, a generalizable approach for understanding lipid-protein interactions has emerged. First, the

interaction configuration of the membrane-bound protein is proposed by combining molecular dynamics simulations with experimental methods such as x-ray reflectivity from lipid monolayers with proteins adsorbed. Then the lipid-protein interaction is decomposed into various possible interactions components, such as electrostatic and hydrophobic interactions, which are then probed using biochemical assays by varying relevant physical and chemical parameters as possible modulators for the membrane selectivity of the protein. The newly identified interaction components are treated as new dimensions in the parameter space that determines the lipid-protein interaction. Lastly, possible interplay and synergetic effects among the previously known and added parameters are further examined for a holistic view of the possible behavior of the protein within the parameter space.

1.2 A physical chemistry approach to study the PS recognition mechanism of the Tim and MFG-E8 proteins

This framework of interaction components is borne out of studying four proteins with vital roles in the innate immunity system. These four proteins include three members of the T cell immunoglobulin and mucin domain (TIM) family - Tim1, Tim3, and Tim4 - and milk fat globule epidermal growth factor 8 (MFG-E8), with Tim3 (Chapter 3) and MFG-E8 (Chapter 4) being the focus of this thesis. While these proteins share the common interaction with PS, essential for their distinct physiological functions, their differential sensitivity to PS was elusive. The Tim proteins are cell surface proteins that survey the local environment with a calcium-dependent PS binding domain, while MFG-E8 is soluble and contains a calcium-independent PS binding domain. As these four proteins play critical roles in disease and health, they were chosen to give

a collective and compelling example that validates the physicochemical/biophysical approach to understanding immunological regulation mechanisms. Adopting the approach of structure-informed physicochemical sensitivity studies, the lipid-protein interactions involving the four proteins, in particular Tim3 and MFG-E8, have been systematically examined and elucidated. New plausible factors in Tim3 regulation were discovered, confirming the suitability of the approach.

A diverse set of physical chemistry tools were prerequisite for carrying out studies for the various parts of the framework. Molecular dynamics simulations and x-ray reflectivity methods were employed together for resolving protein configuration upon interacting with lipids. Fluorescence spectroscopy experiments were used to quantify the binding free energy of the lipid-protein interaction under various conditions. X-ray fluorescence experiments were performed and data analytical software tools were built to provide information on the sequestration of Ca^{2+} by the lipid membrane.

1.3 Organization

Chapter 2 provides an introduction to the experimental and computational tools employed in this thesis work, and the relevant background information in understanding the role of PS in physiological pathways involving various receptors, especially the Tim's and MFG-E8. Chapter 3 introduces the methodology of the thesis. Chapter 4 describes findings pertinent to the interaction of the Tim proteins with PS-containing membranes, with a particular focus on Tim3. Chapter 5 details findings on the interaction of MFG-E8 with PS-containing membranes. Chapter

6 presents the development and testing of a graphical user interface tool for intuitive and accurate analysis of total external reflection x-ray fluorescence data to aid the analysis of calcium sequestration in the context of Tim binding to the lipid membrane, as well as to make this highly accurate and widely applicable experimental technique more accessible.

1.4 References

1. Stace, C. and Ktistakis, N. Phosphatidic acid- and phosphatidylserine-binding proteins. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*. **2006**, *1761*, 8, 913–926.
2. Zhou Z. New Phosphatidylserine receptors: clearance of apoptotic cells and more. *Developmental Cell.* **2007**, *13*, 6, 759-760.
3. Armstrong, A. and Ravichandran, K.S. Phosphatidylserine receptors: what is the new RAGE? *EMBO reports*. **2011**, *12*, 4, 287–288.
4. Singer SJ; Nicolson GL. The fluid mosaic model of the structure of cell membranes. *Science*. **1972**, *175*, 720-731.

CHAPTER 2

PS-SENSING PROTEINS AND THEIR PHYSIOLOGICAL RELEVANCE

2.1 Phospholipids in the cellular membrane

Cellular membranes are mixtures of protein and lipids. Among these lipids, about 70% are phospholipids, while the other 30% are mainly cholesterol and sphingomyelin. Among the phospholipids, the majority are neutral in charge. These phospholipids contain a hydrophobic tailgroup with two acyl chains attached to the glycerol group, and a headgroup containing a phosphate group and another characteristic group. While lipid molecules used to be viewed as the inert backdrop for membrane protein functioning proposed in the influential mosaic fluid model in the 1970s [1], their active roles in regulating crucial aspects of cellular functioning have been increasingly studied and recognized.

The diversity of lipids often evades attention. To date, over forty thousand lipid molecules have been studied or described in the LIPID MAPS Structure Database, of which close to ten thousand are phospholipids [2]. Phospholipids can be broadly classified by their headgroup type. As such, the zwitterionic phosphatidylcholine (PC) is the most abundant, representing 40 to 50% of all phospholipids [3]. Next to PC is phosphatidylethanolamine (PE), which represents 20 to 50% of all phospholipids [4]. A significant portion of cellular phospholipids is negatively charged, among which the first two most abundant are phosphatidylserine (PS) and phosphatidic acid (PA). PS is the most abundant negatively charged lipid in cellular membranes, representing about 7 to 15% of phospholipids [3, 4], and PA represents about 1 to 4% [4]. Representative

structures of these lipids are seen in **Figure 2.1**. The tailgroup of phospholipids are very diverse in the lengths of both acyl chains, and the number of saturation sites on each chain. Both the chain length and the number of saturation sites are tightly controlled in eukaryotic cells, reflecting the importance of both parameters [5]. In vertebrates, the average acyl chain length is around 18 carbons, with around 40% saturated chains, and 60% unsaturated chains, including mono- and poly-unsaturated [5].

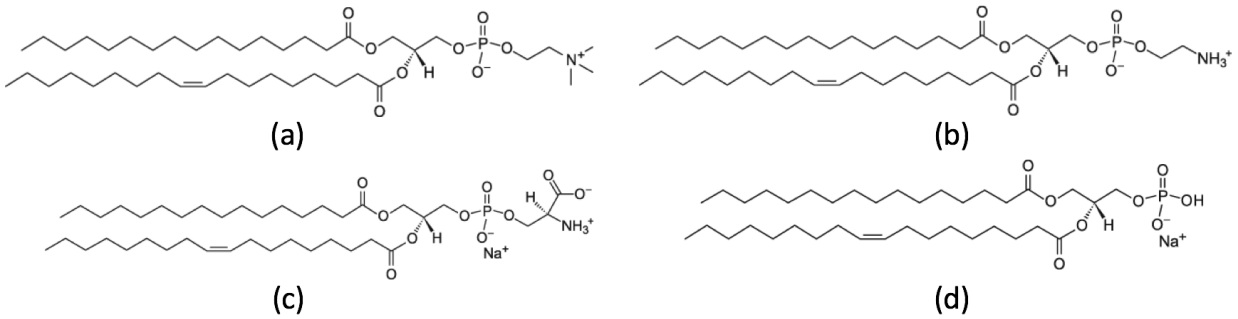

Figure 2.1: representative lipid structures. (a): 1-pamiltoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC). (b): 1-pamiltoyl-2-oleoyl-*sn*-glycero-3-phosphatidylethanolamine (POPE). (c): 1-pamiltoyl-2-oleoyl-*sn*-glycero-3-phosphatidylserine (POPS). (d): 1-pamiltoyl-2-oleoyl-*sn*-glycero-3-phosphatic acid (POPA) (sodium salt).

Eukaryotic cells generally maintain a compositional asymmetry across the inner leaflet and outer leaflet of the plasma membrane, facilitated by ATP-consuming enzymes termed flippase and floppase [6]. In this asymmetry, phospholipids in the outer leaflet contains predominantly PC, and low levels PE, perhaps to avoid nonspecific adhesion from positively charged entities in the extracellular environment [7]. In contrast, PS lipids almost exclusively reside in the inner leaflet [7]. Exposure of PS from the inner leaflet to the outer leaflet can be facilitated by enzymes called scramblases [7], and is a strong cellular signal for the induction of a variety of cellular events, depending on the cell type and possibly the particular physiologic condition. On platelet

membranes, exposure of PS marks the onset of a cascade of interactions leading to blood coagulation [8]. PS exposure is also a hallmark of apoptosis that various apoptotic receptors recognize to the effect of guiding efficient phagocytosis of these dying cells [9]. This phagocytosis pathway could also be hijacked by viruses with PS exposed in their lipid coating to coax engulfment by phagocytes, thus leading to the infection of these cells [10]. Other important biological processes that dependence on the exposure of PS include myotube formation, vesicle fusion, cell division, sperm capacitation, and many others [11].

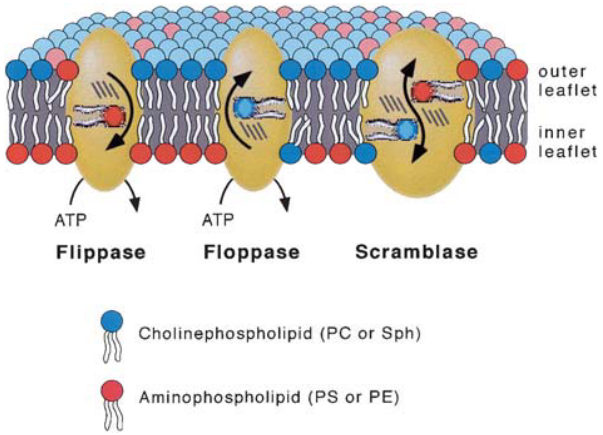

Figure 2.2: schematic of the asymmetry of the plasma membrane and the enzymes involved. *Adapted from [7]. Written permission granted by the publisher.*

2.2 PS-binding proteins

PS exposure induces downstream events, such as phagocytosis of apoptotic cells and blood coagulation, via interacting with PS receptors in the extra cellular fluid, in the blood, or on the membrane surface of other cells. A growing array of more than forty proteins has been found to recognize PS [12, 13, 14]. These receptors have been found to interact with PS via two general ways: 1) charge-charge interaction between basic amino acid residues and the PS headgroup, and

2) via recognizable domains, such as the annexin core domains, the protein kinase C (PKC)-like C2 domains, the discoidin-like C2 domains, the Gla domains, and the immunoglobulin domains [12]. Such PS-binding domains are shared within a protein family, or across multiple families to serve as building blocks for more complex functional proteins. For example, the annexin core domain is shared within the annexins, while the PS-binding PKC-like C2 domain is shared among diverse proteins, including PKC, synaptotagmin, and rabphilin [12]. The T cell immunoglobulin and mucin domain (Tim) family protein share a calcium-dependent PS-binding immunoglobulin (Ig) domain [3, 4], while MFG-E8 contains a calcium-independent PS-binding PKC-like C2 domain [15]. Associated with these diverse PS receptors are diverse cellular events mediated by these players, as discussed in the previous section.

2.3 Model System 1: Tim Proteins

The Tim protein family in humans includes Tim1, Tim3 and Tim4. The three Tim proteins share common structural features, including a well folded immunoglobulin (Ig) domain at their N-terminus, a mucin domain connection to the plasma membrane, a transmembrane domain, and a C-terminal tail buried in the membrane, as seen in **Figure 2.3a** [16]. In addition, they share a calcium-dependent PS binding site on the Ig domain [17, 18]. Despite similarities in their structure and ligand binding, the Tim proteins serve distinct and often opposing physiological functions. In particular, Tim4 enables macrophages and other phagocytes to attach to target membranes bearing high levels of PS (such as apoptotic cells). Tim1 is a positive effector of T cell immunity [16, 19, 20, 21], while Tim3 is an immune checkpoint receptor that dampens immune response [22].

This family of protein was chosen for the studies undertaken in this thesis due to their significant physiologic relevance, and their puzzling differential interaction pattern with PS despite of their similar structures. The crystal structures of the Ig domains of Tim1 (protein data bank code: 2OR8) [18], Tim3 (3KAA) [23], and Tim4 (PDB code: 3BI9) [24] have been resolved. As seen in the overlaid crystal structures in **Figure 2.3b**, the Ig domains of the three Tim proteins looked very similar. The root mean square deviations (RMSD's) between of the pair-wise comparison are all smaller than 1.5Å. The main deviation among them comes from the flexible peripheral loops.

As this thesis is focused in particular on Tim3, additional physiologic background information on Tim3 is provided here to better contextualize the studies presented later on in Chapter 3.Tim3 is an immune checkpoint receptor expressed on the surface of a variety of immune cell types including T helper 1 cells, effector T cells, regulatory T cells, macrophages, and natural killer cells [22]. It functions in the negative regulation of immune responses [25]. Tim3 suppresses autoimmune and alloimmune responses and promotes peripheral immune tolerance [26]. Both its over- and under-expression evoke pathological responses. On the one hand, blocking Tim3 activity with specific antibodies leads to spontaneous autoimmune reactions [26, 27]. On the other hand, overexpression of Tim3 is associated with various disease states, such as tumor growth [28, 29, 30], autoimmune disorders [27], type I diabetes [27], and chronic viral infections like the hepatitis C [31, 32, 33], hepatitis B [34, 35], AIDS [36], and influenza [37]. Immunosuppression via Tim3 could be exploited in the treatment of immune disorders [38, 39], such as various types of cancer [40, 41, 42, 43, 44], multiple sclerosis [45, 46], and acute alcoholic hepatitis [47].

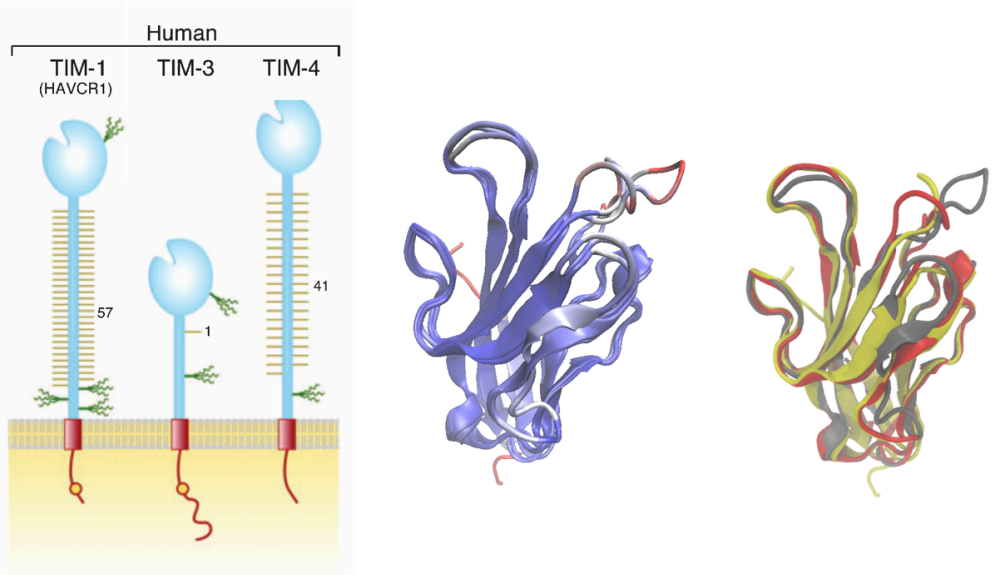

Figure 2.3: structures of the Tim proteins. (a): (adapted from [16]. *Written permission for reuse in thesis granted by publisher*.) Schematic representation of Tim protein structures. The glycosylation sites in IgV domains are approximately positioned according to the crystal structures. Glycosylation sites in the mucin domain were predicted with NetOglyc and NetNglyc and are positioned approximately. (b): Overlay of the three Tim proteins by minimizing RMSD. Here deep blue represents high structural similarity, while red represents high level of disparity. The pair-wise RMSD values were 1.34, 1.27, 1.40 for Tim1-Tim3, Tim1-Tim3, and Tim3-Tim4 respectively. (c): Overlay of the Tim proteins by color. Yellow: Tim1, gray: Tim3, red: Tim4.

While the role of Tim3 as an immune checkpoint receptor is widely accepted, its mechanism of action is poorly understood and, at present, confusing due to its potential interaction with a variety of ligands [16]. Among these are galectin-9 [48], certain carbohydrates [49], PS [50], high-mobility group protein-1 (HMGB-1) [51], Ceacam-1 [52] and Tim4 [53]. Tim3 could participate in immunosuppression by galectin-9 [48, 54, 55, 56, 57, 58, 59]; however, galectin-9 does not interact with Tim3 specifically, and participates in several critical immune functions independent of Tim3 [60, 61, 62, 63, 64]. Moreover, the function of Tim3 as a receptor for galectin-9 has been questioned [65]. The binding of PS by the Tim3 on the plasma membrane of

macrophages facilitates their phagocytosis of apoptotic cells [66]. The interaction of Tim3 with HMGB-1 interferes with the recruitment of nucleic acids into the endosomes of dendritic cells, thereby diminishing the immunogenicity of nucleic acids released by dying tumor cells [51]. Ceacam-1 was reported to bind Tim3 with 1:1 stoichiometry, thereby promoting the maturation and surface expression of Tim3 [52]. These diverse players complicate the analysis of the molecular basis for the physiological functions of Tim3.

2.4 Model System 2: MFG-E8

As the Tim proteins share the calcium-dependent PS binding Ig domain, milk fat globule-epidermal growth factor 8 (MFG-E8) was studied in this thesis to offer a contrast as a calcium-independent PS receptor. MFG-E8, also known as lactadherin, is a secreted protein from multiple cell types, including macrophages [67] and apoptotic cells [68]. The full length MFG-E8 contains one to three epidermal growth factor (EGF)-like domains at the N-terminus [69], and two C-type membrane targeting domains, a protein kinase C (PKC) conserved 1 (C1) and PKC conserved 2 (C2) on the C-terminus [67]. MFG-E8 acts as a linker between apoptotic cells and macrophages [67], with its C1 and C2 domains binding to the exposed phosphatidylserine (PS) on apoptotic cells, and its EGF-like domains binding to $\alpha_v\beta_3$ and $\alpha_v\beta_5$ integrins on macrophages [67, 70, 71, 72, 73], leading to the phagocytosis of the attached apoptotic cells. Besides apoptotic cells, MFG-E8 has been found to be involved in the clearance of vesicular debris, such as platelet derived vesicles [74], the photosensitive outer segments in the mammalian retina [75, 76]. MFG-E8 has also been found to promote obesity by mediating the uptake of dietary fats and serum fatty acids [77]. Normal expression and functioning of MFG-E8 is crucial for homeostasis, as

evidenced by a strong correlation of MFG-E8 aberrant functioning and diseases, such as the accumulation of apoptotic cells [78], autoimmune diseases [79], Alzhermier's disease [80], Parkinson's disease [81], neural inflammation [82], neuron loss [83], diabetes [84], and cancer [85, 86, 87], raising the prospects of MFG-E8 as a potential therapeutic target [88, 89, 90, 91, 92]. Due to its exceptional binding affinity with PS, it has also been considered as an alternative Ca^{2+} independent marker for phosphatidylserine [93, 94, 95].

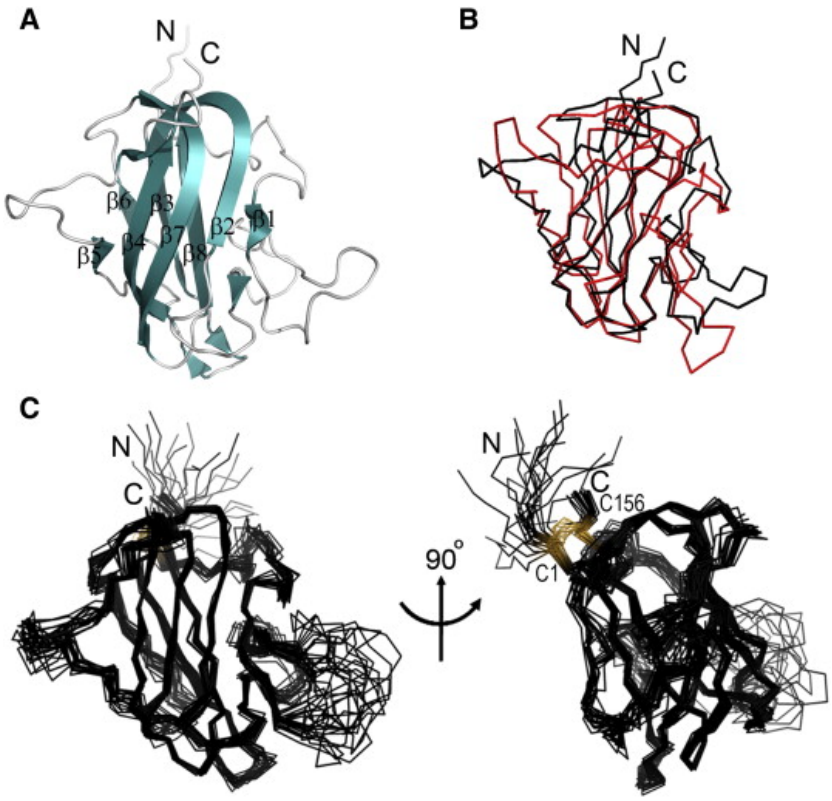

Figure 2.4: NMR Structure of the MFG-E8 C2 domain. (A) Structure of the MFG-E8 C2 domain conformer with the lowest energy. (B) Structural alignment of mouse MFG-E8 C2 (black) and bovine MFG-E8 C2 (red; PDB code: 3BN6) domains. (C) Ensemble of the top 20 MFG-E8 C2 domain conformers. Cys1 and Cys156 with disulfide bridges are displayed in yellow. (Adapted from [15], *Written permission for reuse in thesis granted by publisher.*)

MFG-E8 interacts with PS-containing lipid membranes mainly via its C2 domain [70, 71, 96, 97, 98]. The interaction of the C2 domain with PS has been studied using four kinds of model systems, and results vary with different model systems. Originally the C2 domain was shown to bind to microplate wells coated with PS, phophatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidic acid (PA), phosphatidylglycerol (PG), and cardiolipin (CL), but not with phosphatidylcholine (PC) and sphingomyelin [70]. Later on, the domain was shown to be stereoselective to phosphatidyl-*L*-serine in lipid vesicles and not to PC [95]; however, it was also shown to bind various lipids, including PS, PC, and PG in micelles [95, 98, 15]. Therefore, it seems that the C2 domain is selective to PS in vesicle systems, but binds a variety of lipids, including PC, PS, PG, PE, PA in lipid-coated microplate wells and micelles. As C2 domain binding is dependent on vesicle curvature [95], the difference in its binding to micelles *vs.* vesicle might be explained partly in a curvature argument.

The crystal structure of its bovine copy was resolved [96, 97], and more recently, the solution structure of the C2 domain of the murine MFG-E8 was resolved by NMR [15], as shown in **Figure 2.4.** These structures provide a basis for molecular interpretations of the interaction of the C2 domain with lipid membranes. Using mutagenesis, Shao et al. found the possible involvement of hydrophobic residues in the PS-binding of the bovine C2 domain using lipid-coated glass microspheres of 1.6 μM diameter [97]. In the study of Ye et al., interactions between the C2 domain and PS were analyzed by mixing the protein with micelles made of short PS lipids with a 6-carbon tail (6:0 PS) [15]. The authors discovered a surface on the domain containing several hydrophobic, aliphatic, and basic residues that comes into contact with the lipid micelles, and proposed a putative PS binding groove on the C2 domain. The use of lipid micelles, however,

could artificially enhance the affinity due to the increased curvature as discussed above. As different model systems gave dramatically different results, it is critical to design and employ a model system that closely mimic the physiologic conditions, thus the use of liposomes should be encouraged for the study of lipid interactions involving the C2 domain.

2.5 References

1. Singer SJ; Nicolson GL. The fluid mosaic model of the structure of cell membranes. *Science*. **1972**, *175*, 720-731.
2. Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill A. H. Jr.; Murphy, R. C.; Raetz, C. R.; Russell, D. W.; Subramaniam, S. LMSD: Lipid maps structure database. *Nucleic Acids Res.* **2007,** *35*, D527-D532.
3. Vance, J. E. Phosphatidylserine and phosphatidylethanolamine in mammalian cells: Two metabolically related aminophospholipids. *J. Lipid Res.* **2008,** *49*, 1377-1387.
4. Voelker, D. R. Organelle biogenesis and intracellular lipid transport in eukaryotes. *Microbiol. Rev.* **1991,** *55*, 543-560.
5. Pamplona, R. Membrane phospholipids, lipoxidative damage and molecular integrity: A causal role in aging and longevity. *Biochim. Biophys. Acta*. **2008,** *10*, 1249-1262.
6. Devaux, P. F.; Herrmann, A.; Ohlwein, N.; Kozlov, M. M. How lipid flippases can modulate membrane structure. *Biochim. Biophys. Acta*. **2008,** *1778*, 1691-1600.
7. Zwaal, R. F.; Comfurius, P.; Bevers, E. M. Surface exposure of phosphatidylserine in pathological cells. *Cell Mol. Life Sci.* **2005,** *62*, 971-988.
8. Lentz, B. R. Exposure of platelet membrane phosphatidylserine regulates blood coagulation. *Prog. Lipid Res.* **2003,** *42*, 423-438.
9. Segawa, K.; Nagata, S. An apoptotic 'eat me' signal: Phosphatidylserine exposure. *Trends Cell Biol.* **2015,** *25*, 639-650.
10. Morizono, K.; Chen, I. S. Role of phosphatidylserine receptors in enveloped virus infection. *J. Virol.* **2014,** *88*, 4275-4290.

11. Fadel, B.; Xue, D. The ins and outs of phospholipid asymmetry in the plasmamembrane: roles in health and disease. *Crit. Rev. Biochem. Mol. Biol.* **2009**, *44*, 264-277.
12. Stace, C. and Ktistakis, N. Phosphatidic acid- and phosphatidylserine-binding proteins. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*. **2006**, *1761*, 8, 913–926.
13. Zhou Z. New Phosphatidylserine receptors: clearance of apoptotic cells and more. *Developmental Cell.* **2007**, *13*, 6, 759-760.
14. Armstrong, A. and Ravichandran, K.S. Phosphatidylserine receptors: what is the new RAGE? *EMBO reports*. **2011**, *12*, 4, 287–288.
15. Ye, H.; Li, B.; Subramanian, V.; Choi, B.-H.; Liang, Y.; Harikishore, A.; Chakraborty, G.; Baek, K.; Yoon, H. S. NMR solution structure of C2 domain of MFG-E8 and insights into its molecular recognition with phosphatidylserine. *Biochim. Biophys. Acta*. **2013,** *1828*, 1083–1093.
16. Freeman, G.; Casasnovas, J.; Umetsu, D.; DeKruyff, R. TIM genes: a family of cell surface phosphatidylserine receptors that regulate innate and adaptive immunity. *Immunol. Rev*. **2010,** *235*, 172–189.
17. Cao, E.; Zang, X.; Ramagopal, U. A.; Mukhopadhaya, A.; Fedorov, A.; Federov, E.; Zencheck, W. D.; Lary, J. W.; Cole, J. L.; Deng, H.; Xiao, H.; Dilorenzo, T. P., Allison, J. P.; Nathenson, S. G.; Almo, S. C. T cell immunoglobulin mucin-3 crystal structure reveals a galectin-9-independent ligand-binding surface. *Immunity*. **2007,** *26*, 311–321.
18. Santiago, C.; Ballesteros, A.; Tami, C.; Martínez-Muñoz, L.; Kaplan, G. G.; Casasnovas, J. M. Structures of T Cell immunoglobulin mucin receptors 1 and 2 reveal mechanisms for regulation of immune responses by the TIM receptor family. *Immunity.* **2007,** *25,* 299-310.
19. Kuchroo, V. K.; Umetsu, D. T.; DeKruyff, R. H.; Freeman, G. J. 2003. The TIM gene family: emerging roles in immunity and disease. *Nat. Rev. Immunol.* **2003,** *3*, 454-462.
20. Kuchroo, V.; Meyers, J.; Umetsu, D.; DeKruyff, R. TIM family of genes in immunity and tolerance. *Adv. Immunol*. **2006,** *91*, 227–249.
21. Rodriguez-Manzanet, R.; DeKruyff, R.; Kuchroo, V.; Umetsu, D. The costimulatory role of TIM molecules. *Immunol. Rev*. **2009,** *229*, 259–270.
22. Anderson, A. C.; Joller, N; Kuchroo, V. K. Lag-3, Tim-3, and TIGIT: co-inhibitory receptors with specialized functions in immune regulation. *Immunity*. **2016,** *44*, 989–1004.

23. DeKruyff, R.; Bu, X.; Ballesteros, A.; Santiago, C.; Chim, Y. -L.; Lee, H. -H.; Karisola, P.; Pichavant, M.; Kaplan, G.; Umetsu, D.; Freeman, G.; Casasnovas, J. T. Cell/transmembrane, Ig, and mucin-3 allelic variants differentially recognize phosphatidylserine and mediate phagocytosis of apoptotic cells. *J. Immunol*. **2010,** *184*, 1918–1930.
24. Santiago, C.; Ballesteros, A.; Martínez-Muñoz, L.; Mellado, M.; Kaplan, G.; Freeman, G.; Casasnovas, J. Structures of T Cell Immunoglobulin Mucin Protein 4 Show a Metal-Ion-Dependent Ligand Binding Site Where Phosphatidylserine Binds. *Immunity* **2007**, *27*, 941–951.
25. Monney, L.; Sabatos, C.; Gaglia, J.; Ryu, A.; Waldner, H.; Chernova, T.; Manning, S.; Greenfield, E.; Coyle, A.; Sobel, R.; Freeman, G.; Kuchroo, V. Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease. *Nature*. **2002,** *415*, 536–541.
26. Sánchez-Fueyo, A.; Tian, J.; Picarella, D.; Domenig, C.; Zheng, X. X.; Sabatos, C. A.; Manlongat, N.; Bender, O.; Kamradt, T.; Kuchroo, V. K.; Gutiérrez-Ramos, J. -C. C.; Coyle, A. J.; Strom, T. B. Tim-3 inhibits T helper type 1-mediated auto- and alloimmune responses and promotes immunological tolerance. *Nat. Immunol*. **2003,** *4*, 1093–1101.
27. Sabatos, C.; Chakravarti, S.; Cha, E.; Schubart, A.; Sánchez-Fueyo, A.; Zheng, X.; Coyle, A.; Strom, T.; Freeman, G.; Kuchroo, V. Interaction of Tim-3 and Tim-3 ligand regulates T helper type 1 responses and induction of peripheral tolerance. *Nat. Immunol*. **2003,** *4*, 1102–1110.
28. Yan, J.; Zhang, Y.; Zhang, J. -P.; Liang, J.; Li, L.; Zheng, L. Tim-3 Expression Defines Regulatory T Cells in Human Tumors. *PLoS One*. **2013,** *8*, e58006.
29. Anderson, A. C. Tim-3: an emerging target in the cancer immunotherapy landscape. *Cancer Immunol. Res*. **2014**, *2*, 393-398.
30. Cheng, L.; Ruan, Z. Tim-3 and Tim-4 as the potential targets for antitumor therapy. *Hum. Vaccin. Immunother*. **2015,** *11*, 2458–2462.
31. Ji, X.; Ma, C.; Wang, J.; Wu, X.; Niki, T.; Hirashima, M.; Moorman, J.; Yao, Z. HCV-infected hepatocytes drive CD4+CD25+Foxp3+ regulatory T-cell development through the Tim-3/Gal-9 pathway. *Eur. J. Immunol*. **2013,** *43*, 458–467.
32. Golden-Mason, L.; Palmer, B. E.; Kassam, N.; Townshend-Bulson, L.; Livingston, S.; McMahon, B. J.; Castelblanco, N.; Kuchroo, V.; Gretch, D. R.; Rosen, H. R. Negative

immune regulator Tim-3 is overexpressed on T cells in hepatitis C virus infection and its blockade rescues dysfunctional CD4+ and CD8+ T cells. *J. Virol.* **2009,** *83*, 9122-9130.

33. McMahan, R. H.; Golden-Mason, L.; Nishimura, M. I.; McMahon, B. J.; Kemper, M.; Allen, T. M.; Gretch, D. R.; Rosen, H. R. Tim-3 expression on PD-1+ HCV-specific human CTLs is associated with viral persistence, and its blockade restores hepatocyte-directed in vitro cytotoxicity. *J. Clin. Invest.* **2011,** *120*, 4546-4657.
34. Nebbia, G.; Peppa, D.; Schurich, A.; Khanna, P.; Singh, H. D.; Cheng, Y.; Rosenberg, W.; Dusheiko, G.; Gilson, R.; ChinAleong, J.; Kennedy, P.; Maini, M. K. Upregulation of the Tim-3/galectin-9 pathway of T cell exhaustion in chronic hepatitis B virus infection. *PLoS One*. **2012,** *7*, e47648.
35. Wu, W.; Shi, Y.; Li, S.; Zhang, Y.; Liu, Y.; Wu, Y.; Chen, Z. Blockade of Tim-3 signaling restores the virus-specific CD8+ T-cell response in patients with chronic hepatitis B. *Eur. J. Immunol.* **2012,** *42*, 1180-1191.
36. Rahman, A. N.; Clayton, K.; Mujib, S.; Fong, I. W.; Ostrowski, M. A. TIM-3 and its immunoregulatory role in HIV infection. *J. Clin. Cell Immunol.* **2012,** *S7*, (2012).
37. Cho, J. L.; Roche, M.I.; Sandall, B.; Brass, A. L. Enhanced Tim3 activity improves survival after influenza infection. *J. Immunol.* **2012,** *189*, 2879-2889.
38. Wang, F.; He, W.; Zhou, H.; Yuan, J.; Wu, K.; Xu, L.; Chen, Z. The Tim-3 ligand galectin-9 negatively regulates CD8+ alloreactive T cell and prolongs survival of skin graft. *Cell. Immunol.* **2007,** *250*, 68–74.
39. Oikawa, T.; Kamimura, Y.; Akiba, H.; Yagita, H.; Okumura, K.; Takahashi, H.; Zeniya, M.; Tajiri, H.; Azuma, M. Preferential involvement of Tim-3 in the regulation of hepatic CD8+ T cells in murine acute graft-versus-host disease. *J. Immunol.* **2006,** *177*, 4281–4287.
40. Das, M.; Zhu, C.; Kuchroo, V. K. Tim-3 and its role in regulating anti-tumor immunity. *Immunol. Rev*. **2017,** *276*, 97-111.
41. Kim, J. E.; Patel, M. A; Mangraviti, A.; Kim, E. S.; Theodros, D. et al. Combination therapy with anti-PD-1, anti-TIM-3, and focal radiation results in regression of murine gliomas. *Clin. Cancer Res.* **2017,** *23*, 124-136.
42. Liu, Y.; Cai, P.; Wang, N.; Zhang, Q.; Chen, F.; Shi, L.; Zhang, Y.; Wang, L.; Hu, L.. Combined blockade of Tim-3 and MEK inhibitor enhances the efficacy against melanoma. *Biochem. Biophys. Res. Commun.* **2017,** *484*, 378-384.

43. Kleponis, J.; Skelton, R.; Zheng, L. Fueling the engine and releasing the break: combinational therapy of cancer vaccines and immune checkpoint inhibitors. *Cancer Biol. Med.* **2015,** *12*, 201-208.
44. Pardoll, D. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer*. **2012,** *12*, 252–264.
45. Koguchi, K.; Anderson, D. E.; Yang, L.; O'Connor, K. C.; Kuchroo, V. K.; Hafler, D.A. 2006. Dysregulated T cell expression of TIM3 in multiple sclerosis. *J. Exp. Med*. **2006,** *203*, 1413–1418.
46. Yang, L.; Anderson, D. E.; Kuchroo, J.; Hafler, D. A. Lack of TIM-3 immunoregulation in multiple sclerosis. *J. Immunol*. **2008,** *180*, 4409–4414.
47. Markwick, L. J.; Riva, A.; Ryan, J. M.; Cooksley, H.; Palma, E.; Tranah T. H.; Manakkat Vijay, G. K.; Vergis, N.; Thursz, M.; Evans, A.; Wright, G.; Tarff, S.; O'Grady, J.; Williams, R.; Shawcross, D. L., Chokshi, S. Blockade of PD1 and TIM3 restores innate and adaptive immunity in patients with acute alcoholic hepatitis. *Gastroenterology* **2015,** *148,* 590-602.
48. Zhu, C.; Anderson, A.; Schubart, A.; Xiong, H.; Imitola, J.; Khoury, S.; Zheng, X.; Strom, T.; Kuchroo, V. The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nature Immunol*. **2005,** *6*, 1245–1252.
49. Wilker, P. R.; Sedy, J. R.; Grigura, V.; Murphy, T. L.; Murphy, K. M. Evidence for carbohydrate recognition and homotypic and heterotypic binding by the TIM family. *Int. Immunol*. **2007,** *19*, 763–773.
50. DeKruyff, R.; Bu, X.; Ballesteros, A.; Santiago, C.; Chim, Y. -L.; Lee, H. -H.; Karisola, P.; Pichavant, M.; Kaplan, G.; Umetsu, D.; Freeman, G.; Casasnovas, J. T. Cell/transmembrane, Ig, and mucin-3 allelic variants differentially recognize phosphatidylserine and mediate phagocytosis of apoptotic cells. *J. Immunol*. **2010,** *184*, 1918–1930.
51. Chiba, S.; Baghdadi, M.; Akiba, H.; Yoshiyama, H.; Kinoshita, I.; Dosaka-Akita, H.; Fujioka, Y.; Ohba, Y.; Gorman, J. V.; Colgan, J. D.; Hirashima, M.; Uede, T.; Takaoka, A.; Yagita, H.; Jinushi, M. Tumor-infiltrating DCs suppress nucleic acid-mediated innate immune responses through interactions between the receptor TIM-3 and the alarmin HMGB1. *Nature Immunol*. **2012,** *13*, 832–842.
52. Huang, Y. H.; Zhu, C.; Kondo, Y.; Anderson, A. C. et al. CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. *Nature*. **2014,** *517*, 386–390.

53. Ge, R. T.; Zeng, L.; Mo, L. H.; Xu, L. Z.; Zhang, H. P.; Yu, H. Q.; Zhang, M.; Liu, Z. G.; Liu, Z. J.; Yang, P. C. Interaction of TIM4 and TIM3 induces T helper 1 cell apoptosis. *Immunol. Res*. **2016,** *64*, 470–475.
54. Elahi, S.; Niki, T.; Hirashima, M.; Horton, H. Galectin-9 binding to Tim-3 renders activated human CD4+ T cells less susceptible to HIV-1 infection. *Blood*. **2012,** *119*, 4192–4204.
55. Wang, F.; He, W.; Zhou, H.; Yuan, J.; Wu, K.; Xu, L.; Chen, Z. The Tim-3 ligand galectin-9 negatively regulates CD8+ alloreactive T cell and prolongs survival of skin graft. *Cell. Immunol*. **2007,** *250*, 68–74.
56. Jayaraman, P.; Sada-Ovalle, I.; Beladi, S.; Anderson, A. C.; Dardalhon, V.; Hotta, C.; Kuchroo, V. K.; Behar, S. M. Tim3 binding to galectin-9 stimulates antimicrobial immunity. *J. Exp. Med.* **2010**, *207*, 2343-2354.
57. Kikushige, Y.; Miyamoto, T.; Yuda, J.; jabbarzadeh-Tarizi, S.; Shima, T.; Takayanagi, S.; Niiro, H.; Yurino, A.; Miyawaki, K.; Takenaka, K.; iwasaki, H.; Akashi, K. TIM-3/Gal-9 autocrine stimulatory loop drives self-renewal of human myeloid leukemia stem cells and leukemic progression. *Cell Stem. Cell*. **2015,** *17***,** 341-352.
58. Dardalhon, V.; Anderson, A.; Karman, J.; Apetoh, L.; Chandwaskar, R.; Lee, D.; Cornejo, M.; Nishi, N.; Yamauchi, A.; Quintana, F.; Sobel, R.; Hirashima, M.; Kuchroo, V. Tim-3/galectin-9 pathway: regulation of th1 immunity through promotion of CD11b+Ly-6G+ myeloid cells. *J. Immunol*. **2010,** *185*, 1383–1392.
59. Li, Y. H.; Zhou, W. H.; Tao, Y.; Wang, S. C.; Jiang, Y. L. The Galectin-9/Tim-3 pathway is involved in the regulation of NK cell function at the maternal–fetal interface in early pregnancy. *Cell. Mol. Immunol.* **2016,** *13*, 73-81.
60. Bi, S.; Hong, P. W.; Lee, B.; Baum, L. G. Galectin-9 binding to cell surface protein disulfide isomerase regulates the redox environment to enhance T-cell migration and HIV entry. *Proc. Natl. Acad. Sci*. **2011,** *108*, 10650-10655.
61. Su, E. W.; Bi, S.; Kane, L. P. Galectin-9 regulates T helper cell function independently of Tim-3. *Glycobiology*. **2011,** *21*, 1258–1265.
62. Oomizu, S.; Arikawa, T.; Niki, T.; Kadowaki, T.; Ueno, M.; Nishi, N.; Yamauchi, A.; Hirashima, M. Galectin-9 suppresses Th17 cell development in an IL-2-dependent but Tim-3-independent manner. *Clin. Immunol.* **2012,** *143*, 51–58.

63. Vaitaitis, G. M.; Wagner, D. H. Galectin-9 controls CD40 signaling through a Tim-3 independent mechanism and redirects the cytokine profile of pathogenic T cells in autoimmunity. *PLoS One*. **2012,** *7*, e38708.
64. Golden-Mason, L.; McMahan, R. H.; Strong, M. Galectin-9 functionally impairs natural killer cells in humans and mice. *J. Virol.* **2013,** *87*, 4835-4845.
65. Leitner, J.; Rieger, A.; Pickl, W. F.; Zlabinger, G. TIM-3 does not act as a receptor for galectin-9. *PLoS Pathogens.* **2013,** *9*, e1003253.
66. Nakayama, M.; Akiba, H.; Takeda, K.; Kojima, Y.; Hashiguchi, M.; Azuma, M.; Yagita, H.; Okumura, K. Tim-3 mediates phagocytosis of apoptotic cells and cross-presentation. *Blood.* **2009,** *113*, 3821-3830.
67. Hanayama, R.; Tanaka, M.; Miwa, K.; Shinohara, A.; Iwamatsu, A.; Nagata, S. Identification of a factor that links apoptotic cells to phagocytes. *Nature*. **2002,** *417*, 182–187.
68. Brissette, M. J.; Lepage, S.; Lamonde, A. S.; Sirois, I.; Groleau, J.; Laurein, L.-P.; Cailhier, J.-F. MFG-E8 released by apoptotic endothelial cells triggers anti-inflammatory macrophage reprogramming. *PLoS One.* **2012,** *7*, e36368.
69. Wang P. MFG-E8 and Inflammation. Springer Nature, Dordrecht, 2014.
70. Andersen, M. H.; Berglund, L.; Rasmussen, J. T.; Petersen, T. E. Bovine PAS-6/7 binds $\alpha_v\beta_5$ integrin and anionic phospholipids through two domains. *Biochemistry*. **1997,** *36*, 5441-5446.
71. Andersen, M.; Graversen, H.; Fedosov, S.; Petersen, T.; Rasmussen, J. Functional analyses of two cellular binding domains of bovine lactadherin. *Biochemistry*. **2000,** *39*, 6200-6206.
72. Taylor, M.; Couto, J.; Scallan, C.; Ceriani, R.; Peterson, J. A. Lactadherin (formerly BA46), a membrane-associated glycoprotein expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp (RGD)-dependent cell adhesion. *DNA Cell Biol*. **1997,** *16*, 861-869.
73. Borisenko, G. G.; Iverson, S. L.; Ahlberg, S.; Kagan, V. E.; Fadeel, B. Milk fat globule epidermal growth factor 8 (MFG-E8) binds to oxidized phosphatidylserine: implications for macrophage clearance of apoptotic cells. *Cell Death Differ.* **2004,** *11*, 943-945.
74. Dasgupta, S. K.; Abdel-Monem, H.; Niravath, P.; Le, A.; Bellera, R. V.; Langlous, K.; Nagata, S.; Rumbaut, R. E.; Thiagarajan, P. Lactadherin and clearance of platelet-derived microvesicles. *Blood*. **2009,** *113*, 1332-1339.

75. Ruggiero, L.; Connor, M. P.; Chen, J.; Langen, R.; Finnemann, S. C. Diurnal, localized exposure of phosphatidylserine by rod outer segment tips in wild-type but not Itgb5−/− or Mfge8−/− mouse retina. *Proc. Natl. Acad. Sci.* **2012,** *109*, 8145-8148.
76. Nandrot, E. F.; Anand, M.; Almeida, D.; Atabai, K.; Sheppard, D.; Finnemann, S. C. Essential role for MFG-E8 as ligand for alphavbeta5 integrin in diurnal retinal phagocytosis. *Proc. Natl. Acad. Sci.* **2007,** *104*, 12005–12010.
77. Khalifeh-Soltani, A.; McKleroy, W.; Sakuma, S.; Cheung, Y. Y.; Tharp, K.; Qiu, Y.; Turner, S. M.; Chawla, A.; Stahl, A.; Atabai, K. Mfge8 promotes obesity by mediating the uptake of dietary fats and serum fatty acids. *Nat. Med.* **2014**, *20*, 175-183.
78. Ait-Oufella, H.; Kinugawa, K.; Zoll, J.; Simon, T.; Boddaert, J.; Heeneman, S.; Blanc-Brude, O.; Barateau, V.; Potteaux, S.; Merval, R.; Esposito, B.; Teissier, E.; Daemen, M. J., Leseche, G.; Boulanger, C.; Tedgui, A.; Mallat, Z. Lactadherin deficiency leads to apoptotic cell accumulation and accelerated atherosclerosis in mice. *Circulation.* **2007,** *115*, 2168-2177.
79. Hanayama, R.; Tanaka, M.; Miyasaka, K.; Aozasa, K.; Koike, M.; Uchiyama, Y.; Nagata, S. Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-deficient mice. *Science*. **2004,** *304*, 1147–1150.
80. Boddaert, J.; Kinugawa, K.; Lambert, J. C.; Boukhtouche, F.; Zoll, J.; Merval, R.; Blanc-Brude, O.; mann, D.; Berr, C.; Vilar, J.; Garabedian, B.; Journiac, N.; Charue, D.; Silvestre, J. S.; Duyckaerts, C.; Amouyel, P.; Mariani, J.; Tedgui, A.; Mallat, Z. Evidence of a role for lactadherin in Alzheimer's disease. *Amer. J. Pathol.* **2007,** *170*, 921-929.
81. Kinugawa, K.; Monnet, Y.; Lu, L.; Bekaert, A. J.; Théry, C.; Mallat, Z.; Hirsch, E. C.; Hunot, S. MFGE8 does not orchestrate clearance of apoptotic neurons in a mouse model of Parkinson's disease. *Neurobiol. Dis.* **2013,** *51*, 192-201.
82. Fricker, M.; Neher, J. J.; Zhao, J. W.; Théry, C.; Tolkovsky, A. M.; Brown, G. C. MFG-E8 mediates primary phagocytosis of viable neurons during neuroinflammation. *J. Neurosci.* **2012,** *32*, 2657-2666.
83. Neniskyte, U.; Brown, G. Lactadherin/MFG-E8 is essential for microglia-mediated neuronal loss and phagoptosis induced by amyloid β. *J. Neurochem*. **2013,** *126*, 312–317.
84. Das, A.; Ghatak, S.; Sinha, M.; Chaffee, S.; Ahmed, N. S.; Parinandi, N. L.; Wohleb, E. S.; Sheridan, J. F.; Sen, C. K.; Roy, S. Correction of MFG-E8 resolves inflammation and promotes cutaneous wound healing in diabetes. *J. Immunol.* **2016,** *196*, 5089–5100.

85. Sugano, G.; Bernard-Pierrot, I.; Lae, M.; Battail, C.; Allory, Y.; Stransky, N.; Krumeich, S.; Lepage, M. L.; Maille, P.; Donnadieu, M. H.; Abbou, C. C.; Benhamou, S.; Lebret, T.; Sastre-Garau, X.; Amigorena, S.; Radvanyi, F.; Thery, C. Milk fat globule—epidermal growth factor—factor VIII (MFGE8)/lactadherin promotes bladder tumor development. *Oncogene*. **2011,** *30*, 642-653.
86. Neutzner, M.; Lopez, T.; Feng, X.; Bergmann-Leitner, E.; Leitner, W.; Udey, M. MFG-E8/lactadherin promotes tumor growth in an angiogenesis-dependent transgenic mouse model of multistage carcinogenesis. *Cancer Res*. **2007,** *67*, 6777–6785.
87. Shi, J.; Shi, Y.; Waehrens, L. N.; Rasmussen, J. T.; Heegaard, C. W.; Gilbert, G. E. Lactadherin detects early phosphatidylserine exposure on immortalized leukemia cells undergoing programmed cell death. *Cytometry A*. **2006,** *69*, 1193-1201.
88. Johnson, B. L.; Midura, E. F.; Prakash, P. S.; Rice, T. C.; Kunz, N.; Kalies, K.; Caldwell, C. C. Neutrophil derived microparticles increase mortality and the counter-inflammatory response in a murine model of sepsis. *Biochim. Biophys. Acta*. **2017.**
89. Cen, C.; Aziz, M.; Yang, W.-L.; Zhou, M.; Nicastro, J. M.; Coppa, G. F.; Wang, P. Milk fat globule-epidermal growth factor-factor VIII attenuates sepsis-induced acute kidney injury. *J. Surg. Res.* **2017,** *213*, 281-289.
90. Hendricks, L.; Aziz, M.; Yang, W.-L.; Nicastro, J.; Coppa, G. F.; Symons, M.; Wang, P. Milk fat globule-epidermal growth factor-factor VIII–derived peptide MSP68 is a cytoskeletal immunomodulator of neutrophils that inhibits Rac1. *J. Surg. Res.* **2017,** *208*, 10-19.
91. Wang, Y.; Luo, G.; Chen, J.; Jiang, R.; Zhu, J.; Hu, N.; Huang, W.; Cheng, G.; Jia, M.; Su, B.; Zhang, N.; Cui, T. Cigarette smoke attenuates phagocytic ability of macrophages through down-regulating Milk fat globule-EGF factor 8 (MFG-E8) expressions. *Sci. Rep.* **2017,** *7*, 42642.
92. Forbes, S. J. Milk fat globule-EGF factor 8 for liver fibrosis therapy: creaming off the beneficial effects of mesenchymal stromal cells. *Gastroenterology.* **2017,** *152*, 943-946.
93. Dasgupta, S. K.; Guchhait, P.; Thiagarajan, P. Lactadherin binding and phosphatidylserine expression on cell surface-comparison with annexin A5. *Transl Res.* **2006,** *148*, 19-25.
94. Shi, J.; Gilbert, G.E. Lactadherin inhibits enzyme complexes of blood coagulation by competing for phospholipid-binding sites. *Blood*. **2003,** *101*, 2628–2636.

95. Shi, J.; Heegaard, C. W.; Rasmussen, J. T.; Gilbert, G. E.. Lactadherin binds selectively to membranes containing phosphatidyl-L-serine and increased curvature. *Biochim. Biophys. Acta*. **2004,** *1667*, 82-90.
96. Lin, L.; Huai, Q.; Huang, M.; Furie, B.; Furie, B. C. Crystal structure of the bovine lactadherin C2 domain, a membrane binding motif, shows similarity to the C2 domains of factor V and factor VIII. *J. Mol. Biol.* **2007,** *371*, 717-724.
97. Shao, C.; Novakovic, V. A.; Head, J. F.; Seaton, B. A.; Gilbert, G. E. Crystal structure of lactadherin C2 domain at 1.7 Å resolution with mutational and computational analyses of its membrane-binding motif. *J. Biol. Chem*. **2008,** *283*, 7230–7241.
98. Nanga, R. P.; Vivekanandan, S.; Yoon, H. S. Expression, purification and characterization of C2 domain of milk fat globule-EGF-factor 8-L. *Protein Expr. Purif*. **2007,** *52*, 329-333.

CHAPTER 3

METHODOLOGY

A diverse set of physicochemical/biophysical tools have been employed in this thesis. To build the protein-membrane binding configuration for Tim3 and MFG-E8, molecular dynamics (MD) simulation and x-ray reflectivity methods were employed together. To study the binding free energy of proteins with various lipid systems, a binding assay based on spectral shift of the intrinsic fluorescence of the Tim's and MFG-E8 was used. To analyze the calcium distribution at the surface of lipid membranes in the context of PS-Tim proteins interaction, x-ray fluorescence experiments were performed and an x-ray fluorescence software was developed. The underlying principles, data analytical methods, and experimental procedures, for these techniques are treated here, with the exception of x-ray fluorescence, which exists in a separate chapter (Chapter 6).

3.1 Molecular Dynamics Simulations

While x-ray resolved crystal structures and nuclear magnetic resonance (NMR) resolved structures are invaluable for guiding scientists connecting observed properties of protein molecules with their structural motifs, the static nature of these structures often obscure important aspects of such interactions in a dynamic world. In addition, resolving the structure of a peripheral binding protein such as Tim3 and MFG-E8 adsorbed to the membrane is ill-suited for either x-ray crystallography or NMR techniques, due, respectively, to the infeasibility of crystalizing a lipid membrane with the protein, and the slow gyration of protein molecules partially immobilized by a lipid membrane surface. Using the crystal or NMR structure as a

starting point for MD simulations, however, information on the detailed dynamics of the interaction could be obtained. In addition, MD simulations were used to provide an ensemble average structure for fitting the x-ray reflectivity data (Section 3.2). An exciting new development in MD simulations involving lipid membrane systems was the development of the highly mobile membrane mimetic (HMMM) method in 2012 [1], which significantly speeds up the lipid dynamics by shortening the lipid acyl chains while filling the evacuated space in the hydrophobic core of the membrane with organic solvents (**Figure 3.1**). This method made possible the simulation of membrane binding events in a reasonable computational time frame.

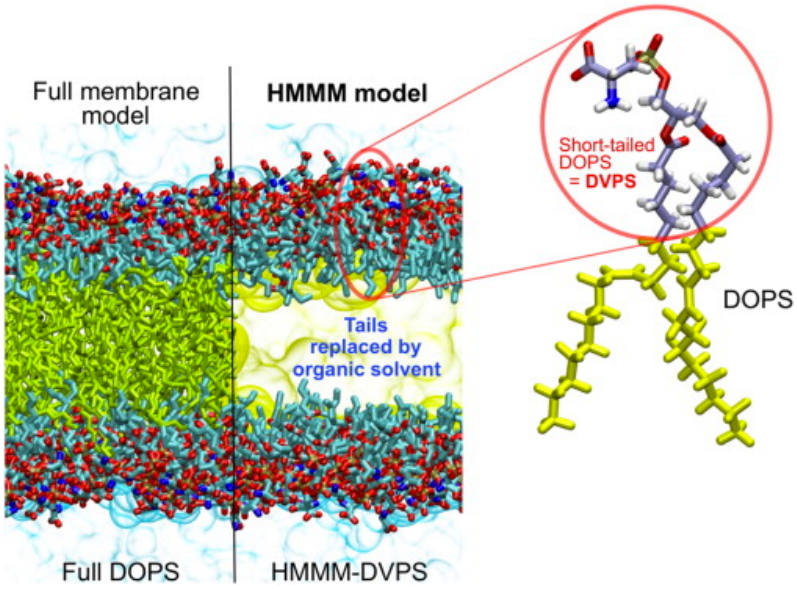

Figure 3.1: Contrasting the HMMM and conventional representations of a membrane. In an HMMM model, a large fraction of the acyl tails (left, yellow sticks) of the membrane-forming lipids is replaced by a liquid organic phase (right, yellow area). In this study, a full DOPS lipid molecule (inset) is represented by a short-tailed DVPS molecule with only 5 carbons (the circled fraction of the molecule), and the space vacated by the removal of the C6–C18portion of the acyl tails is filled with an organic solvent (yellow). Oxygen atoms are red, nitrogen blue, phosphorus gold, and carbon ice blue, except for C6–C18, which are in yellow, representing the organic solvent. Bulk water molecules are shown in light blue. (reprint from [1], *permission granted by publisher*).

MD simulations were performed using NAMD 2.9 for Tim3, and NAMD 2.11 for MFG-E8 [2], with the CHARMM27 force field for Tim3, and the CHARMM36 force field for MFG-E8 [3], using the HMMM method [1]. Parallelization was generally over 640 cores on the Beagle Supercomputer at the University of Chicago. An asymmetric membrane bilayer was used, with 7:3 POPC:POPS on the leaflet facing the protein, while neat POPC on the other leaflet away from the protein. HMMM input files were generated using the CHARMM-GUI HMMM protein/membrane builder [4, 5]. The protein docking events were simulated with at 310 K (36.85 ℃) with the Tim3 Ig domain 10 nm away, or the MFG-E8 C2 domain 20 nm away from the membrane. Each trajectory of the short tail lipid system was simulated for 180 to 200 ns, at which point the proteins were bound to the membrane, and equilibration had reached judging by the stabilization of the root-mean-square deviation (RMSD) of the superimposed protein structure. One to three extended tail systems were generated from each short tail system using the CHARMM-GUI HMMM protein/membrane builder, and simulated for another 200 ns. MD trajectory processing was done using VMD [6] and the MDToolBox [7]. Sequence alignment was performed using the MultiSeq plugin for VMD [8]. Image rendering in VMD was performed using Tachyon [9]. Custom MATLAB code was used to perform RMSD analysis.

To locate the optimal protein structure from the MD simulations for x-ray reflectivity (XR) fitting, first the protein structure from all the extended tail frames were overlaid in the plane horizontal to the membrane (x-y plane), then a mean structure of all of these overlaid frames. This 2D overlay was performed using the "superimpose2d" function of the MDToolBox [7]. The protein structures were superimposed in the x-y plane, because XR is only sensitive to the electron density along the axis normal to the membrane surface (z-axis), thus variations in 2D

should be eliminated. This mean structure represents an ensemble mean, yet was unrealistic due to overlapping atomic positions produced in the averaging process. Therefore, the frame from all extended tail trajectories that minimized RMSD with the mean structure was located and used for fitting XR data. This new protein structure was referred to as MD-optimized, and it was significantly different from the NMR structure in the case of the MFG-E8 C2 domain (**Figure 3.2a**). In comparison to the protein structure in other frames, this MD-optimized structure gives an overall smaller RMSD for the full course of the simulation, as seen in **Figure 3.2b**.

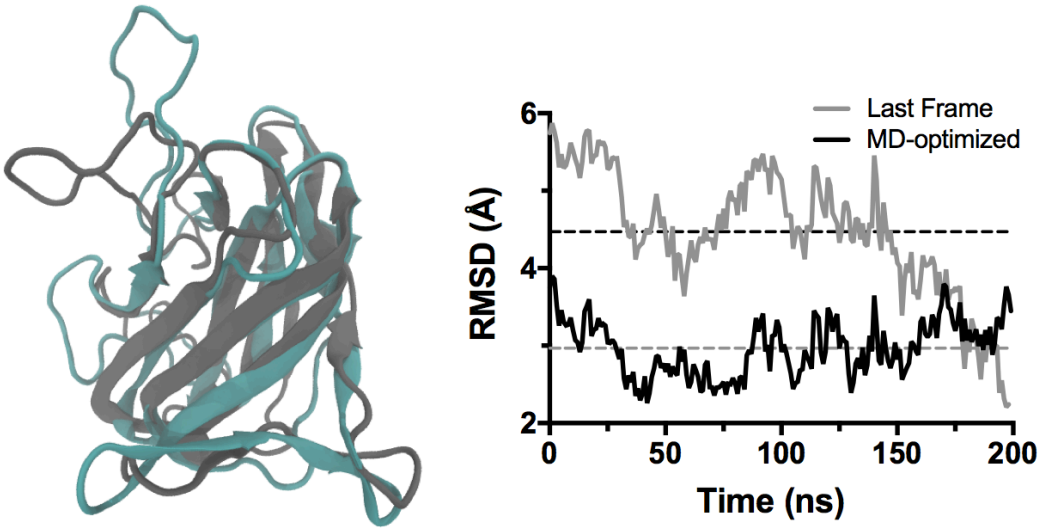

Figure 3.2: Characteristics of the MD-optimized structure. (a): Aligning of the C2 domain NMR structure with the membrane MD equilibrated bound structure. The teal colored one is the MD-optimized structure, while the gray colored one is the NMR structure. The backbone RMSD between these two structures are calculated to be 5.28 Å. (b): Backbone RMSD using either the MD-optimized structure or the last frame as the reference. Here the last frame is used as an example to show that the MD-optimized better capture the average structure of the protein with lower overall RMSD. The horizontal lines indicate the average RMSD for frames sampled every 1 ns, or 50 frames. When referencing MD-optimized structure, the averaged RMSD was 2.97 Å, and when referencing the last frame, the averaged RMSD was 4.47 Å.

3.2 X-ray Reflectivity (XR)

The theory of x-ray reflectivity is thoroughly treated in this reference [10]. Briefly, when highly monochromatic x-ray encounters an air/liquid interface, reflection and refraction happens at the interface largely observing classical optics rules. Because x-rays scatter off electrons and the wavelengths of x-ray are in the Angstrom (Å) range, it is sensitive to electron density changes at the molecular level. The specular reflection of x-rays, described in **Figure 3.3**, represents a momentum transfer in the direction normal to the interface (z-axis), and thus is sensitive to the arrangements of the surface at the molecular detail. As the refraction index of water is smaller than 1, there exists a critical angle below which total reflection happens. By convention, the momentum transfer in the z-direction is defined as

$$Q_z = \frac{4\pi}{\lambda} \sin\theta,$$

where θ is the incidence angle, and λ is the wavelength of the x-ray, and the critical angle is represented correspondingly by Q_c. The intensity of reflectivity depends on both the incidence angle and the electron density profile. For perfectly sharp interfaces, the reflectivity intensity is referred to as the Fresnel reflectivity, R_f, as calculated by

$$R_f = \left| \frac{Q_z - \sqrt{Q_z^2 - Q_c^2}}{Q_z + \sqrt{Q_z^2 - Q_c^2}} \right|^2,$$

Assuming a homogeneous surface with only electron density variations in the z-direction, the reflectivity intensity, *R*, can be expressed as the square of the Fourier transform of the electron density gradient multiplied by the bulk Fresnel reflectivity:

$$R(Q_z) = R_f \left| \frac{1}{\rho(z)} \int \frac{d\rho(z)}{dz} e^{iQ_z z} dz \right|^2,$$

where $\rho(z)$ is the electron density at position z. In computational implementation of the above equation, the electron density profile is discretized, and treated as small slabs.

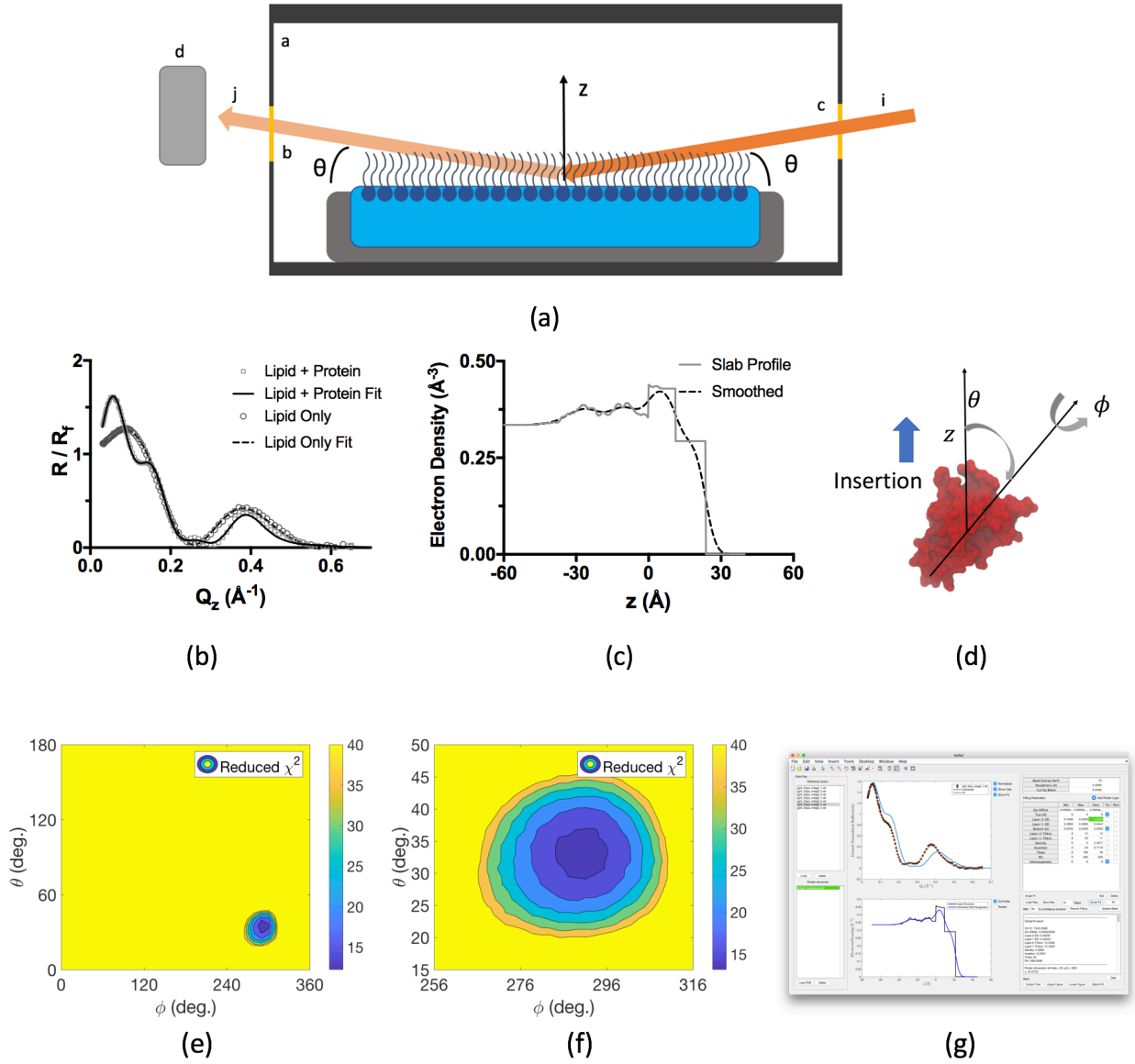

Figure 3.3: X-ray reflectivity methods. (a): Geometry of x-ray reflectivity. In this figure, a – the Helium filled enclosure; b, c – Kapton windows; i – incidence beam; j – reflected beam; θ – incidence angle and reflection angle; d – detector. (b): Sample XR generated by lipid only and lipid + protein systems. (c): The electron density with protein adsorbed to the monolayer. (d): Definition of the protein orientation. θ is the angle of rotation angle about the x-axis, and ϕ is the angle of rotation about the z-axis. (e): A representative reduced χ^2 map over the two-dimensional (θ, ϕ) grid. (f): XR fitting GUI.

In application to lipid monolayer studies, the lipid monolayer is generally modeled with a two-box model, assuming a uniform electron density for the headgroup region, and another for the tailgroup region. The whole electron density profile is then smeared by convoluting with the thermal fluctuations arising from the capillary waves. If proteins molecules were adsorbed to the membrane, the added electron density would generate much different XR patterns (**Figure 3.3b-c**).

X-ray crystallography and NMR methods are generally not applicable to solving membrane bound protein structures. A recent study using a slab model for representing static protein structures analysis XR data from a lipid adsorbed protein monolayer has provided a much-needed experimental method for resolving the orientation and insertion depth of membrane bound protein [11]. In this method, protein binding is defined by four parameters, namely the angle of rotation, θ, about the principal x-axis, the angle of rotation, ϕ, about the principal z-axis (**Figure 3.3d**), the insertion depth of the domain, and the surface number density of the protein. The use of surface number density is a modification from the original area fraction, and is more representative of the protein concentration at the surface. There are also four lipid parameters, namely the lengths and electron densities of the tailgroup and the headgroup. These four protein parameters and four lipid parameters, together with a *Qz* offset value that accounts for systematic deviations of the instrument form the parameter space determining the reflectivity pattern. As the search algorithm over these two angles, θ and ϕ, are not efficient at the moment, the preferred way to fit XR data is to first step through a grid of (θ, ϕ) pairs and find the minimum sum of residual squares (χ^2) value, by use of a well implemented search algorithm, such as the trust

region reflective method as adopted here. Then zoom in for finer steps in the around possible reduced χ^2 minima (**Figure 3.3e-f**). In a typical scan, the number of points, N, is around 100, and the obtained minimal χ^2 values are general between 400 to 2000. A ration χ^2/N larger than one generally indicates deviation of the model from physical reality. As the χ^2/N ratio is generally larger than 4 but smaller than 20, the underlying slab model and the static protein structure captured the main features of the XR data, but not the whole physical reality. These algorithms have been custom built and integrated in a graphical user interface tool for quicker fitting by the author (**Figure 3.3g**).

These XR experiments were performed at ChemMatCARS of Sector 15ID at the Advanced Photon Source of Argonne National Laboratory. A custom trough with a movable barrier and surface pressure control loop was used to control the surface pressure of the lipid monolayer. Surface pressure was controlled at three values, 20, 25, and 30 mN/m. For each film, a reflectivity scan on a lipid-only system was performed prior to injecting the protein samples to the subphase. Following the injection, an equal amount of buffer was removed from the subphase to restore the initial liquid level and 2 hours of equilibration time was allowed. Then, XR scans were taken within a Q_z range of 0.018 – 0.55 or 0.65 $Å^{-1}$. Two curves were gathered on a given film, which were generally quite similar, and averaged. In all of XR experiments, phospholipids with 1-stearoyl-2-oleoyl (SO) tails were used to minimize radiation damage [11].

3.3 Binding Assay

A binding assay based on the intrinsic tryptophan fluorescence of the Tim proteins and MFG-E8 was used with variations. When tryptophan residues move from the more polar water environment into the membrane headgroup/tailgroup interfacial region, its fluorescence spectrum displays a blue shift and an increase in intensity, which could be utilized to measure the bound fraction of the burying entity [12]. As the Tim proteins and MFG-E8 all possess tryptophan residues that help anchor the protein into the membrane, this tryptophan spectral shift provides a nonintrusive way to monitor the interaction between lipids and the four proteins. A Horiba Fluorolog-3 spectrophotometer with a 1 cm path-length quartz cuvette was used. For maximal excitation efficiency, the excitation wavelength was set at 280 nm. The collection range is generally set at 300 to 420 nm to capture the full emission peak. **Figure 3.4a** shows the difference in fluorescence spectra between the free and the C2 domain of MFG-E8.

In all binding assays, large unilamellar vesicles (LUV) were used. Uniformly sized LUVs were produced by extrusion [13]. Briefly, lipids in the powder form were dissolved in chloroform, mixed to the desired concentration and composition, then dried to form a thin film inside the glass vials under nitrogen flow and constant rotation by hand. The lipids were then further dried under vacuum for 1 h. A buffer of 150 mM NaCl plus 10 mM Hepes at pH 7.5 was added to the lipids and vortexed at a moderate frequency for 1 h at 40 °C. The lipid/buffer mixtures were then subjected to five freeze-thaw cycles, using a dry ice/ethanol slurry. Finally, the lipids were extruded 21 times using 100-nm-diameter filters (Avanti Polar Lipids). Vesicle diameter and polydispersity were around 130 nm and 10% as characterized by dynamic light scattering (Zen3600 Malvern Nano Zetasizer) for each batch.

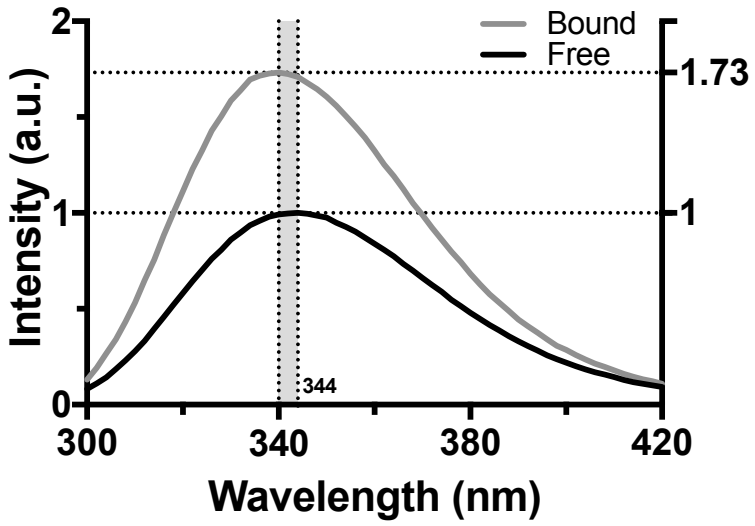

Figure 3.4: Difference in tryptophan fluorescence spectra between the bound and free protein. These are data from the C2 domain of MFG-E8, showing a blueshift of 4 nm, and an increase in quantum yield of 73%.

The binding assay was generally performed in a titration series, for better characterization of the effects of a range of the titrants. The varying parameters could be lipid surface density holding total lipid concentration constant, Ca^{2+} concentration, or total lipid concentration. For each titration, first the unbound spectrum, $U(\lambda)$, was established without the titrant for at least three repeats. Then the titrant was added and the spectra, $X(\lambda)$, became a linear combination of the bound and free protein spectra (**Figure 3.5a**). As the further addition of titrant did not further increase the intensity and cause the blueshift, the tentative bound spectrum $B(\lambda)$ was established as the last one in the series. Spectra with mixed bound and unbound components, $X(\lambda)$, was decomposed into a linear combination of the bound and unbound spectra:

$$X(\lambda) = b \cdot B(\lambda) + (1 - b) \cdot U(\lambda), \qquad 3.4$$

where b is the bound fraction. The cuvette was thoroughly cleaned between samples, and spectra were collected after 2 minutes of equilibration. For each series, the bound fractions and the dissociation constant were estimated by curve fitting, assuming specific binding to a single-site

(**Figure 3.5b**), and a bmax value, which stands for the real bound fraction of the tentative bound spectrum, using this equation:

$$Bound\ Fraction = bmax\frac{[titrant]}{[titrant]+K_d}, \quad 3.5$$

Or if a Hill model (**Figure 3.5c**) needs to be used, data will be fit to:

$$Bound\ Fraction = bmax\frac{[titrant]^h}{[titrant]^h+K_d}, \quad 3.6$$

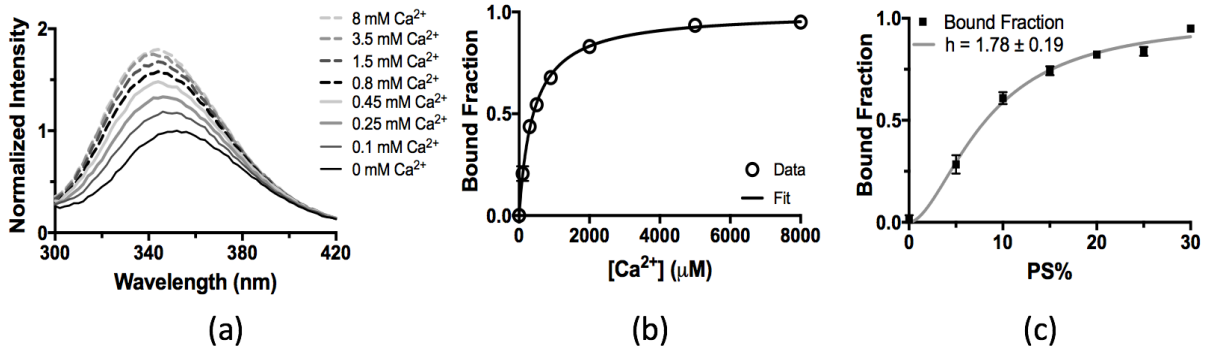

Figure 3.5: Binding titration series and curve fitting. (a): Progression of the spectrum from free (lowest intensity) to bound (highest intensity). In this example, a Ca^{2+} titration series was performed for studying Tim3 sensitivity to Ca^{2+}. The spectra in the middle could be expressed as a linear combination of the free spectrum and the bound spectrum. (b): Single site binding produces a hyperbolic saturation curve. (c): Multiple site binding fits to a Hill model. In this case, the titrant is the surface density, and the larger than 1 Hill coefficient is indicative of a larger than one stoichiometry between the protein and lipid.

In two typical single-site binding studies, a Ca^{2+} titration is performed for the Tim proteins, and a vesicle titration for MFG-E8, as MFG-E8 binds PS in a calcium-independent way. As the concentration of the titrant increases, the protein spectra became a mixture of bound and unbound.

3.4 Treatment of the Ca^{2+} dependent binding of the Tim proteins to PS containing vesicles

As the interaction between the Tim proteins and vesicles are a three-body interaction, involving the Ca^{2+} ion, the vesicle, and the protein, the use of the dissociation constant from a Ca^{2+} titration experiment to represent the binding strength of the reaction needs justification. Assuming only one Ca^{2+} is needed for the interaction, the binding equilibrium can be described by

$$\mathrm{P} + \mathrm{Ca} + \mathrm{V} \rightleftharpoons \mathrm{PCaV},$$

with dissociation constant

$$\mathrm{K_d} = \frac{[\mathrm{P}][\mathrm{Ca}][\mathrm{V}]}{[\mathrm{PCaV}]},$$

where P stands for the protein, and V stands for vesicle. In all the experiments, protein concentration is kept at 170 nM, while lipid concentration is kept at 300 μM, and Ca^{2+} concentration ranges from 4 μM to 8 mM. At the lowest Ca^{2+} concentration of 4 μM, where the bound Ca^{2+} to free Ca^{2+} ratio is the largest, at most 11.5% of the protein is bound, as in the Tim4 case. Therefore, the protein would only bind <20 nM Ca^{2+} so that the free Ca^{2+} concentration is close to that of the total:

$$[\mathrm{Ca}] \approx [\mathrm{Ca}]_\mathrm{t},$$

where $[\mathrm{Ca}]$ is the free Ca^{2+} concentration and $[\mathrm{Ca}]_\mathrm{t}$ is the total Ca^{2+} concentration.

At the lowest PS content (2 mol%), where the ratio of Tim3-bound to free PS is the largest, the total PS concentration in the outer leaflet would be $300 \times 0.05 \times 0.2 = 3.75$ μM. When the protein is fully bound, 170 nM PS would be associated with its central binding pocket. Therefore, the highest bound PS to free PS ratio is around 4.6%. The concentration of unbound PS would then be very close to its total. Therefore,

$$[\mathrm{V}] \approx [\mathrm{V}]_\mathrm{t},$$

where $[V]$ is the free vesicle PS concentration, while $[V]_t$ is the total vesicle PS concentration. Therefore, the dissociation constant

$$K_d \approx \frac{[P][Ca]_t[V]_t}{[PCaV]},$$

We define Kd(Ca) as

$$K_d(Ca) = \frac{K_d}{[V]_t},$$

or

$$K_d(Ca) = \frac{[P][Ca]_t}{[PCaV]},$$

and the bound fraction of protein is

$$b = \frac{[Ca]}{K_d(Ca)+[Ca]},$$

which is a typical one-site specific binding equation, and $K_d(Ca)$ is the Ca^{2+} concentration where half of the protein is bound. Therefore, the global K_d differs with $K_d(Ca)$ only by a constant multiplier, and the use of $K_d(Ca)$ to represent the global binding strength is justified.

3.5 References

1. Ohkubo, Y. Z.; Pogorelov, T. V.; Arcario, M. J.; Christensen, G. A.; Tajkhorshid. E. Accelerating membrane insertion of peripheral proteins with a novel membrane mimetic model. *Biophys. J.* **2012,** *102*, 2130-2139.
2. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem*. **2005,** *26*, 1781–1802.
3. Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Leumkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pane, V. S.; Case, D. A.; Brooks, C. L. III;

MacKerell, A. D. Jr.; Klauda, J. B.; Im, W. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **2016,** *12*, 405-413.

4. Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem*. **2008,** *29*, 1859–1865.
5. Qi, Y.; Cheng, X.; Lee, J.; Vermaas, J. V.; Pogorelov, T. V.; Tajkhorshid, E.; Park, S.; Klauda, J. B.; Im, W. CHARMM-GUI HMMM builder for membrane simulations with the highly mobile membrane-mimetic model. *Biophys. J.* **2015,** *109*, 2012-2022.
6. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Molec. Graphics*. **1996,** *14*, 34-38.
7. *Github.* https://github.com/ymatsunaga/mdtoolbox (Accessed April 2017)
8. Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*. **2006,** *7*, 382.
9. Tachyon Parallel/Multiprocessor Ray Tracing System. http://www.photonlimited.com/~johns/tachyon/ (Accessed April 2017)
10. Als-Nielsen, J.; Jacquemain, D.; Kjaer, K.; Leveiller, F.; Lahav, M.; Leiserowitz, L. Principles and Applications of Grazing Incidence X-Ray and Neutron Scattering from Ordered Molecular Monolayers at the Air-Water Interface. *Physics Reports* **1994**,*246*, 251313.
11. Málková, Š.; Long, F.; Stahelin, R.; Pingali, S.; Murray, D.; Cho, W.; Schlossman, M. X-Ray Reflectivity Studies of cPLA2α-C2 Domains Adsorbed onto Langmuir Monolayers of SOPC. *Biophys. J.* **2005,** *89*, 1861-1873.
12. Kraft, C.; Garrido, J.; Leiva-Vega, L.; Romero, G. Quantitative Analysis of Protein-Lipid Interactions Using Tryptophan Fluorescence. *Sci. Signal.* **2009,** *2*, p14.
13. MacDonald, R. C.; MacDonald, R. I.; Menco, B. M.; Takeshita, K.; Subbarao, N. K.; Hu, L. Small-volume extrusion apparatus for preparation of large, unilamellar vesicles. *Biochim. Biophys. Acta*. **1991,** *1061*, 297–303.

CHAPTER 4

PHYSICOCHEMICAL SELECTIVITY OF TIM3 TO PHOSPHATIDYLSERINE-CONTAINING LIPID MEMBRANES

4.1 Overview

T cell immunoglobulin mucin 3 (Tim3) is a cell surface protein that serves as an immune checkpoint receptor. It is an emerging target for cancer immunotherapy. Tim3 contains a calcium-dependent phosphatidylserine (PS) binding site on its immunoglobulin (Ig) domain critical for its intercellular interactions, as with its homologs, Tim1 and Tim4. Here, by use of a tryptophan fluorescence spectral shift assay, we quantified the binding affinity of the Ig domains of Tim1, Tim3, and Tim4 to find Tim3 affinity to be much lower than Tim1 and Tim4, consistent with previous observations [1, 2]. Nevertheless, membrane binding by Tim3 is indeed Ca^{2+}-mediated and selective to the PS headgroup. We built a molecular-level picture of its binding based on analysis of x-ray reflectivity of the Tim3 Ig domain adsorbing to lipid monolayers to elucidate potential modulators for Tim3 binding. We found Tim3 binding to be highly enhanced by decreased membrane packing density and increased negative charge density. These two sensitivity factors originate from four modes of Tim3/bilayer interaction (i.e., calcium-dependent, electrostatic, hydrophobic, and mechanical), giving Tim3 its unique physicochemical selectivity to differentiate itself from other PS-binding proteins. The characteristic interactions of Tim3 with PS and PA may tune Tim3 to specifically respond to cellular signals as in neutrophil activation and its subsequent apoptosis.

4.2 Computational and Experimental Study Results

4.2.1 Tim3 binding to PS-containing lipid membranes

Earlier studies suggested that the binding of Tim3 to PS is weaker than that of the Tim1 and Tim4 proteins [1, 2]. However, an accurate quantitative comparison was lacking. Using the tryptophan fluorescence spectral shift assay [2, 3], we determined the binding affinity of the Ig domains of the three TIM proteins for large unilamellar vesicles composed of 70% 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 30% 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS), henceforth denoted as 7:3 POPC:POPS. We determined the Ca^{2+} dependence of protein binding to estimate the dissociation constant, $K_d(Ca^{2+})$ at which half of the protein bound. The use of $K_d(Ca^{2+})$ as a metric for the overall binding affinity is justified in the **Section 4.3.2**. As seen in **Figure 4.1a**, a hyperbolic binding curve was obtained, strongly suggesting that all three TIMs associate with the membrane with a 1:1 Ca^{2+} stoichiometry. The dissociation constant for Tim3 and 30% POPS vesicles was five times larger than the $K_d(Ca^{2+})$ for Tim1 and about 15 times larger than that for Tim4, which translates to a difference in binding free energy of 3.96 kJ/mol for Tim1 and -6.39 kJ/mol for Tim4.

4.2.2 Tim3 PS binding site is specific to the PS headgroup and Ca^{2+}

To further understand the nature of the PS binding site on Tim3, we tested the specificity of the binding site to Ca^{2+} and the PS headgroup. To test its lipid headgroup specificity, we substituted the negatively charged PS headgroup with the negatively charged PA or phosphatidylglycerol

(PG) headgroups, so we used vesicles with composition 7:3 POPC:POPA or 7:3 POPC:POPG. In both cases, no measurable affinity was observed. To test the specificity of the binding site to Ca^{2+}, we studied the dependence of the binding affinity on Mg^{2+} and Ba^{2+} and found that neither ion was able to facilitate Tim3 binding to 7:3 POPC:POPS. These results are summarized in **Table 4.1**.

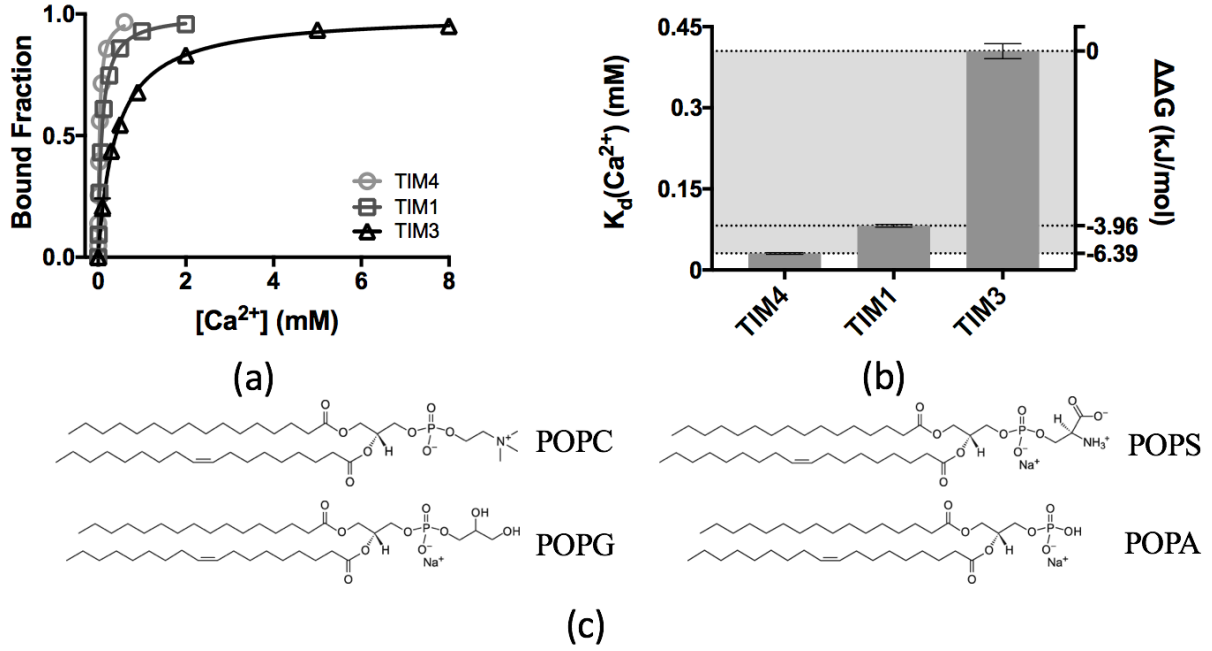

Figure 4.1: Tim1, Tim3, and Tim4 affinities for PS. (**a**): Ca^{2+} dependence of Tim1, Tim3, and Tim4. The protein concentration was 170 nM protein and the lipid concentration was 300 µM lipid with 7:3 POPC:POPS. (**b**): $K_d(Ca^{2+})$ and $\Delta\Delta G$ for the three TIM proteins. The average $K_d(Ca^{2+})$ values for Tim1, Tim3, and Tim4 were 81.9±2.4, 405±14, and 30.7±1.4 µM, respectively, as seen from the left vertical axis. Using Tim3 binding free energy as the reference, the $\Delta\Delta G$ of Tim1 and Tim4 are -3.96 and -6.39 kJ/mol, respectively, as seen from the right vertical axis. (**c**): Structure of PO lipids, POPC, POPS, POPG, and POPA.

Tim3 Binding Affinity	POPC : POPS = 7 : 3	POPC : POPG = 7 : 3	POPC : POPA = 7 : 3
Ca^{2+}	Measurable	Negligible	Negligible
Mg^{2+}	Negligible	N/A	N/A
Ba^{2+}	Negligible	N/A	N/A

Table 4.1: Tim3 binding site is specific to PS and Ca^{2+}.

4.2.3 Tim3 binding geometry reveals insertion of two hydrophobic residues into the membrane and positioning of four basic residues in close proximity to the membrane

As Tim3 affinity to PS-containing membrane is much weaker than Tim1 and Tim4, we tried to explore additional modulators for the interaction. To guide such exploration, we first built a molecular interaction picture of Tim3 with PS-containing membranes using x-ray reflectivity. X-ray reflectivity patterns are generated by electron density profiles along the axis normal to the interface. We used phospholipids with 1-stearoyl-2-oleoyl (SO) tails because their longer acyl chains help reduce radiation damage [2]. We used a conventional 2-box model for the lipids and multiple slabs to represent the protein in each of its orientations [4]. The orientation, insertion, and surface coverage of the protein are varied to best fit the x-ray data. Initially, we used the crystal structure of the Ig domain of Tim3 [1] as the model but found its orientation to be unresolvable. We therefore employed Highly Mobile Membrane Mimetic (HMMM) Molecular Dynamics (MD) simulations to resolve a fully relaxed membrane bound model of Tim3 based on the crystal structure, which we posit to be more representative of membrane-bound Tim3.

This structure conveys key hints for possible modulators for Tim3 binding additional to its PS binding site. The Ig domain of Tim3 is inserted into the membrane with its hydrophobic residues on the FG loop and CC' loop. The tryptophan tracked in the fluorescence assay, W35, is located on the CC' loop of Tim3 and projects into the membrane. A leucine residue, L99, is inserted further into the interface of the headgroup and tailgroup. Note that the binding site could only adopt the optimal binding conformation with the two hydrophobic tips well inserted into the membrane, suggesting Tim3 could be sensitive to membrane packing density. We also have identified two lysine residues, K103 and K104, and two arginine residues, R50 and R54, in close proximity to the membrane, which could mediate electrostatic interactions.

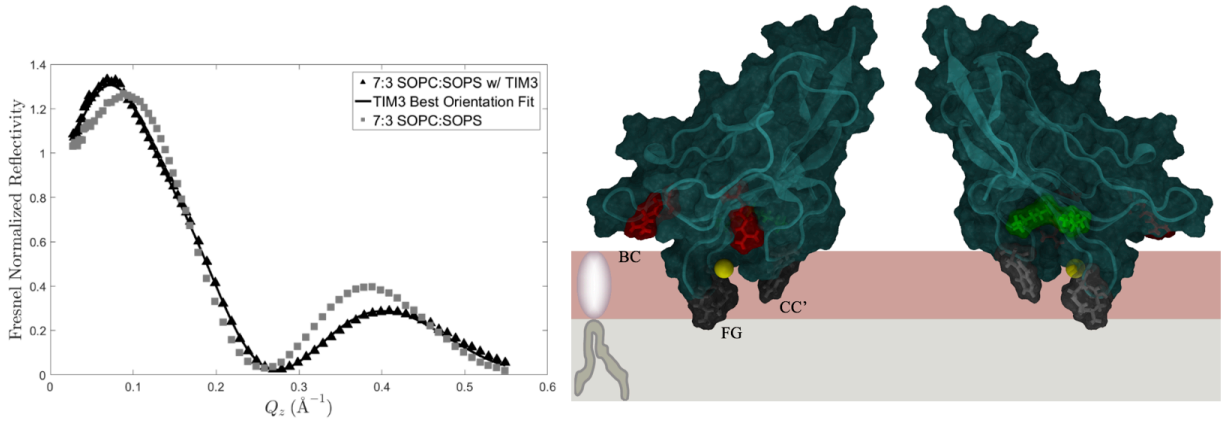

Figure 4.2: The structure of membrane bound Tim3 resolved by X-ray reflectivity. (**a**) X-ray reflectivity data and fitting for SOPC:SOPS = 7:3 at surface pressure 25 mN/m. (**b**) The inferred configuration of Tim3 bound to the lipid layer. The yellow sphere denotes the Ca^{2+} coordinated by the binding site. Pink denotes the assumed head group region while the gray denotes the tail group region. Hydrophobic residues that insert into the membranes are W41 and L99, highlighted in gray, the two lysine residues close to the membrane are K103 and K104, highlighted in green, and the two arginine residues close to the membranes are R50 and R54, highlighted in red. BC, CC' and FG are the three flexible loops extending from the protein.

4.2.4 The insertion of hydrophobic residues is supported by the binding free energy dependence on urea concentration

The interaction model in **Figure 4.2b** suggests the burying of a tryptophan and a leucine residue into the relatively nonpolar region of the membrane, which would create favorable hydrophobic interactions. To test whether hydrophobic interaction is indeed involved in Tim3 binding, we measured the dependence of binding affinity to POPC:POPS=7:3 vesicles on urea concentration up to where the protein unfolded. As seen in **Figure 4.3**, $K_d(Ca^{2+})$ increased with urea concentration, and the change in binding free energy relative to no urea showed a linear increase with a slope of 2.26 kJ/mol. At the 2.5 M urea, the protein spectra became less consistent, thus the large error bars for this data point. Beyond 2.5 M urea, the unbound protein spectra became significantly widened, indicating unfolding of the domain.

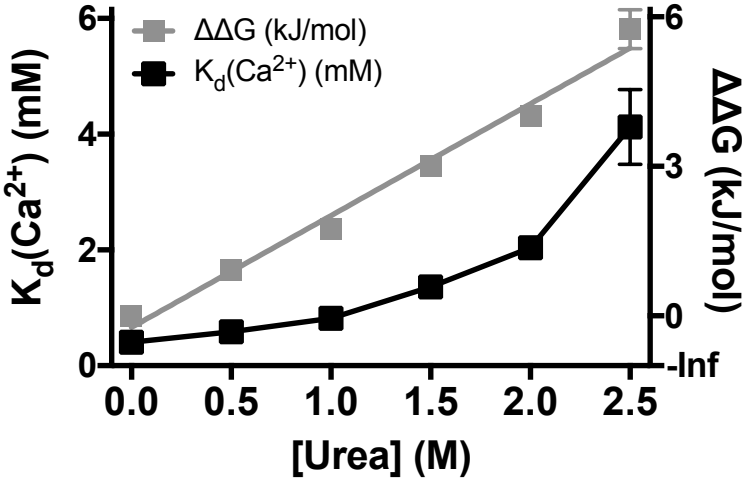

Figure 4.3: The effect of urea on Tim3 binding to POPC:POPS=7:3 vesicles. The experiment was performed as in **Figure 4.1** using Ca^{2+} titrations. Free energy changes were calculated from $K_d(Ca^{2+})$ values using the equation $\Delta G = RT \ln K_d$. The data was fitted to the equation $\Delta G = 2.26[\text{Urea}] - 0.23$ (kJ/mol). The standard deviation for the slope is 0.119, and the standard deviation for the intercept is 0.178. To match the osmotic pressure inside and outside the vesicles, vesicles were prepared in buffer of with given urea concentration at each point.

4.2.5 Tim3 binding is highly sensitive to membrane packing density

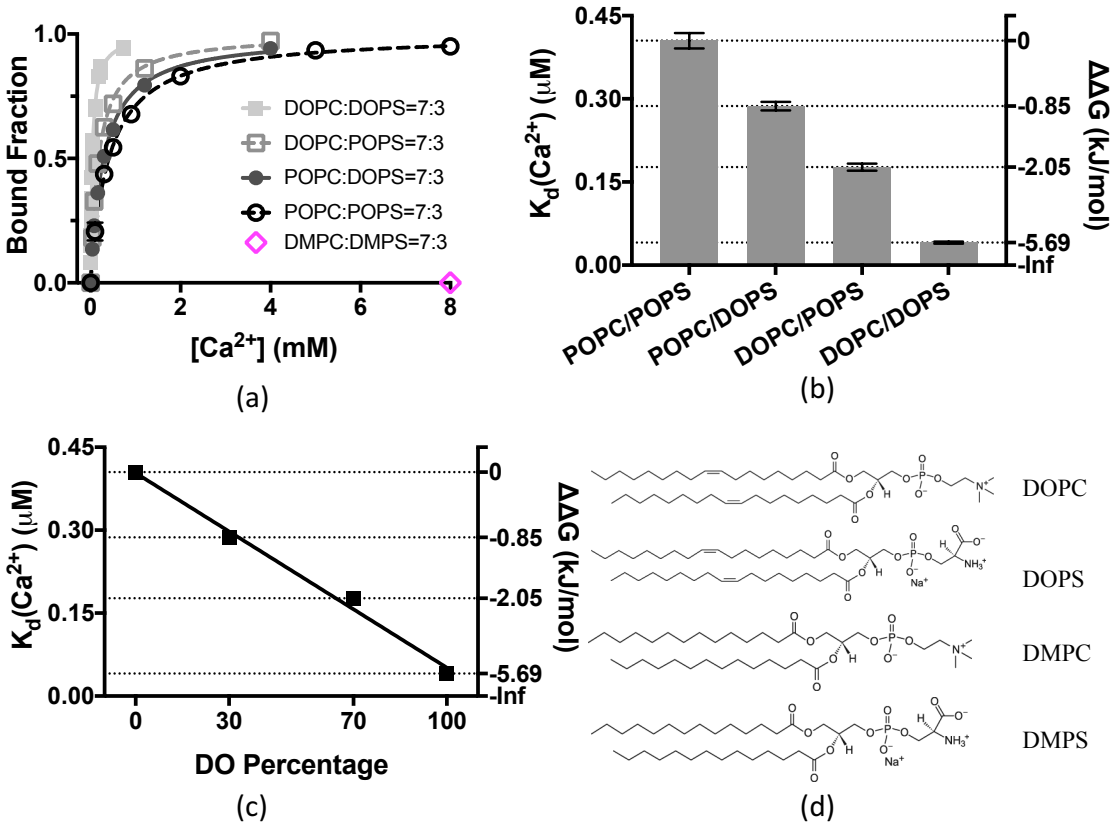

Figure 4.4: The binding of Tim3 to vesicles of phospholipids with loose and tight packing. (a): The Ca^{2+}-dependent association of Tim3 with PS:PC=7:3 vesicles containing saturated (DM), one unsaturated (PO), or two unsaturated (DO) acyl chains. (**b**): From the data in Panel a, $K_d(Ca^{2+})$ values were calculated as follows: for POPC:POPS, 405 ± 14 μM; POPC:DOPS, 287 ± 7.7 μM; DOPC:POPS, 177 ± 6.5 μM; and DOPC:DOPS, 40.7 ± 1.7 μM. (**c**): $K_d(Ca^{2+})$ values from panel b were plotted as a function of DO content. (**d**): Structure of DO and DM lipids.

The interaction model, the tryptophan fluorescence spectral shift, and Tim3 binding dependence on urea strongly supported the insertion of hydrophobic residues into the nonpolar region of the membrane. This insertion entails to work against the membrane stretching modulus, yet is necessary for favorable hydrophobic interactions, as well as to bring the PS binding site to proper

alignment with the membrane headgroup region. As protein insertion into lipid membranes entails a sensitivity to membrane packing density or the closely related membrane stretching modulus [5, 6, 7], we determined the binding affinity of Tim3 to vesicles with different lipid packing density/stretching modulus assuming that tighter packing and higher stretching modulus will reduce the affinity of the protein for the membrane. For this study, we substituted the tailgroup from the mono-unsaturated PO to the more loosely packed bi-unsaturated 1,2-dioleoyl (DO), or the more densely packed saturated 1,2-dimyristoyl (DM) tailgroups. As seen in **Figure 4.4a-c**, substituting DO lipids for PO lipids significantly increased the membrane binding affinity of Tim3, while substituting DM lipids for PO lipids reduced the binding lower than our detection limit. Substituting DO tails for all of the PO tails produced a 10-fold decrease in the dissociation constant, from 405 μM to 41 μM.

4.2.6 Tim3 binding is sensitive to membrane negative charge density and buffer ionic strength

The interaction model in **Figure 4.2b** also suggested potential interaction with negative charges on the membrane due to the two lysine and two arginine residues in proximity to the membrane. Here we first confirmed that Tim3 did not bind appreciably to neat zwitterionic POPC vesicles [2]. Previously, it was found that Tim3 affinity to POPC:POPS=95:5 vesicles was much weaker than POPC:POPS=7:3 vesicles [2], which is confirmed here, as seen **Figure 4.5a**. We next determined the affinity of Tim3 to POPC:POPS:POPG=70:5:25 vesicles. POPG occupies nearly the same area as POPS [8], and bears one negative charge as POPS. As seen in **Figure 5a**, POPC:POPS:POPG=70:5:25 vesicles was equal in ability to bind Tim3 as POPC:POPS=7:3

vesicles, which meant 25% POPG was equally capable of binding Tim3 as 25% POPS in addition to 5% POPS.

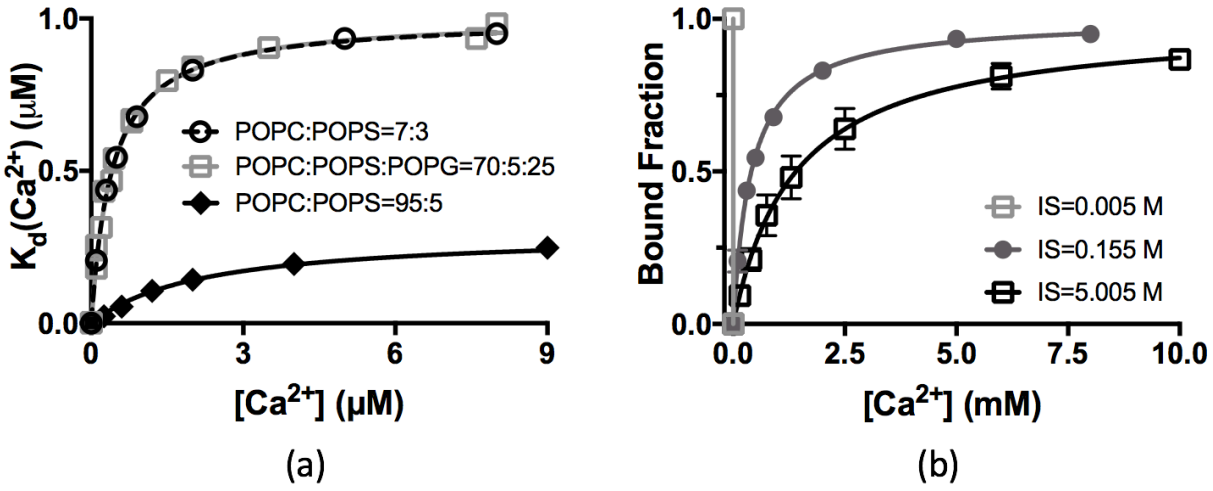

Figure 4.5: Tim3 exhibited electrostatic interaction with PS-containing vesicles. (a): Tim3 affinity to POPC:POPS=7:3 is similar to POPC:POPS:POPG=70:5:25 vesicles. The $K_d(Ca^{2+})$ for POPC:POPS=7:3 and POPC:POPS:POPG=70:5:25 vesicles was 416 ± 16 and 405 ± 13 μM respectively. The $K_d(Ca^{2+})$ is not calculated for the POPC:POPS=95:5 as the binding is too low to yield an accurate estimate. (**b**) Tim3 binding to POPC:POPS=7:3 vesicles is dependent on ionic strength. As HEPES has a pK_{a_2} at 7.5, half of the HEPES will be in its acid form, while the other half in dissociated form, so the ionic strength (IS) of 10 mM HEPES buffer will be 5 mM without NaCl. The $K_d(Ca^{2+})$ increased to 1.39 ± 0.12 mM at IS = 5.005 M, from 0.405 ± 0.014 mM at IS = 0.155 M. This increase in $K_d(Ca^{2+})$ represent a binding free energy change ($\Delta\Delta G$) of 3.05 kJ/mol. At IS = 0.005 M, the binding affinity became so strong that the background Ca^{2+} at 1.2 μM in the buffer was enough to recruit all protein to the membrane. Chelating away this background Ca^{2+} using 1 mM EGTA, however, released all bound Tim3.

We then studied the affinity of Tim3 to POPC:POPS=7:3 vesicles under different ionic strength by adding different amounts of NaCl to the 10 mM HEPES buffer at pH 7.5. As show in **Figure 4b**, binding affinity is reduced at high ionic strength, while dramatically enhanced at low ionic strength. Without adding more Ca^{2+}, the background Ca^{2+} in the HEPES buffer at 1.2 μM (**Figure 4.6**) is enough to bring all Tim3 to bind at ionic strength 0.005 M. Due to limitations in

our experimental reagents and setup, we were not able to conduct a fine stepped Ca^{2+} concentration serie at 0.005 M, but we found adding 1 mM EGTA to remove all free Ca^{2+} released all bound protein.

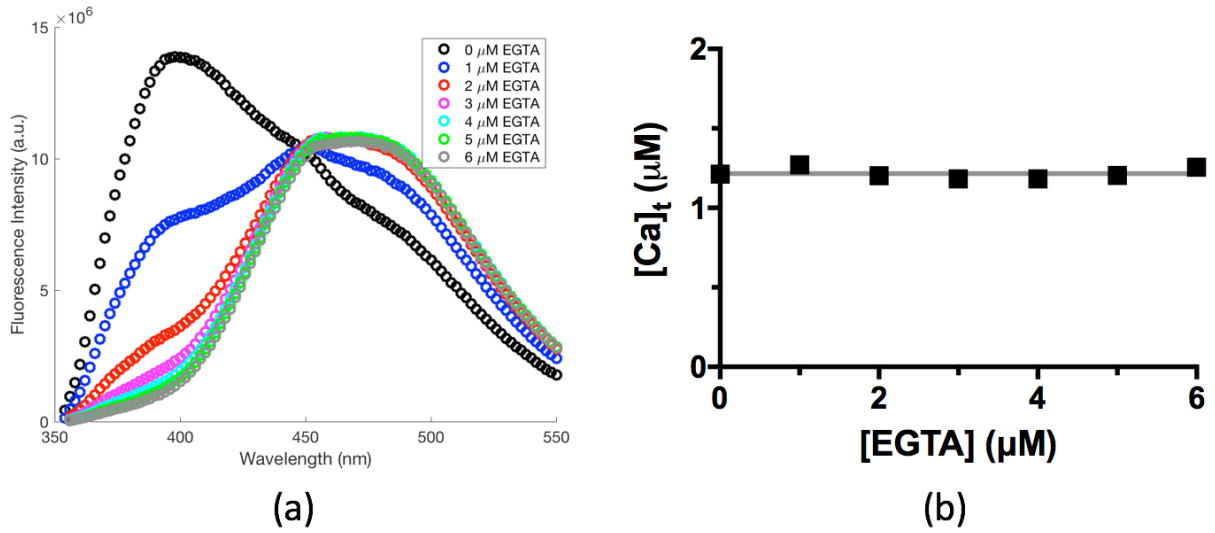

Figure 4.6: measurement of Ca^{2+} concentration using Indo-1 and EGTA. (**a**): The fluorescence spectra of Indo-1 with the 10 mM HEPES 150 mM NaCl buffer with various amounts of EGTA. Indo-1 concentration was 1 μM. As EGTA concentration is increased, the Indo-1 fluorescence spectra shift from the Ca^{2+} chelating form to the Ca^{2+} free form, with decreasing intensity at the 400 nm peak, and increasing intensity for the 450 – 500 nm peak. (**b**): Calculated total calcium concentration and fit to the a constant value based on mass action law. The dissociation constant used for Indo-1 was $K_d(\text{Indo}) = 230$ nM [Paredes 2008, Grynkiewicz 1985], and that for EGTA was $K_d(\text{EGTA}) = 38.5$ nM [Tien 1989]. By assuming the Indo-1 bound fraction is b_max when no EGTA is added, and b_min when 6 μM, we obtained the hypothetically bound and free spectra. We then fit the spectra with EGTA concentrations in between assuming those in the middle are linear combinations of the bound and unbound spectra. These bound fractions are used to fit to the mass action law equation to generate the total calcium concentration. The total background Ca^{2+} concentration was fit to be 1.2 μM.

4.2.7 Tim3 binding is highly sensitive to small increments in membrane PA

With a deeper understanding of the physicochemical modulators for Tim3 binding, we next determined the affinity of Tim3 for large unilamellar vesicles composed of various phospholipids bearing 1-pamiltoyl-2-oleoyl acyl chains but different headgroups. As shown above, POPC:POPS=7:3 vesicles were avid for Tim3 in the presence of calcium (**Figure 4.1a**). We then substituted for POPS increasing amounts of the anionic phospholipids: POPG or POPA. We varied the Ca^{2+} concentration in each series to determine the Tim3 dissociation constant $K_d(Ca^{2+})$ as a measure of the strength of the overall binding reaction. We already showed that Tim3 required PS to bind to the membranes in that Tim3 did not bind to vesicles composed of binary mixtures of PA/PC or PG/PC. That is, $K_d(Ca^{2+})$ diverges to infinity when no PS is present. Substituting a small fraction of the PS with PG or PA (to make 30 mol% total anionic phospholipid) decreased the dissociation constant for Tim3 (i.e., increased its binding affinity) as also seen in **Figure 4.5**. PA was more effective than PG in this regard. Indeed, substituting PA for 1% out of the 30 mol% PS reduced the dissociation constant, $K_d(Ca^{2+})$, for Tim3 to from 405 μM to 138 μM, or a $\Delta\Delta G$ of 2.67 kJ/mol, and substituting 10% PA decreased the $K_d(Ca^{2+})$ to 77.5 μM, or a $\Delta\Delta G$ of 4.10 kJ/mol. In contrast, Tim1 and Tim4 were not sensitive to PA (**Figure 4.8**). Note that, as PA was increased above 20 mol% and PG increased above 10 mol%, the dissociation constant for Tim3 increased. No Tim3 binding was observed with 30 mol% of either PA or PG in the absence of PS. Thus, the Ca^{2+} dependent membrane binding of Tim3 appeared to be specific for PS. Secondary interactions between Tim3 and PA and PG presumably strengthened its membrane binding but this enhancement diminished with decreasing levels of PS.

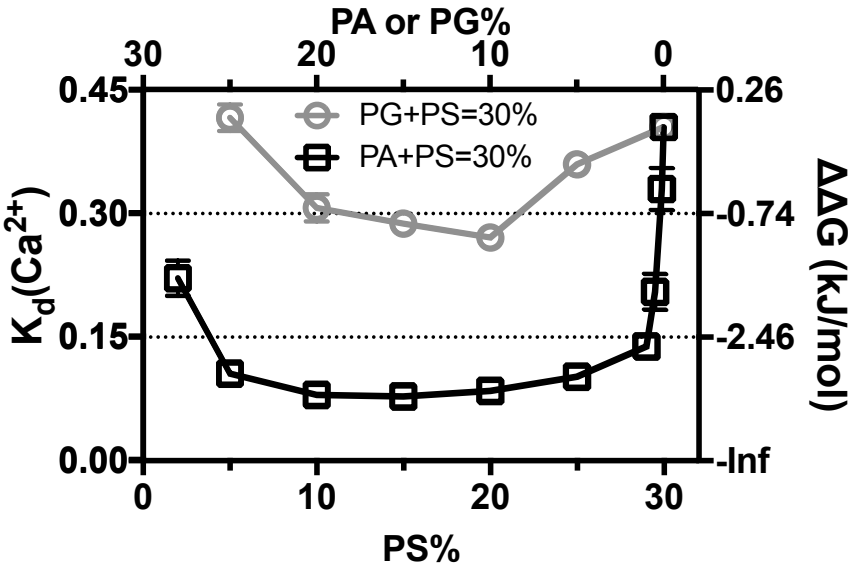

Figure 4.7: The association of Tim3 with lipid vesicles containing PS, PA and PG. $K_d(Ca^{2+})$ was determined using vesicles containing 70% POPC and 30 mol% anionic phospholipids distributed between PS and PA or PG. The total lipid concentration was 300 μM, Tim3 concentration was 170 nM. The temperature was 23 ˚C. No binding was detected for 30% PG and PA (not plotted). The change in binding free energy was calculated in reference to the POPC:POPS=7:3 system.

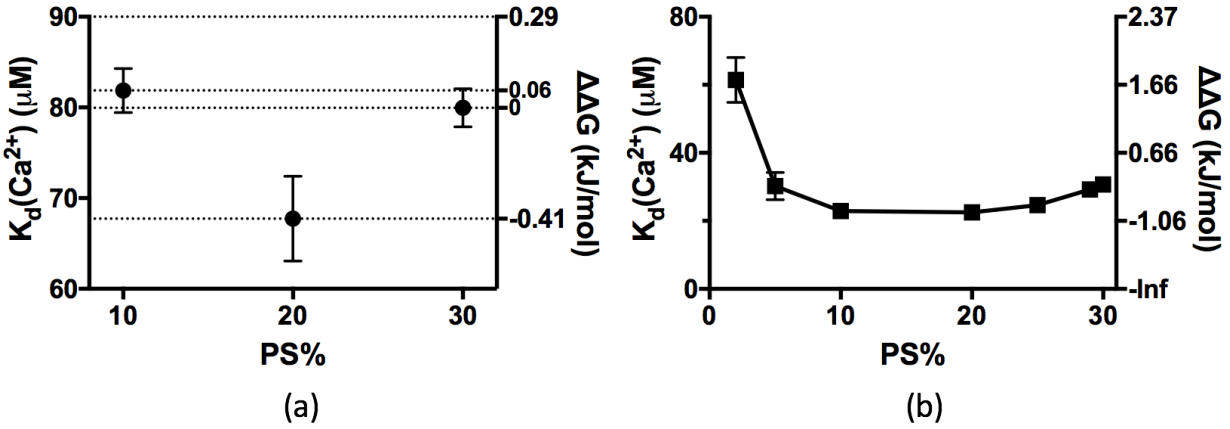

Figure 4.8: Tim1 and Tim4 do not have strong sensitivity to PA. (a): Tim1 is not sensitive to small amounts of PA in the lipid membrane. Here PA+PS is kept at a constant 30%. $K_d(Ca^{2+})$ of Tim1 to PC/PS/PA vesicles, with values 81.87 $\pm$ 2.4 μM at 20% PA, 67.7 $\pm$ 4.7 μM at 10% PA, and 80.0 $\pm$ 2.1 μM at 0% PA. (**b**): **Tim4 is not sensitive to small amounts of PA in the lipid membrane.** Here PA + PS is kept at a constant 30%. PA% used were: 28%, 25%, 20%, 10%, 5%, 1%, and 0% from left right.

4.2.8 Ca^{2+} sequestration of PA, PS, PG, and PC measured by total reflection x-ray fluorescence (TXRF)

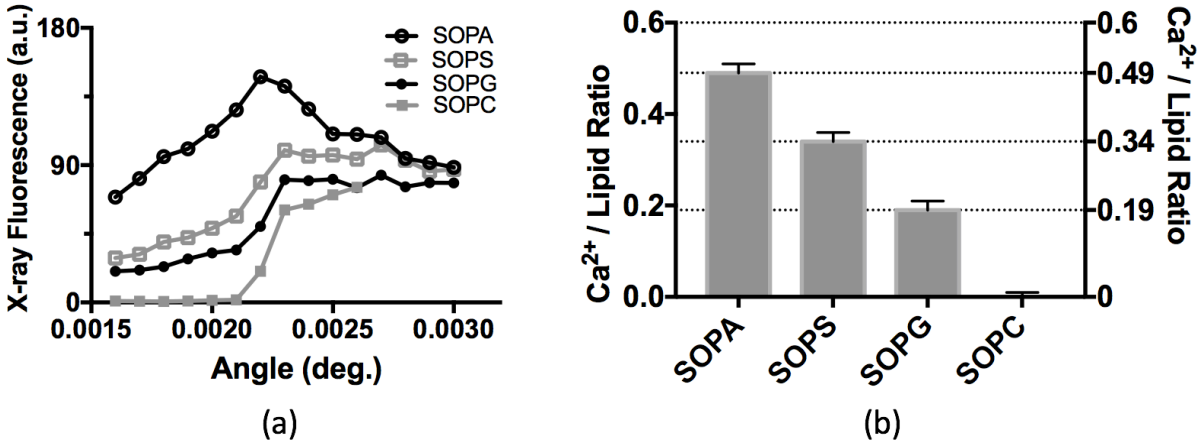

Figure 4.9: measurement of Ca^{2+} sequestration by SO monolayers by TXRF. (a): Ca^{2+} x-ray fluorescence signal around the critical angle from four DO lipid films. The buffer contains 10 mM HEPES, 150 mM NaCl at pH 7.5. (**b**): Ca^{2+}/lipid ratio for four DO lipids as a result of analyzing data from panel **(a)**. The analysis was carried out using *XeRay* using a two box model for the lipid, and that all sequestered Ca^{2+} reside in the headgroup layer. The ratios between number of Ca^{2+} cations and lipid molecules are 0.49 ± 0.02, 0.34 ± 0.02, 0.19 ± 0.02, and 0 ± 0.01 respectively.

We next used TXRF to measure Ca^{2+} sequestration by SO lipid monolayers [9, 10, 11, 12]. Development of this technique is discussed in Chapter 5. As in the x-ray reflectivity measurements, SO lipids are used to limit radiation damage to the monolayer. As seen in **Figure 4.9a**, Ca^{2+} fluorescence signal around the critical angle decreases in the order: SOPA > SOPS > SOPG, while SOPC showed no measurable signal around the critical angle. We analyzed the amount of Ca^{2+} sequestration for each film using the graphical user interface tool *XeRay* [13], with the assumption that all sequestered Ca^{2+} reside in the headgroup region of the monolayer

[10, 13]. The surface sequestered amount of Ca^{2+} is converted to the ratio between the number of Ca^{2+} ions and the number of lipid molecules, as seen in **Figure 4.9b**.

4.3 Discussion

4.3.1 Tim3 interaction model with PS-containing lipid membrane offers hints in key modulators of binding

Tim3's much weaker affinity to PS compared to Tim1 and Tim4 as seen in **Figure 4.1a-b** led us to examine the structural determinants of the overall binding affinity. Homeostatic, free Ca^{2+} concentration in the human serum is around 1 mM [14]. With a $K_d(Ca^{2+})$ of 0.405 mM, assuming excess POPC:POPS=7:3 vesicles, the equilibrium bound fraction of Tim3 would be around 70% at the physiological calcium level. This relatively weak interaction allows room for other modulators in addition to the PS binding site. To better grasp the possible additional modulators of Tim3 binding, we built a molecular interaction model based on fitting the protein orientation to x-ray reflectivity data from monolayer bound Tim3 Ig domains. Results of the model offered strong hints that these possible modulators could be membrane packing density and membrane negative charge density. As seen in **Figure 4.2**, the insertion of two hydrophobic tips on the protein is necessary to bring the PS binding site to proper alignment with the headgroup region. Thus, a more tightly packed membrane could reduce binding affinity by preventing proper protein alignment, while a more loosely packed membrane could augment binding affinity. The two lysine residues and two arginine residues close to the membrane were

hypothesized to sense the negatively charged species on the membrane. These two possible modulators were subsequently studied and their effects were found to be quite pronounced.

4.3.2 Hydrophobic interaction between Tim3 and lipid membranes

According to the interaction model in **Figure 4.2**, the main insertion residues of the protein are likely to be the hydrophobic tips of the FG and CC' loops, corresponding to a leucine and a tryptophan residue respectively. The insertion of hydrophobic residues into the relatively nonpolar membrane environment is supported by two lines of evidence. Firstly, the tryptophan spectral shift tracked in the binding studies indicates the relocation of the tryptophan residue from a polar to nonpolar membrane environment. Secondly, the reduced binding affinity with increasing urea concentration, as seen in **Figure 4.3**, suggests the existence of hydrophobic interactions between the lipid bilayer and Tim3. As urea does not significantly alter the physical properties of the lipid bilayer [16], but acts to solubilize hydrophobic entities [17, 18, 19, 20, 21], it has the effect of making the exposed tryptophan and leucine residues on Tim3 more energetically favorable in the bulk. The change in binding free energy from 0 to 5 M urea is 5.65 kJ/mol, which is likely the lower bound of the contribution of hydrophobic interaction to the total binding free energy.

4.3.3 Tim3 is selective for more loosely packed membranes, possibly due mainly to a reduction in mechanical resistance for membrane insertion

Our results showed a ten-fold higher affinity of Tim3 to the bi-unsaturated DO vesicles than the mono-unsaturated PO vesicles, while no detectable affinity towards the saturated DM lipids, as shown in **Figure 4.4a-c**. The area per molecule and charge of the different lipids are summarized in **Table 4.2**. For lipids of the same headgroup, the area per molecule decreases as the number of unsaturation sites decreases. POPG has the same size as POPS, while POPA is roughly 10% smaller. Therefore, the packing density of lipid membrane seems to be a key parameter in modulating Tim3 binding. Peripheral binding proteins that shallowly insert into lipid membranes are known to be capable of sensing membrane packing density [5, 6, 7, 15], for which two major mechanisms have been proposed. Lower membrane packing density generally allows for greater degree of protein insertion [5], since imperfect lipid packing presents more spatial defects in the membrane [6].

This increase in binding free energy with increased packing density may also be explained mechanically. To insert into the lipid membrane, Tim3 would work against the stretching modulus of the membrane. The stretching moduli of mixed PC and PS bilayers could not be found, however, measurement or simulation results on the stretching moduli of pure DOPC, POPC, and DMPC can be found in the literature. The stretching moduli of DOPC [22, 23], POPC [24, 25, 26], and DMPC [27] are about 176, 255, and 855 mN/m respectively. Assuming the stretching moduli of 7:3 DOPC:DOPS and 7:3 POPC:POPS to be close to those of DOPC and POPC, and further assuming the change in binding free energy between the all DO lipid vesicles and the all PO lipid was due to the difference in work against the stretching moduli, the insertion area of Tim3 would be

$$\Delta A = \frac{\Delta\Delta G}{N_A \cdot \Delta m} \approx \frac{5.69 kJ/mol}{N_A \cdot 79 mN/m} = 12\ \text{Å}^2.$$

Lipid	Charge	Area per Molecule ($Å^2$)	Stretching Modulus (mN/m)	Phase at 23 ℃
DOPC	0	72.5* (no salt) [28]	176 [22], 188 [23]	Fluid
DOPS	-1	64.1* (no salt) [28]		Fluid
POPC	0	64# [8]	272 [24], 180 [25], 330 [25], 240 [26]	Fluid
POPS	-1	55# [8, 29]		Fluid
POPG	-1	55# [29]		Fluid
POPA	-1 to -2	50±2# [8]		Fluid
DMPC	0	48.1(gel)*, 59.7 (fluid)* [28]	855 [27]	Gel (T_m = 24℃)
DMPS	-1	40.8* [28]		Gel

Table 4.2: charge, area per molecule, and stretching moduli of lipids in binding studies. *: experimental result. #: simulation result.

Incidentally, 12 $Å^2$ is around the size of the indole group of a tryptophan residue. As the indole group would not insert flat into the membrane, and the other hydrophobic residues would also contribute to the total insertion area, the insertion area might be close to 12 $Å^2$. On the other hand, assuming the insertion area is indeed 12 $Å^2$, the change in binding free energy from PO to DM lipids would be

$$\Delta\Delta G(hypothetical) = \Delta A \cdot N_A \cdot \Delta m \approx 43.3\text{kJ/mol},$$

or bout 17 kT. This large increase in binding free energy would result in binding affinity below the detection limit by the fluorescence binding assay.

4.3.4 Tim3 is selective for high membrane negative charge density likely due to electrostatic interaction with negatively charged lipid headgroups

Tim3 affinity to POPC:POPS=95:5 vesicles was much lower than that to POPC:POPS=7:3 vesicles as seen in [2], and **Figure 4.5a**. However, also seen in **Figure 4.5a**, Tim3 affinity to POPC:POPS:POPG=70:5:25 vesicles is similar to that of POPC:POPS=7:3 vesicles, while POPC:POPG=7:3 vesicles did not show detectable binding (**Table 4.1**). These pieces of

evidence, taken together, suggest that Tim3 needs a base amount of PS lipids (such as 5%), beyond which the protein only seems be affected by the overall negative charge of the headgroup, so the additional 25% lipid being POPS or POPG does not alter the binding affinity to first order approximation.

Two possible mechanisms exist for this sensitivity to membrane negative charge density. Firstly, negatively charged lipids sequester significant amounts of Ca^{2+} under physiological salt conditions, as shown in **Figure 4.9a-b**. The Ca^{2+}/SOPA ratio of 0.49 ± 0.02 agrees with the stoichiometry between DMPA liposomes and Ca^{2+} under low availability of Ca^{2+}, but is half of that under high availability of Ca^{2+} [30]. This stoichiometry is also half of the same monolayer film under a 10 mM HEPES buffer and no NaCl at pH 7.5 [13]. For SOPS, the result of 0.34 ± 0.02 Ca^{2+} per headgroup, a ratio of less than 0.5 agreed with the measurements of PS headgroups in the sequestration of Ca^{2+} and other cations [31, 32], as well as with molecular dynamics simulations [23, 34]. However, this mechanism seemed not likely to be the major cause for enhancing Tim3 affinity, as SOPG sequesters less Ca^{2+} while it has very similar ability to stabilize bound Tim3. Secondly and more probably, Tim3 interacts directly with the negatively charged residues via its four basic residues identified in the molecular interaction model (see **Figure 4.2b**). Moreover, the existence of this electrostatic interaction is also supported by the dependence of binding on ionic strength, as shown in **Figure 4.5b**.

4.3.5 Tim3 physicochemical selectivity

So far our evidence supports the physicochemical selectivity of Tim3 to PS-containing lipid membranes. That is, (1) the binding site is chemically specific to both Ca^{2+} and the PS headgroup, (2) the insertion of hydrophobic residues brings the PS binding site and the hydrophobic residues to proper alignment with the membrane, and (3) the peripheral residues sense the negative charge distribution on the membrane. These three selectivity factors combined to give four modes of interaction that could make up the total binding free energy: specific Ca^{2+} mediated binding to PS, electrostatic interactions, hydrophobic contacts, and mechanical stretching of the membrane. Among these four modes, the PS binding site seems to contribute the majority of the binding energy, as Tim3 did not bind to vesicles without the PS headgroup in the PO vesicle binding studies (**Table 4.1**). Even though lowering the ionic strength to 0.005 M dramatically increased the binding affinity, such that 1.2 μM Ca^{2+} saturated the binding, removing the free Ca^{2+} using EGTA could still free all Tim3, as seen in **Figure 4.5b**.

The combined energies of these four modes of interaction provide Tim3 with a parameter space that tunes its interaction with plasma membranes that is significantly different from the other two TIM proteins as well as from more than 40 other PS binding proteins [35]. The use of a combination of the four modes of interaction, in addition to other possible modes of interaction such as curvature sensing [36, 37], might well be shared by some of the PS binding proteins, giving them the parameter space to be adequately differentiable in a physiological environment.

4.3.6 Connection between Tim3 physicochemical selectivity and its sensitivity to PA

Small increments of PA, such as 1%, would not be able to change significant affect the amount of sequestered Ca^{2+} concentration, as seen in **Figure 4.9b**, but significantly reduced the $K_d(Ca^{2+})$ from 405 μM to 138 μM, a 3-fold decrease (**Figure 4.7**). Thus, the PA enhancement effect is likely not due to increased local Ca^{2+} concentrations. Instead, the PA enhancement effect could be the combined result of two other elements of its physicochemical selectivity. First, PA could enhance electrostatic interaction. Phosphatidic acid has a pK_{a2} for the second proton in the headgroup of around 7.5 [38]. Thus, at pH 7.5, it would bear a nominal charge of about -1.5. But, due to further deprotonation, this negative charge could be increased significantly through contact with a lysine or arginine residue [38]. This increased charge was found to approach -2 for lysine. At 1% PA, assuming even distribution of PA in the inner leaflet and outer leaflet, the external PA concentration is 1.5 μM, or on average 8.8 PA headgroups for each Tim3. As seen in **Figure 4.7**, addition of PA beyond 5% did not further affect the binding affinity much, which suggested the PA effect could be "saturated", also consistent with Tim3 associating with a given number of PA headgroups. The increased negative charge of PA could result in Tim3 experiencing a significantly higher charge environment. Secondly, PA has a much smaller headgroup cross-sectional area compared to the tail cross-sectional area and so has a negative spontaneous curvature [8, 39], which could facilitate protein insertion either by introducing defects into the membrane, or by virtue of less special hindrance to the inserted parts on the protein. The observed binding enhancement by PG appeared to be secondary, and could be due to slight alterations in the hydrogen bonding or charge/charge interaction between Tim3 and PG or PS headgroups.

4.3.7 Tim3 sensitivity to PA and membrane packing density could be involved in the regulation of neutrophils

Could the high sensitivity of Tim3 to small amounts of PA and membrane packing density be physiologically relevant? PA can be produced rapidly when phosphatidylcholine phospholipase D (PC-PLD) is activated in mammalian cells, resulting in PA levels as high as 1-4% of the lipid in the plasma membrane [35]. Immune receptors are coupled to the activation of PC-PLD, although it is uncertain what immune events ensue [40]. In mammalian cells, a wide range of immune processes have been proposed to be affected by PC-PLD activation and the generation of PA; these include phagocytosis, the respiratory burst in neutrophils, inflammation, and diabetes [41, 42, 43]. That proteins and lipids in the cell membrane self-organize into functional domains with higher packing density is well established [44, 45, 46]. The dynamics of such organized domains might modulate membrane packing density and alter the affinity of Tim3 towards the cell membrane.

Several lines of evidence support the possibility that Tim3 is involved in the regulation of neutrophils. Neutrophils, the most abundant white blood cells, are highly mobile phagocytes that act as first responders to acute inflammation where they rapidly engulf foreign materials [47]. To do so, they rapidly generate copious highly oxidative molecules that attack the internalized particles. Neutrophil apoptosis follows shortly [48]. An important hallmark of neutrophil apoptosis is the activation of PC-PLD, which significantly elevates membrane PA content. Furthermore, at later stages of apoptosis, PS is exposed on the outer leaflet of neutrophils, and the tightly packed self-organized cell membrane domains start losing integrity These factors

would promote the binding of Tim3 to such apoptotic neutrophils. This increased reaction would result in the phagocytosis of the apoptotic neutrophils cells by Tim3-expressing macrophages, such as monocytes, dendritic cells, and natural killer cells. Since the regulation of neutrophils is highly correlated with immune-deficient diseases like chronic viral infection [49, 50] and various types of cancer [51], and Tim3 has been repeatedly linked to these diseases, the involvement of PA in Tim3 mediated pathways merits further investigation.

4.4 Conclusion

We quantified and confirmed that Tim3 binding to PS-containing membranes is the weakest compare to Tim1 and Tim4. Its PS binding site specifically requires Ca^{2+} and a PS headgroup, which contributes the majority of the binding energy. Tim3 has a high sensitivity to membrane packing density, which might arise from the shallow insertion of the hydrophobic tips of the FG and CC' loop into the membrane. Tim3 also has significant electrostatic interaction with the negative charges on the membrane, which is attributed to the four basic residues on Tim3 are in close proximity to the membrane. This chemical specificity, membrane packing sensitivity, and negative charge dependence combined to give Tim3 a unique physicochemical selectivity to PS-containing membranes, so that it could be adequately differentiable from the other PS binding proteins. We observed that small increments of membrane PA could significantly enhance Tim3 binding. The combined effect of the membrane packing sensitivity and charge sensitivity could explain the Tim3's high sensitivity to PA. The high sensitivity of Tim3 to PA in lipid membranes indicates that the exposure of PA at the plasma membrane might serve as a second signal in the immunological functions of Tim3. A change in target cell plasma membrane PA exposure or in

its lipid packing density could be involved in various pathophysiological processes. Future efforts in the study of the sensitivity of Tim3 to PA and membrane packing density, such as the phagocytosis of apoptotic neutrophils, could lead to a better understanding of its involvement in health and disease.

4.5 Methods and Materials

Materials. Analytical grade 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), $CaCl_2$, NaCl, and urea and solvents including chloroform, methanol, acetone were procured from Fisher Scientific (Hampton, NH). Lipids including POPC, POPS, POPG, POPA, SOPC, SOPS, SOPG, SOPA, DOPC, DOPS, DMPC, DMPS were obtained from Avanti Polar Lipids (Alabaster, AL). The buffers used in most studies were composed of 10 mM HEPES and 150 mM NaCl at pH 7.5, (termed HBS).

Protein Production. The cDNA of Tim4, Tim1, and Tim3 were a generous gift from Shigekazu Nagata (Osaka University, Osaka). The cDNA corresponding to the murine form of the PS recognition domains (Ig domain) of Tim4 [residues D4-A112; Protein Data Bank code 3BIB (23)], Tim1 [residues D2- E112; Protein Data Bank code 2OR8 (38)], and Tim3 [residues G6-K112; Protein Data Bank code 3KAA (30)] were cloned in frame with a C-terminal 6× His tag into the pAcGP67A vector (BD Biosciences). Recombinant Tim4, Tim1, and Tim3 were all produced in Hi5 cells via baculovirus transfection, as previously described [52], with the exception that a higher salt buffer was used for purification (10 mM Hepes at pH 7.2 plus 450 mM NaCl buffer). The higher concentration of NaCl buffer was critical for maintaining protein

solubility during the purification process. After purification, protein was stored at 4 °C and was stable for several months. All mutant proteins were purified via the same procedures and showed identical elution profiles over FPLC and nearly equivalent protein yields, suggesting minimal disruption to the underlying protein structure.

4.6 References

1. DeKruyff, R.; Bu, X.; Ballesteros, A.; Santiago, C.; Chim, Y. -L.; Lee, H. -H.; Karisola, P.; Pichavant, M.; Kaplan, G.; Umetsu, D.; Freeman, G.; Casasnovas, J. T. Cell/Transmembrane, Ig, and Mucin-3 Allelic Variants Differentially Recognize Phosphatidylserine and Mediate Phagocytosis of Apoptotic Cells. *J. Immunol*. **2010,** *184*, 1918–1930.
2. Tietjen, G.; Gong, Z.; Chen, C. H.; Vargas, E.; Crooks, J. E.; Cao, K. D.; Heffern, C. T.; Henderson, J. M.; Meron, M.; Lin, B.; Roux, B.; Schlossman, M. L.; Steck, T. L; Lee, K. Y.; Adams, E. J. Molecular mechanism for differential recognition of membrane phosphatidylserine by the immune regulatory receptor Tim4. *Proc. Natl. Acad. Sci.* **2014,** *111*, E1463–E1472.
3. Kraft, C.; Garrido, J.; Leiva-Vega, L.; Romero, G. Quantitative Analysis of Protein-Lipid Interactions Using Tryptophan Fluorescence. *Sci. Signal.* **2009,** *2*, p14.
4. Málková, Š.; Long, F.; Stahelin, R.; Pingali, S.; Murray, D.; Cho, W.; Schlossman, M. 2005. X-Ray Reflectivity Studies of cPLA2α-C2 Domains Adsorbed onto Langmuir Monolayers of SOPC. *Biophys. J*. **2005,** *89*, 1861-1873.
5. Pande, A. H.; Qin, S.; Tatulian, S. A. Membrane fluidity is a key modulator of membrane binding, insertion, and activity of 5-lipoxygenase. *Biophys. J*. **2005,** *88*, 4084-4094.
6. Vanni, S.; Vamparys, L.; Gautier, R.; Drin, G.; Etchebest, C.; Fuchs, P. F.; Antonny, B. Amphipathic lipid packing sensor motifs: probing bilayer defects with hydrophobic residues. *Biophys. J.* **2013**, *104*, 575-584.
7. Sánchez, S. A.; Tricerri, M. A.; Ossato, G.; Gratton, E. Lipid packing determines protein–membrane interactions: challenges for apolipoprotein AI and high density lipoproteins. *Biochim. Biophys. Acta*. **2010,** *1798*, 1399-1408.

8. Dickey, A.; Faller, R. Examining the contributions of lipid shape and headgroup charge on bilayer behavior. *Biophys. J*. **2008,** *95*, 2636-2646.
9. Yoneda, Y.; Horiuchi, T. Optical Flats for Use in X Ray Spectrochemical Microanalysis. *Rev. Sci. Instrum*. **1971,** *42*, 1069–1070.
10. Bu, W.; Flores, K.; Pleasants, J.; Vaknin, D. Preferential affinity of calcium ions to charged phosphatidic acid surface from a mixed calcium/barium solution: X-ray reflectivity and fluorescence studies. *Langmuir.* **2009,** *25*, 1068–1073.
11. Novikova, N. N.; Zheludeva, S. I.; Konovalov, O. V. Total reflection X-ray fluorescence study of Langmuir monolayers on water surface. *J. Appl. Cryst*. **2003,** *36*, 727-731.
12. Graber, Z. T.; Wang, W.; Singh, G.; Kuzmenko, I.; Vaknin, D. Competitive cation binding to phosphatidylinositol-4, 5-bisphosphate domains revealed by X-ray fluorescence. *RSC Advances*. **2015,** *5*, 106536-106542.
13. Gong, Z.; Kerr, D.; Hwang, H. L.; Henderson, J. M.; Suwatthee, T.; Slaw B. R.; Cao, K. D.; Lin, B.; Bu, W.; Lee K, Y. Quantitative analysis of total reflection X-ray fluorescence from finely layered structures using XeRay. *Rev. Sci. Instrum.* **2017,** *88*, 033112
14. Larsson, L.; Ohman, S. Serum ionized calcium and corrected total calcium in borderline hyperparathyroidism. *Clin. Chem*. **1978,** *24*, 1962-1965.
15. Heron, D. S.; Shinitzky, M. Lipid fluidity markedly modulates the binding of serotonin to mouse brain membranes. *Proc. Natl. Acad. Sci.* **1980,** *77*, 7463-7467.
16. Costa-Balogh, F. O.; Wennerström, H.; Wadsö, L.; Sparr, E. How Small Polar Molecules Protect Membrane Systems against Osmotic Stress: The Urea− Water− Phospholipid System. *J. Phys. Chem. B.* **2006,** *110*, 23845-23852.
17. Sharp, K. A.; Madan, B.; Manas, E.; Vanderkooi, J. M. Water structure changes induced by hydrophobic and polar solutes revealed by simulations and infrared spectroscopy. *J. Chem Phys*. **2001,** *114*, 1791–1796.
18. Rezus, Y.; Bakker, H. J. Effect of urea on the structural dynamics of water. *Proc. Natl. Acad. Sci*. **2006,** *103*, 18417-18420.
19. Lee, M. E. E.; van der Vegt, N. F. Does urea denature hydrophobic interactions? *J. Am. Chem Soc*. **2006,** *128*, 4948–4949.
20. Rezus, Y. L.; Bakker, H. J. Observation of immobilized water molecules around hydrophobic groups. *Phys. Rev. Lett*. **2007,** *99*, 148301.

21. Guinn, EJ; Pegram, LM; Capp, MW. Quantifying Why Urea Is a Protein Denaturant, Whereas Glycine Betaine Is a Protein Stabilizer. *Proc. Nat. Acad. Sci.* **2011,** *108*, 16932-16937.
22. Fa, N; Lins, L; Courtoy, PJ; Dufrêne, Y. Decrease of Elastic Moduli of DOPC Bilayers Induced by a Macrolide Antibiotic, Azithromycin. *Biochimica et Biophysica Acta.* **2007**, *1768*, 1830-1838.
23. Tristram-Nagle, S.; Petrache, H. I.; Nagle, J. F. Structure and Interactions of Fully Hydrated Dioleoylphosphatidylcholine Bilayers. *Biophysical Journal.* **1998**, *75*, 917–925.
24. Janosi, L.; Gorfe, A. Simulating POPC and POPC/POPG Bilayers: Conserved Packing and Altered Surface Reactivity. *J Chem Theory Comput.* **2010**, *6*, 3267–73.
25. Binder, H.; Gawrisch, K. Effect of Unsaturated Lipid Chains on Dimensions, Molecular Order and Hydration of Membranes. *J Phys Chem B.* **2001**, *105*, 12378–12390.
26. Rawicz, W; Olbrich, KC; McIntosh, T; Needham, D. Effect of Chain Length and Unsaturation on Elasticity of Lipid Bilayers. *Biophysical Journal.* **2000**, *79*, 328-339.
27. Rosoff, M. Vesicles. **1996**, *CRC Press*. Brooklyn, New York.
28. Petrache, H. I.; Tristram-Nagle, S.; Gawrisch, K.; Harries, D. Structure and fluctuations of charged phosphatidylserine bilayers in the absence of salt. *Biophys. J.* **2004,** *86*, 1574-1586.
29. Mukhopadhyay, P.; Monticelli, L.; Tieleman, D. P. Molecular dynamics simulation of a palmitoyl-oleoyl phosphatidylserine bilayer with Na+ counterions and NaCl. *Biophys. J.* **2004,** *86*, 1601-1609.
30. Garidel, P.; Blume, A. Calcium induced nonideal mixing in liquid-crystalline phosphatidylcholine− phosphatidic acid bilayer membranes. *Langmuir*. **2000,** *16*, 1662-1667.
31. McLaughlin, S.; Mulrine, N.; Gresalfi, T.; Vaio, G.; McLaughlin, A. Adsorption of divalent cations to bilayer membranes containing phosphatidylserine. *J. Gen. Physiol*. **1981,** *77*, 445–473.
32. Roux, M.; Bloom, M. Ca2+, Mg2+, Li+, Na+, and K+ distributions in the headgroup region of binary membranes of phosphatidylcholine and phosphatidylserine as seen by deuterium NMR. *Biochemistry*. **1990,** *30*, 7077-7089.
33. Sinn, C. G.; Antonietti, M.; Dimova, R. Binding of calcium to phosphatidylcholine–phosphatidylserine membranes. *Colloids Surf. A.* **2006,** *282*, 410–419.

34. Martín-Molina, A.; Rodríguez-Beas, C.; Faraudo, J. Effect of calcium and magnesium on phosphatidylserine membranes: experiments and all-atomic simulations. *Biophys. J.* **2012,** *102*, 2095-2103.
35. Stace, C.; Ktistakis, N. Phosphatidic acid- and phosphatidylserine-binding proteins. *Biochim. Biophys. Acta.* **2006,** *1761*, 913–926.
36. Kirchhausen, T. Bending membranes. *Nat. Cell Bio*. **2012,** *14*, 906-908.
37. Antonny, B. Mechanisms of membrane curvature sensing. *Annu. Rev. Biochem*. **2011,** *80,* 101-123.
38. Kooijman, E. E.; Tieleman, P. D.; Testerink, C.; Munnik, T.; Rijkers, D. T.; Burger, K. N.; Kruijff, B. An electrostatic/hydrogen bond switch as the basis for the specific interaction of phosphatidic acid with proteins. *J. Biol. Chem*. **2007,** *282*, 11356–11364.
39. Kooijman, E. E; Chupin, V.; Fuller, N. L.; Kozlov, M. M. Spontaneous curvature of phosphatidic acid and lysophosphatidic acid. *Biochemistry*. **2005,** *44*, 2097-2102.
40. Melendez, A.; Allen, J. Phospholipase D and immune receptor signaling. *Semin. Immunol.* **2002,** *14*, 49–55.
41. Jenkins, G. M.; Frohman, M. A. Phospholipase D: a lipid centric review. *Cell. Mol. Life. Sci.* **2005,** *62,* 2305-2316.
42. McDermott, M.; Wakelam, M. J.; Morris, A. J. Phosphalipase D. *Biochem. Cell Biol.* **2004**, *82*, 225-253.
43. Liscovitch, M.; Czarny, M.; Fiucci, G.; Tang, X. Phospholipase D: molecular and cell biology of a novel gene family. *Biochem. J*. **2000,** *345*, 401-415.
44. Hancock, J. F. Lipid rafts: contentious only from simplistic standpoints. *Nat. Rev. Mol. Cell Biol*. **2006,** *7*, 456-462.
45. Pike, L. J. The challenge of lipid rafts. *J. Lipid. Res*. **2009,** *50,* S323-328.
46. Rajendran, L; Simons, K. Lipid rafts and membrane dynamics. *J. Cell Sci*. **2005,** *118*, 1099-1102.
47. Dahlgren, C.; Karlsson, A. Respiratory burst in human neutrophils. *J. Immunol*. *Met.* **1999,** *232*, 3-14.
48. Fox, S.; Leitch, A. E.; Duffin, R.; Haslett, C. Neutrophil apoptosis: relevance to the innate immune response and inflammatory disease. *J. Innate Immun.* **2010,** *2*, 216-227.

49. Galani, I. E.; Andreakos, E. Neutrophils in viral infections: Current concepts and caveats. *J. Leukoc. Biol*. **2015,** *98*, 557-564.
50. Drescher, B.; Bai, F. Neutrophil in viral infections, friend or foe? *Virus Res*. **2013,** *171*, 1-7.
51. Coffelt, S. B.; Wellenstein, M. D.; de Visser, K. E. Neutrophils in cancer: neutral no more. *Nat. Rev. Cancer*. **2016,** *16*, 432-446.
52. Scharf, L; Li, NS; Hawk, AJ; Garzón, D; Zhang, T; Fox, LM The 2.5 Å structure of cd1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. *Immunity.* **2010**, *33*, 853-862.

CHAPTER 5

STRUCTURAL BASIS FOR FOUR MODES OF INTERACTION BETWEEN THE MFG-E8 C2 DOMAIN AND LIPID MEMBRANES CONTAINING PHOSPHATIDYLSERINE

5.1 Overview

Milk fat globule epidermal growth factor 8 (MFG-E8), is a secreted phosphatidylserine (PS) receptor that acts as a linker between macrophages and apoptotic cells and PS-exposing vesicular structures. Different from the Tim proteins, it binds PS in a calcium-independent manner [1]. The C2 domain of MFG-E8 is mainly responsible for interacting with PS in a lipid membrane. Here by use of molecular dynamics simulations with its NMR structure as the starting point, we obtained a most representative membrane bound structure of the domain. We then obtained the orientation and insertion depth of the C2 domain by fitting the x-ray reflectivity from a lipid monolayer adsorbed with the C2 domain using this membrane bound structure. The resolved membrane bound configuration of the domain enabled us to identify key hydrophobic and basic residues involved in enhancing the overall binding in addition to the putative PS binding groove. Our structural model suggested four modes of relatively independent interaction between the C2 domain and the lipid membrane, namely PS-specific, hydrophobic, electrostatic, and mechanical. Using a binding assay based on its intrinsic fluorescence, we validated the existence of these three models, and found the PS-specific interaction to be the main contributor to the overall binding energy. These four modes of interaction could explain the different PS specificity of the domain observed across different model systems. We further revealed a soft sigmoidal response

of the C2 domain to surface PS density. We believe the unique combination of these four modes of interaction might give the MFG-E8 C2 domain additional membrane selectivity in addition to its specificity to PS.

5.2 Computational and Experimental Studies

5.2.1 MD study of the C2 domain interaction with a lipid bilayer

As protein configuration could change dramatically upon binding to membrane, we used the lowest energy state NMR structure [1] as the starting structure to conduct molecular dynamics simulations to find a membrane-bound state of the domain. The lipid system used in our study is made up of 70 mol% 1-pamiltoyl-2-oleoyl-sn-glycerol-phosphacholine (POPC) and 30 mol% 1-pamiltoyl-2-oleoyl-sn-glycero-phosphatidylserine (POPS), henceforth denoted as 7:3 POPC:POPS. We used the highly mobile membrane mimetic (HMMM) method to speed up the docking event [2]. Five HMMM trajectories were simulated, each for 200 ns. Within 10 ns from the start of simulation, the domain was docked to the membrane for all five trajectories. Then the root mean square distance (RMSD) of the C2 domain in reference to the NMR structure evolved to stabilize around 4.5 Å after 50 ns, after which nine starting frames with different protein orientations were selected for extending the tails and simulated for an additional 200 ns. For all nine extended tail trajectories, we aligned the protein in the plane horizontal to the membrane, calculated a mean structure based on the aligned protein structures, and subsequently located the frame that most resembled the mean structure by minimizing RMSD between the mean structure

and the protein structure in the MD frame. We refer to this new protein structure as MD-optimized.

This MD-optimized structure, seen in **Figure 5.1a**, captured the mean configuration of the protein with the membrane. This membrane-bound domain structure suggested four major modes of interactions of the domain with the lipid membrane. The putative PS binding groove [1], between Loop 2 and 3 as outlined by R79, R146 and K45 indeed faces the membrane, suggesting PS-specific recognition. The tryptophan and phenylalanine residues (W26, W33, and P31) on Loop 1 anchors into the membrane, suggesting hydrophobic interactions, while another two arginine residues (R79, R146) and one lysine residue (K24) could form contacts with the membrane headgroup, suggesting electrostatic interactions. The shallow insertion of the protein into the membrane would also entail work against the stretching modulus of the elastic lipid membrane. As seen in **Figure 5.1b**, the teal colored MD-optimized structure shows significant rearrangements in the extended loops from the lowest energy state NMR structure. The major differences are in the arrangements of the peripheral loops, and thus result in significant redistribution of the electron density profile. This MD-optimized structure differed from the original NMR structure by an RMSD of 5.28 Å, which is around 1.8 times that of the equilibrated state, as seen in **Figure 5.1c**. As the flexible peripheral loops mediate the interaction between the domain and the lipid membrane, significant conformational changes in the peripheral loops upon contacting the membrane are not surprising. As seen later, this MD-optimized structure fits the XR data much better, suggesting it is more representative of the bound state of the domain than the solution structure revealed by NMR.

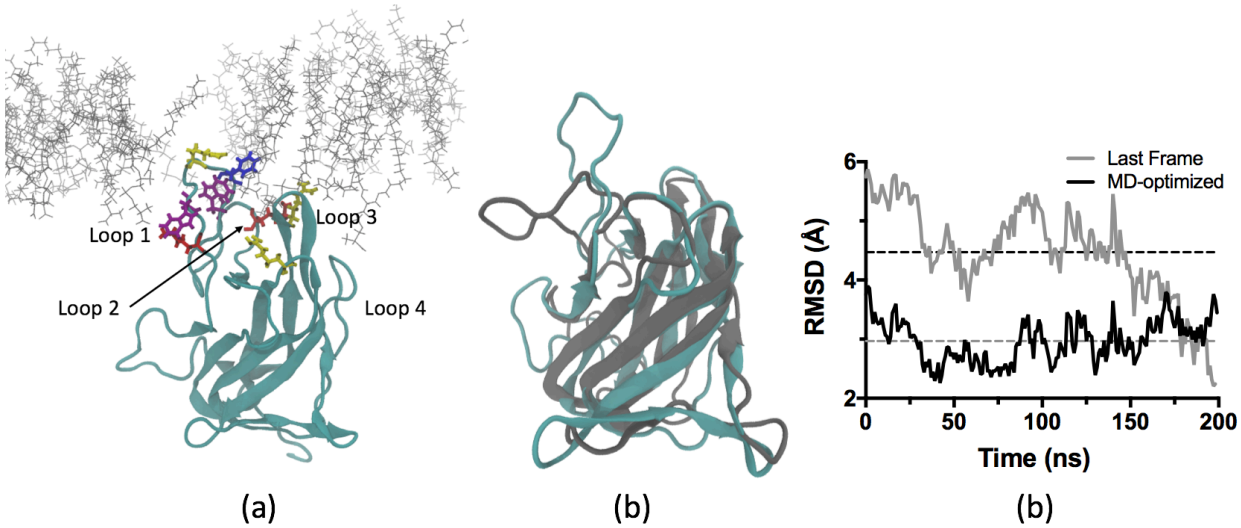

Figure 5.1: C2 domain membrane bound structure from MD simulations. (a) C2 domain bound to a bilayer lipid bilayer of 7:3 POPC:POPS in the MD-optimized frame. Only one leaflet of the membrane was shown. Yellow are arginine residues, namely R29 on Loop 1, R79 on Loop 3, and R146, which is not on a major loop. Red are lysine residues, namely K24 on Loop 1, and K45 on Loop 2. Purple are tryptophan residues, including W26 and W33 on Loop 1. Blue is the phenylalanine residue, P31, on Loop 1. The putative binding [1] site is outlined by R146 and R79 on Loop 3 with K45 on Loop 2. (b): Aligning of the C2 domain NMR structure with the membrane MD equilibrated bound structure. The teal colored one is the MD-optimized structure, while the gray colored one is the NMR structure. The RMSD between these two structures are calculated to be 5.28 Å. (c): The RMSD of a representative trajectory sampled every nanosecond, using either the last frame as the reference, or the MD-optimized strcture as the reference.

5.2.2 Study of the C2 domain adsorbed to lipid monolayer using XR

XR data were collected from lipid monolayers composed of 70 mol% 1-stearylol-2-oleoyl-sn-glycero-3-phosphacholine (SOPC) and 30 mol% 1-stearol-2-oleoyl-sn-glycero-3-phosphatidylserine (SOPS) (7:3 SOPC:SOPS) at varying surface pressure with the C2 domain in the subphase. The SO lipids were used to reduce damage of the film by x-ray [3]. A total of four films were studied, with various protein concentrations and surface pressures, as summarized in **Table 5.1**. Eight datasets with fifteen XR curves were collected for the 7:3 SOPC:SOPS system

on HBS buffer (10 mM HEPES, 150 mM NaCl, pH 7.5). Originally, we tried fitting the XR data using the solved NMR structure [1], which produced a poor visual fitting, and very high reduced χ^2 values, with lowest at 23.76 (see **Figure 5.2**). As significant rearrangements of the peripheral loops were observed upon the C2 domain contacting the membrane, associated with major redistribution in its electron density profile, we deduced that the MD-optimized structure, as described in the previous section should be more suitable.

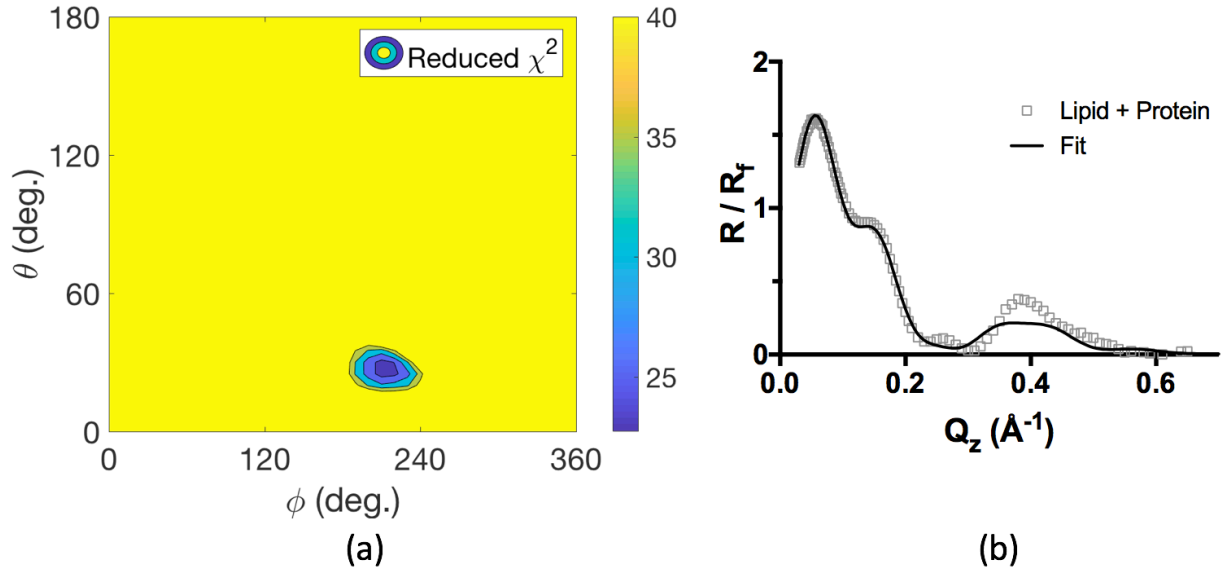

Figure 5.2: Fitting x-ray reflectivity using the NMR structure. (a) χ^2 distribution over (θ, ϕ) pairs. This is the minimized χ^2 over the parameter space while fixing (θ, ϕ). The smallest χ^2 on the map was 3918, which is about 4 times that of the same dataset using the RMSD-minimized structure obtained from MD simulations.

As seen in **Figure 5.3a**, the representative Fresnel normalized reflectivity against the z-component of the momentum transfer ($Q_z = 4\pi/\lambda$) from the monolayer with protein adsorbed is starkly different from that of the lipid monolayer without protein. The contrast between the lipid only and lipid + protein film seen here is much larger than the Tim4 [4]. The interference fringes

in the reciprocal space seen for the protein-adsorbed system were indicative of a protein layer beneath the lipid monolayer. To fit XR data from the protein-adsorbed film, the protein orientation, defined by the angle of rotation, θ, about the principal x-axis, and the angle of rotation, ϕ, about the principal z-axis (**Figure 5.3b**), the insertion depth of the domain, and the surface number density of the protein were fit to the data. A detailed description of the fitting procedure is included in Chapter 3.

Film #	C2 (nM)	Pressure (mN/m)	θ(°)	ϕ(°)	Head (Å)	Tail (Å)	Coverage	Insertion (Å)
1	100	30	32, 31	295, 292	12.58, 12.55	13.20, 13.22	0.07, 0.07	8.27, 8.30
2	200	20	34, 34	289, 293	11.07, 10.71	11.98, 12.40	0.22, 0.21	8.08, 8.07
		25	33, 33	291, 291	11.20, 11.26	13.11, 12.98	0.18, 0.19	7.75, 7.75
		30	35, 34	290, 293	12.01, 11.98	13.54, 13.60	0.15, 0.15	7.69, 7.76
3	1000	30	35, 35	290, 290	11.56, 11.50	13.47, 13.44	0.21, 0.21	8.00, 8.01
4	1000	20	32, 33	295, 293	11.92, 11.99	11.01, 11.00	0.24. 0.24	8.57, 8.55
		25	31, 32	291, 292	12.02, 12.03	12.58, 12.59	0.21, 0.20	8.25, 8.22
		30	30, 35	289, 292	12.30, 12.48	13.50, 13.35	0.17, 0.17	7.94, 7.88

Table 5.1: XR fitting results. For Film 2 and 4, the surface pressure was increased from 20 to 25, then to 30.

As seen in the representative reduced χ^2 map in **Figure 5.3c**, there was a single (θ, ϕ) well where the domain fits the best. The smallest reduced χ^2 value for this particular film was 12.29, similar to previously reported value for using the model [3]. For all sixteen XR curves, the fitted orientation came out to be $(\theta = 33 \pm 1°, \phi = 291 \pm 4°)$. On the other hand, the coverage of the domain, as calculated by the footprint of the protein (which is 5.84 nm^2 at the optimal orientation, see **Figure 5.4**) multiplied by its number density, showed a clear dependence on surface pressure (**Figure 5.3e**). The insertion depth seemed to weakly depend on the surface pressure (**Figure 5.3e**), and it was centered around 7.88 Å. The tail length increased with increasing surface pressure, consistent with a more upright lipid packing (**Figure 5.4**). The corresponding electron density profile of the lipid monolayer adsorbed with protein, with or without thermal smoothing,

is seen in **Figure 5.3d**. The double humps in the electron profile were necessary for capturing all the features in the XR curve. Crucially, the fitted lipid parameters, found in **Table 5.1** were similar to SO monolayers in previous studies [3, 4].

The resolved orientation and insertion in relation to a lipid membrane is depicted in **Figure 5.3e**. The interaction configuration of the domain bears close resemblance to the MD-optimized structure from the simulation, with three significant features for the binding configuration. First, the putative binding site, the cleft outlined by K45 on Loop 2, and R79 and R146 on Loop 3, is in close proximity to the membrane. Second, two tryptophan (W26 and W33) and one phenylalanine (P31) residues on Loop 1 were inserted in the headgroup region. Third, two basic residues, the lysine residue on Loop 1 (K24) was at the surface of the headgroup region, and the arginine residue on the same loop (R29) inserted into the headgroup region.

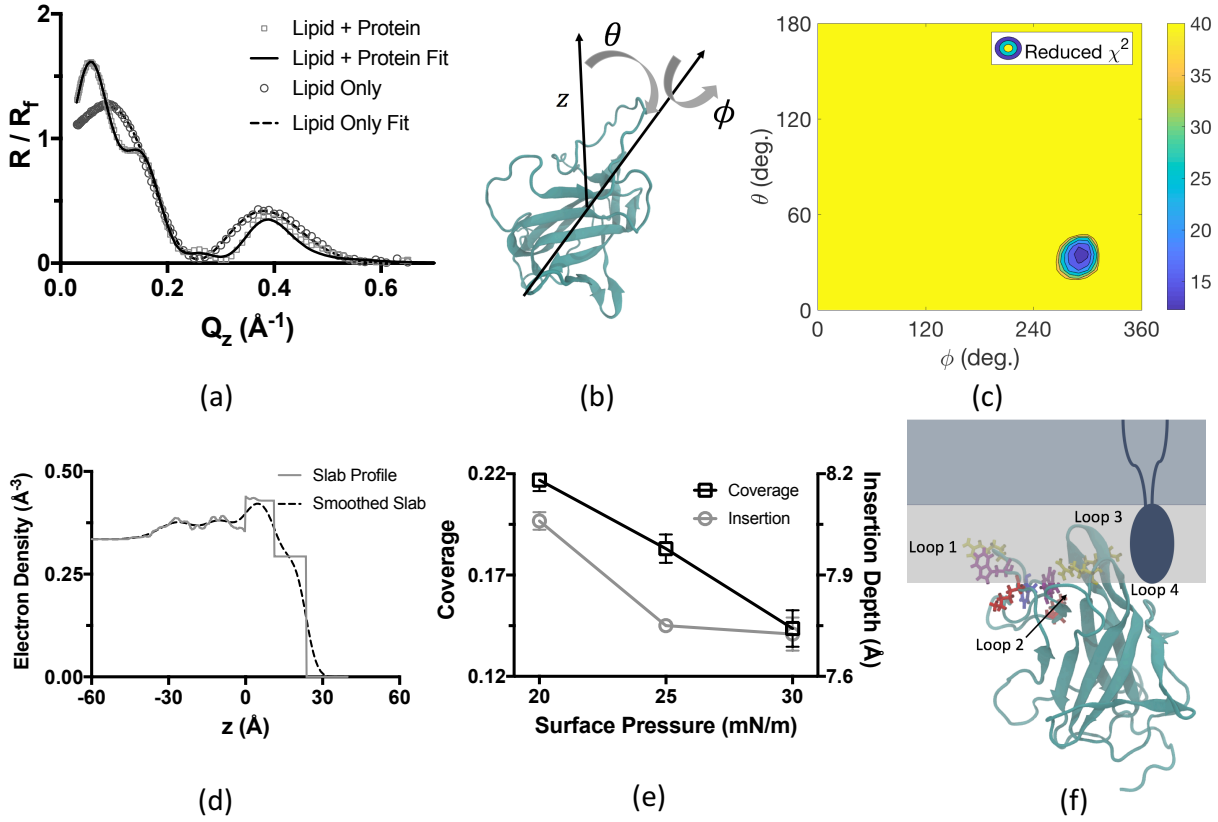

Figure 5.3: XR analysis of the protein bound orientation. (a): Representative XR data from a lipid-only membrane, and a lipid membrane with protein adsorbed. The lipid system was 7:3 SOPC:SOPS. Here the protein concentration was 1 μM, and the surface pressure was kept at 30 mN/m. (b): Definition of the protein orientation. θ is the angle of rotation angle about the x-axis, and ϕ is the angle of rotation about the z-axis. (c): A representative reduced χ^2 map over the two-dimensional (θ, ϕ) grid. Here the lowest reduced χ^2 was 12.29, about one half of that obtained from fitting the NMR structure to the same dataset (**Figure 5.2**). (d): Electron density profile before and after smoothing with thermal fluctuation for the lipid membrane with adsorbed protein at its best fit orientation and level of insertion. (e): Dependence of area coverage and insertion depth on surface pressure from fitting to the XR data sets with 1 μM protein in the subphase. (f): Best-fit configuration of the protein. The protein and lipid were drawn to scale. The color scheme is the same as **Figure 5.1a**. Yellow are arginine residues, namely R29 on Loop 1, R79 on Loop 3, and R146, which is not on a major loop. Red are lysine residues, namely K24 on Loop 1, and K45 on Loop 2. Purple are tryptophan residues, including W26 and W33 on Loop 1. Blue is the phenylalanine residue, P31, on Loop 1. The putative binding [1] site is outlined by R146 and R79 on Loop 3 with K45 on Loop 2.

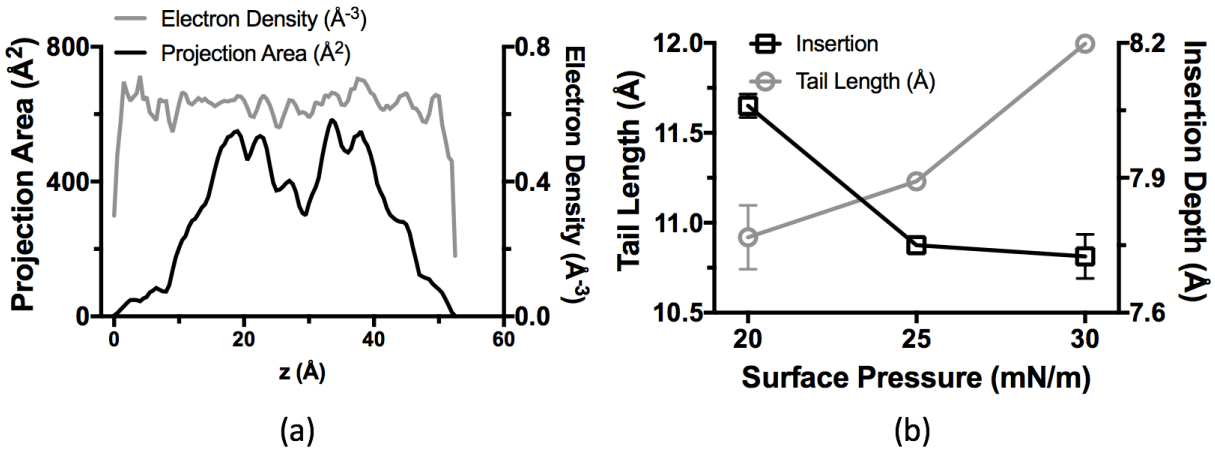

Figure 5.4: C2 domain cross sectional area tail length. (a) C2 domain area profile along the z-axis at ($\theta = 33°, \phi = 291°$). (b): C2 domain electron density protein at ($\theta = 33°, \phi = 291°$). (c): Fitted lipid tail length and insertion depth of the C2 domain as a function of surface pressure for 7:3 SOPC:SOPS vesicle system, with 1 μM C2 protein in the subphase.

5.2.3 Quantification of the C2 domain binding to PO lipid vesicles via its intrinsic tryptophan fluorescence spectral shift

So far with the structural studies, we have found that two tryptophan (W26 and W33) and one phenylalanine (P31) residues on Loop 1 that could insert into the membrane headgroup region. This shallow insertion of hydrophobic residues into the lipid membrane could serve to guide and anchor the protein to the correct interaction geometry [5]. When tryptophan residues move from the more polar water environment into the membrane headgroup/tailgroup interfacial region, its fluorescence spectrum displays a blue shift and an increase in intensity, which could be utilized to measure the bound fraction of the burying entity [6]. Using 7:3 POPC:POPS vesicles, we have quantified the increase in intensity to be 73%, and the blue shift to be 4 nm comparing the fully bound domain with the free domain for 7:3 POPC:POPS vesicles. The PO lipids are used to

match previously conducted experiments with Tim3 (**Chapter 3**), and to cross compare with published data on Tim4 [4]. As the spectra of a given protein sample mixed with lipid would be a linear combination of the bound and the unbound spectra [4, 6], we made use of this metric to quantify the binding affinity of the C2 domain to various vesicle samples.

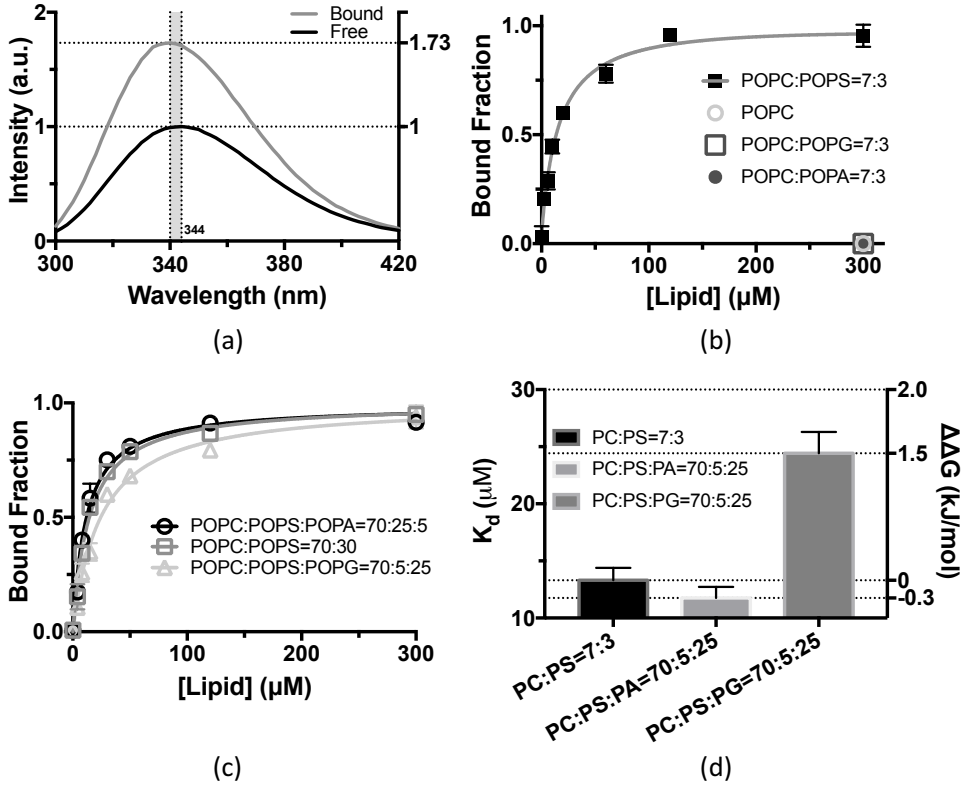

Figure 5.5: binding of the C2 domain to PO lipid vesicles studied by its intrinsic tryptophan fluorescence spectral shift. (a): Tryptophan fluorescence spectral shift of the MFG-E8 C2 domain upon binding to vesicles. Here, the C2 domain concentration was kept at 100 nM, and lipid concentration was 0 μM for the free protein spectrum, and 300 μM for the bound spectrum. The spectra were normalized to the maximum of the free protein spectrum for easier comparison. Upon binding to lipid vesicles, the protein spectra showed a 70% increase in intensity, as well as a 4 nm blueshift. **(b)**: No detectable binding was observed for the POPC only, 7:3 POPC:POPG, nor 7:3 POPC:POPA vesicles. The bound fraction of C2 domain for the 7:3 POPC:POPS vesicle system as a function of lipid concentration was fit to a dissociation constant, K_d, of 13.3 ± 1.1 μM. **(c)**: Dependence of the bound fraction of the C2 domain on lipid concentration. The fitted K_d of 7:3 POPC:POPS, 70:5:25 POPC: POPS: POPA, and 70:5:25 POPC: POPS:POPG systems were 13.3 ± 1.1, 11.75 ± 1.0, and 24.42 ± 1.9 μM, respectively. **(d)**: Change in binding free energy using the 7:3 POPC: POPS system as reference, $\Delta\Delta G = RT\ln(K_d/K_d^0)$.

To better understand the interaction of the C2 domain with biological membranes, we studied the binding affinity of MFG-E8 with vesicles composed of mixtures of POPC, POPS, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylglycerol (POPG), and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidic acid (POPA), namely vesicles composed of 7:3 POPC:POPS, 7:3 POPC:POPG, and 7:3 POPC:POPA. Contrary to micelle studies [1, 7], we found that the C2 domain did not bind to neat POPC vesicles [8], nor did it show detectable affinity to POPC mixed with 30 mol% of either POPG or POPA (see **Figure 5.4b**). In contrast, the C2 domain showed strong affinity to 7:3 POPC:POPS vesicles. For the 7:3 POPC:POPS system, the bound fraction of protein can be fit to a one-site specific binding model with a dissociation constant of 13.3 ± 1.1 μM, as seen in **Figure 5.6b.**

Both our structural studies using molecular dynamics simulations and XR analysis pointed to the likely involvement of the basic residues, in particular lysine (K24) and arginine (R29), on the insertion loop in strengthening the interaction. We therefore tested the affinity of the C2 domain to vesicles with a base level of POPS (5%) and an additional 25% of either the negatively charged POPG or POPA, so that the total charge density was maintained at 30%. As seen in **Figure 5.6c**, all three kinds of vesicles bind well with the C2 domain, a stark contrast to binary POPC/POPG or POPC/POPA vesicles (see **Figure 5.4b**). The dissociation constants of the 70:5:25 POPC:POPS:POPA system was very close to that of the 7:3 POPC:POPS system, while that of the 70:5:25 POPC:POPS:POPG was found to be slightly lower, corresponding to an increase in binding free energy of 1.5 kJ/mol.

5.2.4 Sigmoidal response of MFG-E8 to PS-containing membranes

Our data suggested that the C2 domain has a PS specific mode of interaction, which presumably originated from the PS binding groove, and an electrostatic part that senses negative charge in the membrane. These two modes of interaction would give rise to a C2/PS stoichiometry larger than one. To test this hypothesis, we studied the dependence of binding on the surface density of POPS in vesicles composed of binary mixtures of POPC and POPS. As seen in **Figure 5.6**, the dependence exhibited a sigmoidal behavior, indicating the C2 domain interacting with multiple PS simultaneously. Using the Hill model to fit the bound fraction dependence on PS%, i.e.

$$Bound\ Fraction = \frac{(PS\%)^h}{K_d + (PS\%)^h},$$

where h is the Hill coefficient, $h = 1.78 \pm 0.19$ was obtained.

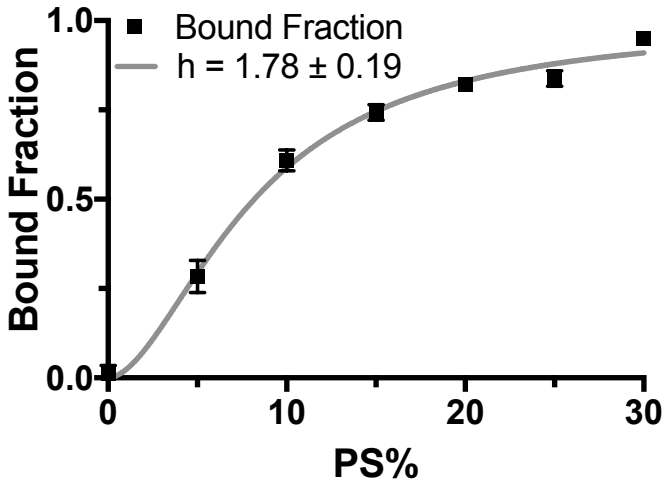

Figure 5.6: Dependence of the C2 domain binding on PS surface density. Here the lipid concentration was kept at a constant 300 μM. Each data point represents the mean and standard deviation of at least three repeats. Fitting the curve to the Hill model gave a Hill coefficient of 1.78 ± 0.19.

5.3 Discussion

Resolving the membrane-bound structure combining MD simulations and XR methods. Fitting the XR using the NMR structure did not yield meaningful results, suggesting that the NMR structure poorly captured the structure of the protein when interacting with the membrane. One the other hand, the MD-optimized structure proved to be adequate in explaining the observed membrane sensitivity of MFG-E8 and was critical for fitting the orientation, insertion, and the coverage of the C2 domain to the XR data. The MD-optimized membrane-bound structure in **Figure 5.1a**, and the x-ray resolved structure in **Figure 5.2f** were very similar, and thus cross-validated well. Both structures pointed to three relatively independent modes of interaction between the C2 domain and lipid membranes: (1) the putative PS binding site between Loop 2 and Loop 3 in close proximity to the lipid headgroup region, (2) the basic residues, especially K24 and R29 on Loop 1, which could mediate electrostatic contacts with negatively charged lipid headgroups, and (3) the hydrophobic residues, especially W26 and P31 on Loop 1 that could anchor the protein and mediate additional hydrophobic interactions.

Major rearrangement of the peripheral loops upon C2 domain binding to the membrane, such as the displacement of Loop 2, as seen in **Figure 5.1b**, could have prevented successful fitting using just the NMR structure. The rearrangements might also explain the two-stage kinetics in C2 domain binding to vesicles observed earlier [8]. The insertion of the loops, especially Loop 1 and Loop 3 into the headgroup region, could give the C2 domain a sensitivity to lipid packing density. As seen in **Figure 5.2e**, the surface area coverage of the C2 domain showed a linear decrease with increasing surface pressure, dropping from 22.3% at 20 mN/m to 14.4% at 30 mN/m. Therefore, one should also expect a dependence of the C2 domain insertion on membrane

packing density, and a selectivity for more loosely packed membranes, such as lipids with more unsaturated bonds, or micelles with higher curvatures.

Four modes of interaction between the C2 domain and PS. The transition of tryptophan residues from the polar water environment to the relatively less polar lipid membrane would entail an increase in its tryptophan fluorescence, and a blueshift [6]. Both the MD-optimized structure and the XR fit structure pointed to membrane insertion of two tryptophan residues on Loop 1, namely W26 and W33 into the membrane (**Figure 5.1a**, **Figure 5.2f**). The observed tryptophan fluorescence spectral shift, with 73% increase in intensity, and 4 nm blueshift, further supported both structures. The intensity increase and the extent of the blueshift observed here were considerably larger than the 10% increase and 2 nm blueshift observed for vesicles composed of bovine brain and egg yolk lipids [8]. The cause of such a discrepancy remained unclear to the author.

As seen in **Figure 5.5a**, the MFG-E8 C2 domain did not bind to neat POPC vesicles, nor did it bind to binary mixtures of POPC with either 30 mol% POPG or 30 mol% POPA. Since all lipids have the same tailgroup, and all of them were at their fluidic state, our data suggested a high level of specificity towards the PS headgroup. This result also agreed with previous report that the C2 domain did not bind to vesicles containing 99:1 PC:PG [8]. Conversely, the C2 domain did bind to PG in micelles [1, 8]. This PG affinity discrepancy between vesicles and micelles likely arose from alterations in lipid packing and curvature between vesicles and micelles [9]. The binding affinity has been demonstrated to increase with increasing membrane curvature [8].

As large unilamellar vesicles likely better resemble apoptotic cells and vesicular debris, the MFG-E8 C2 domain would indeed be specific to the PS headgroup under physiologic conditions.

PA and PG alone could not recruit the C2 domain to the membrane surface (**Figure 5.4b**). However, with the help of 5% PS in the vesicles, membranes with 25% PG or PA making up the rest of the surface charge density were able to facilitate significant binding, as well as did an additional 25% of PS (**Figure 5.4c**). These three lipids are all negatively charge, thus the 25% PG, PA, or PS in addition to the 5% PS likely mediates the electrostatic interactions between the membrane and the basic residues on the C2 domain, such as the K24 and R29 on Loop 1 (**Figure 5.1a**, **5.2f**). In the 7:3 POPC:POPS vesicle system, the PS would serve two roles: one as a ligand for the PS-specific binding site, and the other as a negatively charge contact for the basic amino acid residues outside of the binding site.

The effects of increased surface pressure in the monolayer experiments seemed to be two-fold. For the 200 nM protein case, increasing the surface pressure from 20 to 30 mN/m led to a decrease of the surface coverage of the protein from 22% to around 15%, as well as a decrease in the insertion depth of the protein from 8.08 to 7.73 Å. Both effects would be consistent with the increased work needed to expand the membrane to accommodate the inserted part of the protein.

Our data support the picture of having four modes of interaction working synergistically to recruit the C2 domain to the membrane. First, the PS binding groove appeared to be in close contact with the membrane, which likely provides the essential interaction as the lack of PS in the membrane results in no binding. Second, as seen from both the mean membrane-bound structure from MD simulations and the fitting of the XR data, two tryptophan residues and one

phenylalanine residue on Loop 1 are at the membrane headgroup region, which could mediate hydrophobic interactions with the membrane. Third, one lysine and one arginine residue on the loop are also in close contact with the membrane, which could further provide extra electrostatic contacts to increase the binding affinity. Fourth, the shallow insertion of the protein into the membrane would entail work against the stretching modulus of the membrane, resulting in a mechanical component in the total binding energy.

Characteristics of the binding of MFG-E8 C2 domain to PS. The affinity of the C2 domain to PS is rather strong, and MFG-E8 competes potently with blood coagulation factors V and VIII for phosphatidylserine [10]. As such, bovine MFG-E8 has been used as a calcium-independent imaging agent for phosphatidylserine [11]. MFG-E8 binding is also strong compared to another apoptosis regulator, T cell immunoglobulin mucin 4 (TIM4). The dissociation constant to 7:3 POPC:POPS vesicles of the C2 domain, at 13.3 ± 1.1 μM lipid (**Figure 5.4a**), is one third of that of the IgV domain of TIM4 [4]. This strong affinity is partly a result of the synergistic effect of the four modes of interaction present for the C2 domain. MFG-E8 is soluble, whereas TIM4 is grafted to the plasma membrane of phagocytes. The higher affinity of MFG-E8 for the membrane might be necessary to effectively bridge macrophages with PS-exposing apoptotic cells and vesicular debris, as this bridging interaction is a more difficult three-body interaction.

The soft sigmoidal response to PS surface density, as seen in **Figure 5.6**, might give MFG-E8 a mechanism for targeting more efficiently higher level of PS exposure, while ignoring low levels of PS exposure. A similar sigmoidal response was discovered for TIM4, with a Hill coefficient of 2.40 ± 0.08 [4]. In addition, the hydrophobic insertion, or the insertion of other residues, could

give the C2 domain capability to select more fluidic or more loosely packed membranes, characteristic of the plasma membrane of apoptotic cells [12, 13]. The four modes of interaction between the C2 domain and PS-containing membrane give the C2 domain a four-dimensional parameter space to differentiate from the other PS binding proteins, which could help explain the subtle yet critical differences between the similarly globular PS binding domains of MFG-E8, the Tim proteins, and the other C2 domain bearing proteins such as blood coagulation factor V and VIII and PKC_{α} [14].

5.4 Conclusion

Here we have elucidated the configuration of the C2 domain of MFG-E8 upon its interaction with the membrane using a combination of MD simulations and XR methods. The interaction configuration from the computational model agreed well with that obtained from experimental XR data, while the fitting to the XR data was only possible with the MD-optimized membrane-bound native state. The resolved membrane-bound configuration of the C2 domain suggested three major modes of interaction between the domain and the PS-containing membrane, namely 1) a PS-specific binding site, 2) electrostatic contacts, and 3) insertion of hydrophobic residues, which are confirmed by binding studies using PO lipid vesicles. XR results further demonstrated a preference of the C2 domain to lower surface pressure. Binding studies of the C2 domain with PO lipid vesicles results demonstrated that the PS-specific binding site is necessary for binding, which gives the C2 domain its PS specificity, while the electrostatic interaction contributes significantly to additional binding energy. Future studies should focus on elucidating the

importance of the lipid packing density sensitivity of the C2 domain and its possible involvement in the regulation of MFG-E8 function.

5.5 Methods and Materials

Lipids and chemicals. Analytical grade 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), $CaCl_2$, NaCl, and urea and solvents including chloroform, methanol, acetone were procured from Fisher Scientific (city, state). Lipids including POPC, POPS, POPG, POPA, SOPC, SOPS were obtained from Avanti Polar Lipids (Alabaster, AL). The buffer in most studies was 150 mM NaCl plus 10 mM HEPES at pH 7.5.

MD simulations. All atom MD simulations were performed as described in Chapter 3, Section 3.1. Five trajectories with the C2 domain 20 nm away from the membrane and different protein orientations were simulated for 200 ns. Nine membrane bound states with different protein orientation were selected to extend the tails using the CHARMM-GUI HMMM protein/membrane builder, and simulated for another 200 ns.

XR experiments. XR experiments were performed at ChemMatCARS of Sector 15ID at the Advanced Photon Source of Argonne National Laboratory. A custom trough with a movable barrier and surface pressure control loop was used to control the surface pressure of the lipid monolayer. Surface pressure was controlled at three values, 20, 25, and 30 mN/m. For each film, a reflectivity scan on the lipid-only system was performed prior to injecting the C2 domain to the subphase. The subphase volume was 70 mL, while the protein concentration was 80, so for the 1

μM protein case, at most 0.875 mL of protein was injected. Following the injection, an equal amount of buffer was removed from the subphase to restore the initial liquid level and 2 hours of equilibration time was allowed. Then, XR scans were taken within a Q_z range of 0.018-0.65 $Å^{-1}$. Two curves were gathered on a given film, which were generally quite similar, and averaged.

Production of the MFG-E8L C2 domain in *E. coli.* The cDNA for the *Mus musculus* MFG-E8L C2 domain containing a hexa-histidine tag at the C-terminus in the pET29b vector was a gift from Dr. Sreekanth Rajan and Dr. Yoon Ho Sup at Nanyang Technological University, Singapore. The plasmid was transformed into *E. coli* BL21 (DE3) cells for recombinant protein expression. Expression and purification of the MFG-E8L C2 domain was performed essentially as described previously [23]. Briefly, transformants were selected on a LB agar plate containing 30 μg/ml kanamycin and a single colony was inoculated into a 50 mL LB + kanamycin starter culture and grown at 37 °C in a shaker at 220 rpm for 15 hr. 20 mL of the starter culture was inoculated into 1 L of LB media containing kanamycin and grown to an OD_{600} ~0.7. At this point, the culture was transferred to a room temperature (~25 °C) shaker and cells were induced to express the MFG-E8L C2 domain via the addition of 1 M Isopropyl β-D-1-thiogalactopyranoside (*IPTG*) to a final concentration of 1 mM. After 4 hr of growth, the culture was harvested via centrifugation at 6,000 g for 10 min in a Beckman JLA-8.1 rotor. After removing the supernatant, cell pellets were frozen at -80 °C.

Cell pellets were resuspended at 4 °C via stirring with 10 mL/g of cell pellet in suspension buffer (20 mM Na–PO_4, pH 7.3, 0.5 M NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF) and a pinch of DNAse). Cell lysis was performed by passing the resuspension through a

microfluidizer (EmulsiFlex-C5 homogenizer, AVESTIN) three times. Cell lysate was centrifuged at 22,000 rpm for 30 min using a Beckman JA-25.50 rotor. The supernatant was brought up to 25 mM imidazole, pH 7 and the MFG-E8L C2 domain was extracted from the supernatant via batch binding for 1 hr at 4 °C with 3 mL of Ni-nitrilotriacetic acid (NTA) agarose beads (Qiagen). Ni-NTA beads were then collected and washed with 3 bed volumes of washing buffer (20 mM Na–PO_4, pH 6.5, 0.5 M NaCl, 1 mM PMSF and 25 mM imidazole) and eluted with elution buffer (20 mM Na–PO_4, pH 6.5, 0.5 M NaCl and 0.2 M imidazole). Fractions containing the eluted protein was then pooled and desalted into 20 mM Na–PO_4, pH 6.5, 0.5 M NaCl via Econo-Pac 10DG Desalting Prepacked Gravity Flow Columns (Bio-Rad) to remove the imidazole and concentrated via Centrifugal Filter Units (Millipore) to 1 mL for two 0.5 mL injections onto a S75 gel filtration column. The MFG-E8L C2 domain elutes off of the S75 column at ~22 mL, which is much later than would be predicted based on the size of the MFG-E8L C2 domain. Protein fractions were then analyzed by 12% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and the protein concentration was calculated via the Bradford Protein Assay (Bio-Rad) or the hypothetical extinction coefficient generated by the ExPASy ProtParam tool. Typical protein yields were ~2 mg/L culture medium. The MFG-E8L C2 domain is sensitive to concentration via Centrifugal Filter Units, so protein for experiments was used at the concentration in which it came off of the S75 column (~18 μM or .35 mg/mL).

5.6 References

1. Ye, H.; Li, B.; Subramanian, V.; Choi, B.-H.; Liang, Y.; Harikishore, A.; Chakraborty, G.; Baek, K.; Yoon, H. S. NMR solution structure of C2 domain of MFG-E8 and insights into its

molecular recognition with phosphatidylserine. *Biochim. Biophys. Acta*. **2013,** *1828*, 1083–1093.

2. Ohkubo, Y. Z.; Pogorelov, T. V.; Arcario, M. J.; Christensen, G. A.; Tajkhorshid. E. Accelerating membrane insertion of peripheral proteins with a novel membrane mimetic model. *Biophys. J.* **2012,** *102*, 2130-2139.
3. Málková, Š.; Long, F.; Stahelin, R.; Pingali, S.; Murray, D.; Cho, W.; Schlossman, M. X-Ray Reflectivity Studies of cPLA2α-C2 Domains Adsorbed onto Langmuir Monolayers of SOPC. *Biophys. J.* **2005,** *89*, 1861-1873.
4. Tietjen, G.; Gong, Z.; Chen, C. H.; Vargas, E.; Crooks, J. E.; Cao, K. D.; Heffern, C. T.; Henderson, J. M.; Meron, M.; Lin, B.; Roux, B.; Schlossman, M. L.; Steck, T. L; Lee, K. Y.; Adams, E. J. Molecular mechanism for differential recognition of membrane phosphatidylserine by the immune regulatory receptor Tim4. *Proc. Natl. Acad. Sci.* **2014,** *111*, E1463–E1472.
5. Vanni, S.; Vamparys, L.; Gautier, R.; Drin, G.; Etchebest, C.; Fuchs, P. F.; Antonny, B. Amphipathic lipid packing sensor motifs: probing bilayer defects with hydrophobic residues. *Biophys. J.* **2013,** *104*, 575-584.
6. Kraft, C.; Garrido, J.; Leiva-Vega, L.; Romero, G. Quantitative Analysis of Protein-Lipid Interactions Using Tryptophan Fluorescence. *Sci. Signal.* **2009,** *2*, p14.
7. Reddy Nanga, R. P.; Vivekanandan, S.; Yoon, H. S. Expression, purification and characterization of C2 domain of milk fat globule-EGF-factor 8-L. *Protein Expr. Purif.* **2007,** *52*, 329-333.
8. Otzen, D. E.; Blans, K.; Wang, H.; Gilbert, G. E.; Rasmussen, J. T. Lactadherin binds to phosphatidylserine-containing vesicles in a two-step mechanism sensitive to vesicle size and composition. *Biochim. Biophys. Acta*. **2012,** *1818*, 1019–1027.
9. Garavito, R. M.; Ferguson-Miller, S. Detergents as tools in membrane biochemistry. *J. Biol. Chem*. **2001,** *276*, 32403-32406.
10. Shi, J.; Gilbert, G.E. Lactadherin inhibits enzyme complexes of blood coagulation by competing for phospholipid-binding sites. *Blood*. **2003,** *101*, 2628–2636.
11. Waehrens, L. N.; Heegaard, C. W.; Gilbert, G. E.; Rasmussen, J. T. Bovine lactadherin as a calcium-independent imaging agent of phosphatidylserine expressed on the surface of apoptotic HeLa cells. *J. Histochem. Cytochem.* **2009,** *57*, 907-914.

12. Schlegel, R. A.; Stevens, M.; Lumley-Sapanski, K.; Williamson, P. Altered lipid packing identifies apoptotic thymocytes. *Immunol. Lett*. **1993,** *36*, 283-288.
13. Mower, D. A.; Peckham, D. W.; Illera, V. A.; Fishbaugh, J. K.; Stunz, L. L.; Ashman, R. F. Decreased membrane phospholipid packing and decreased cell size precede DNA cleavage in mature mouse B cell apoptosis. *J. Immunol.* **1994,** *152*, 4832-4842.
14. Lemmon, M. A. Membrane recognition by phospholipid-binding domains. *Nature.* **2008,** *9*, 99–111.
15. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem*. **2005,** *26*, 1781–1802.
16. Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Leumkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pane, V. S.; Case, D. A.; Brooks, C. L. III; MacKerell, A. D. Jr.; Klauda, J. B.; Im, W. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **2016,** *12*, 405-413.
17. Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem*. **2008,** *29*, 1859–1865.
18. Qi, Y.; Cheng, X.; Lee, J.; Vermaas, J. V.; Pogorelov, T. V.; Tajkhorshid, E.; Park, S.; Klauda, J. B.; Im, W. CHARMM-GUI HMMM builder for membrane simulations with the highly mobile membrane-mimetic model. *Biophys. J.* **2015,** *109*, 2012-2022.
19. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Molec. Graphics*. **1996,** *14*, 34-38.
20. Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*. **2006,** *7*, 382.
21. Tachyon Parallel/Multiprocessor Ray Tracing System. http://www.photonlimited.com/~johns/tachyon/ (Accessed April 2017)
22. *Github.* https://github.com/ymatsunaga/mdtoolbox (Accessed April 2017)
23. Nanga, R. P.; Vivekanandan, S.; Yoon, H. S. Expression, purification and characterization of C2 domain of milk fat globule-EGF-factor 8-L. *Protein Expr. Purif*. **2007,** *52*, 329-333.

CHAPTER 6

DEVELOPMENT AND TEST OF A SOFTWARE TOOL (*XERAY*) FOR QUANTITATIVE ANALYSIS OF TOTAL REFLECTION X-RAY FLUORESCENCE (TXRF) FROM FINELY LAYERED STRUCTURES

6.1 Overview

Total reflection x-ray fluorescence (TXRF) is a widely applicable experimental technique for studying chemical element distributions across finely layered structures at extremely high sensitivity. To promote and facilitate scientific discovery using TXRF, we developed a MATLAB-based software package with a graphical user interface, named *XeRay*, for quick, accurate, and intuitive data analysis. *XeRay* allows the user to model any layered system, each layer with its independent chemical composition and thickness, and enables fine-tuned data fitting. The accuracy of *XeRay* has been tested in analysis of TXRF data obtained from both air/liquid interface and liquid/liquid interfacial studies. In an air/liquid interface study, Ca^{2+} sequestration was measured at a Langmuir monolayer of 1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphatidic acid (SOPA) on a buffer solution of 1mM $CaCl_2$ at pH 7.5. Data analysis with *XeRay* reveals that each 1 nm^2 of interfacial area contains $\mathbf{2.38 \pm 0.06}$ Ca^{2+} ions, which corresponds to a 1:1 ratio between SOPA headgroups and Ca^{2+} ions, consistent with several earlier reports. For the liquid/liquid interface study of Sr^{2+} enrichment at the dodecane/surfactant/water interface, analysis using *XeRay* gives a surface enrichment of Sr^{2+} at $\mathbf{68^{+6}_{-5}}$ $\mathbf{Å^2}$ per ion, consistent with the result published for the same dataset. This software tool has

enabled the study Ca^{2+} sequestration in the context of understanding the role of Ca^{2+} in regulating Tim3 interaction with lipid membranes (Chapter 3).

6.2 Introduction of TXRF

Finely layered structures are essential building blocks of both biological and artificial systems. Of particular interest is the distribution of chemical elements inside the layers, as it is critical in the functioning of any of such systems. On the one hand, biological membranes recruit a vast array of membrane binding proteins [1] and messenger metal ions like Ca^{2+} to participate in numerous intracellular functions and intercellular signaling events [2]. On the other hand, in industrial and scientific research, such fine layered structures serve as fundamental design patterns [3], such as thin-layered polymer films [4], silicon wafers [5], heavy metal ion extraction systems [6], and a wide range of surface catalysis and chemistry [7].

X-ray fluorescence is an ideal tool for studying such elemental distributions in finely layered structures due to its relative nondestructive and nonintrusive properties, as well as the ability of the x-ray beam to penetrate condensed matter from sub-nanometers to a few hundred micrometers in thickness [8, 9]. First described by Yoneda and Horiuchi [10], total reflection x-ray fluorescence (TXRF) has become a commonly used method for quantitatively determining the distribution of metal ions and other elements across all kinds of finely layered structures. The versatility of TXRF is manifested in the diversity of systems probed using this technique, including Langmuir monolayers [11, 12, 13], biological membranes [14], membrane binding proteins [15, 16], liquid/liquid interfaces [17], biological tissues [18, 19], environmental samples

[20], archaeological artifacts [21], just to name a few. Its increase in usage in recent decades is due in part to the advent of TXRF-capable beamlines at a number of synchrotron sources [22, 23, 24], which are characterized by brilliant and well-collimated x-ray sources that greatly enhance precision and lower detection limit. The high sensitivity of TXRF to the interfacial elements is a result of the interaction between the probing x-rays and electrons of the substrates that give rise to a refractive index slightly smaller than one, producing a critical angle for total reflection when x-ray enters from a less electron-rich medium into a more electron-rich medium. Below the critical angle, an incident x-ray beam at the interface creates a standing evanescent wave at the surface without significant penetration into the electron-rich material. Above the critical angle, the x-ray largely bypasses the interface and penetrates into the more electron rich medium. The easily tunable sensitivity of x-ray for the interface or the bulk medium forms the basis for using TXRF as a tool for detecting chemical enrichment at the interface. In addition to its high sensitivity to trace amounts of elements, TXRF has the added benefit of being largely non-intrusive and non-destructive compared to electron or molecular beams. In TXRF the excitation energy is generally kept far away from the absorption edge of the element so that the fluorescence intensity is proportional to the density of the element. Conversely, scanning the incidence energy around the absorption edge while fixing the incidence angle produces x-ray absorption near edge spectra (XANES), which provide detailed information on the valence states of the element of interest [25]. Combining these two techniques could reveal both the quantity and state of the element [26]. In this work we are concerned only with TXRF.

Given the unique capability of TXRF to quantify the important phenomenon of atomic element distribution across layered interfaces, we developed an intuitive MATLAB-based analysis

software package for TXRF to facilitate scientific discovery using this technique. To test our software package we designed an experiment intended to measure Ca^{2+} accumulation beneath an anionic SOPA monolayer in a Langmuir trough. Analysis using *XeRay* revealed an approximate 1:1 stoichiometry between the anionic SOPA headgroups and the cationic Ca^{2+} ions. To test the capability of our software for a different kind of layered system, we analyzed a dataset on Sr^{2+} sequestration at the dodecane/surfactant/liquid interface from a previously published report [17] and confirmed that our results obtained from *XeRay* are consistent with the report, which assumed zero thickness for the Sr^{2+} sequestration layer in its x-ray fluorescence analysis.

6.3 TXRF setup and operating protocol for investigating Langmuir monolayers at the air/liquid interface

To demonstrate how a TXRF experiment is carried out, here we describe our air/liquid interface experimental setup and procedure. Similar studies can be conducted on any flat interfaces involving any gaseous or condensed phases. As seen in **Figure 6.1**, the experimental setup for air/liquid experiments includes three major parts: (1) a liquid surface x-ray spectrometer system that controls the geometry and intensity of x-rays that hit a horizontal stage, (2) a Langmuir trough with stirring capability and a surface pressure control system, featuring a feedback loop composed of a Wilhelmy plate and a motorized barrier, all enclosed in an air-tight box filled with helium, and (3) an energy-resolved x-ray detector mounted via an air-tight and height-adjustable holder overhanging the surface. The airtight helium-filled box serves to minimize oxidative damage to the surface, and reduces x-ray scattering by the more electron-dense nitrogen and oxygen in the air.

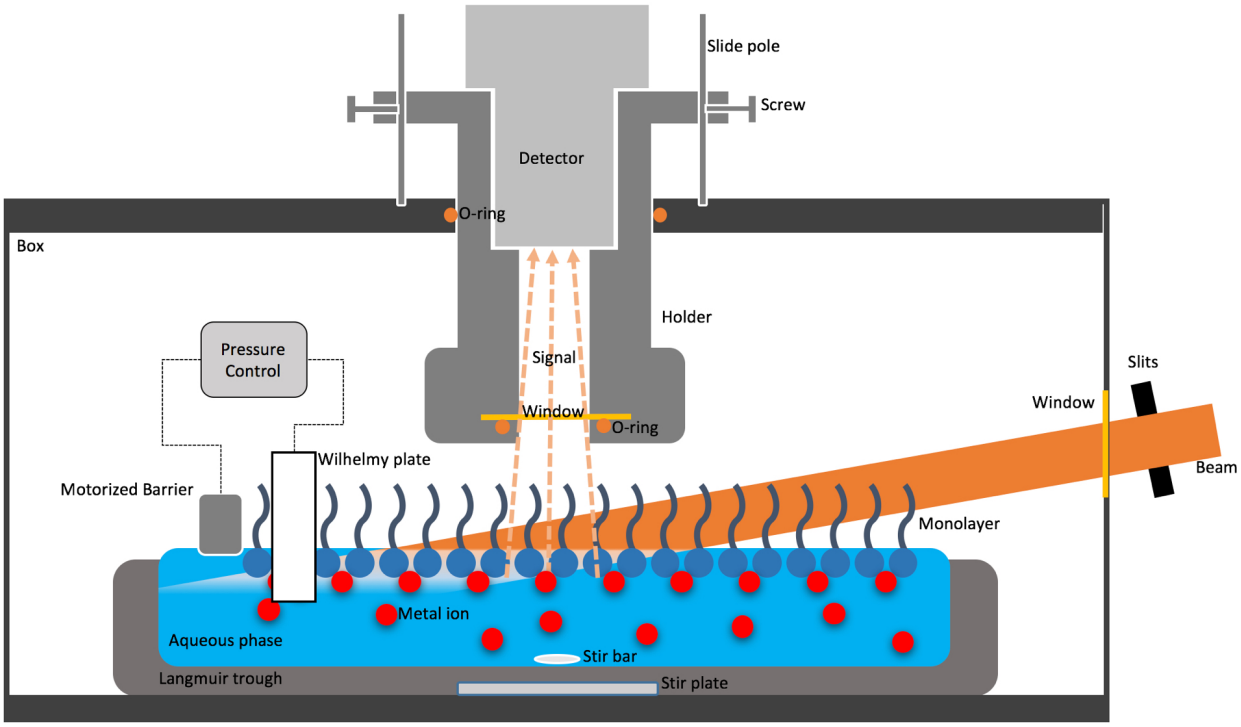

Figure 6.1: Schematic of the air/liquid TXRF setup. The detection side of the detector holder, just above the liquid surface, is sealed by a Kapton window pressed tightly onto the holder by a silica gel O-ring. The entire detector holder can slide vertically along another silica gel O-ring embedded in the box lid. The bottom part of the detector holder is a thin tube measuring 10 mm in diameter (slightly larger than the 8 mm opening on the detector), which effectively acts as a collimator to narrow down the detection area on the surface. At the top of the detector holder is a wider opening, in which the detector tube fits tightly in an upright orientation. The surface tension/pressure is monitored and controlled by a feedback loop composed of a Wilhelmy plate and a motorized barrier. For illustration purposes the barrier and Wilhelmy plate are drawn in the same plane as the beam, while they actually reside on different planes in the setup.

For Langmuir monolayers, quantification of an element bound to the water/lipid interface is of great interest to researchers [8]. First, the scale factor, which describes the excitation efficiency and particular geometry of the system, is determined via a calibration process, whereby x-ray fluorescence data are collected from a solution without a monolayer and with a known bulk concentration of the desired element. Data are then collected from the experimental system of

interest with the insoluble monolayer present, and fitted to yield the surface concentrations. A detailed description of this process is discussed in the theory section, and continued in the software implementation section.

6.4 Theory of TXRF

X-rays can excite atoms from the ground state to higher electronic states to emit fluorescence with unique energy peaks. For example, calcium has a characteristic K_α line at 3.7 keV, and K_β line at 4.0 keV, and strontium has a K_α line at 14.2 keV, and L_α line at 1.8 keV. To obtain analytical closed form solutions, the interfaces are assumed to be perfectly sharp, and all layers are assumed to be homogeneous in composition. The interaction of x-ray with each layer is, to a good approximation, described by classical optics considerations. For each layer, the refractive index is

$$n = 1 - \delta + i\beta, \tag{1}$$

where δ is the dispersion, and β the absorption. δ and β can be calculated additively from the absorption and dispersion properties of each elemental component i in the system, with the dispersion

$$\delta = \sum_i \frac{N_A}{2\pi} r_e \lambda^2 \frac{Z_i}{M_i} \rho_i \,, \tag{2}$$

and the absorption

$$\beta = \sum_i \frac{\lambda}{4\pi} \mu_i, \tag{3}$$

where N_A is Avogadro's number, r_e the classical radius of the electron, λ the wavelength of x-ray, Z_i the atomic number, M_i the atomic mass, ρ_i the mass density, and μ_i the linear mass absorption coefficient. Here we use the database of the mass absorption coefficient of 92 elements from the

periodic table with energies ranging from 30 eV to 30 keV [27]. Typically for condensed matter phases δ and β are much smaller than 1, with δ on the order of 10^{-6}, and β on the order of 10^{-9}. Readers are encouraged to turn to a textbook, like [18], for a more complete treatment.

To calculate fluorescence induced by x-ray, we adopted a previously developed matrix method [28, 29]. First we assume the system is made of N+1 (N $\geq$ 1) layers, the top, layer 0, being the infinite medium via which the x-ray arrives, and the bottom, layer N+1, being the infinite medium via which the x-ray leaves. The refraction angle between two adjacent layers are related by Snell's law

$$\theta_{n+1} = \sqrt{\theta_n^2 - 2(\delta_{n+1} - \delta_n) + 2i(\beta_{n+1} - \beta_n)}, \tag{4}$$

Deriving from Snell's law and the Fresnel relations, the incidence and reflection amplitudes between two adjacent layers are related by

$$\begin{pmatrix} A_n^t \\ A_n^r \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} A_{n+1}^t \\ A_{n+1}^r \end{pmatrix}, \tag{5}$$

where the matrix elements are given by

$$m_{11} = \frac{\theta_n + \theta_{n+1}}{2\theta_n} \exp\left(-i\frac{\pi}{\lambda}(\theta_n d_n + \theta_{n+1} d_{n+1})\right), \tag{6}$$

$$m_{12} = \frac{\theta_n - \theta_{n+1}}{2\theta_n} \exp\left(-i\frac{\pi}{\lambda}(\theta_n d_n - \theta_{n+1} d_{n+1})\right), \tag{7}$$

$$m_{21} = \frac{\theta_n - \theta_{n+1}}{2\theta_n} \exp\left(i\frac{\pi}{\lambda}(\theta_n d_n - \theta_{n+1} d_{n+1})\right), \tag{8}$$

$$m_{22} = \frac{\theta_n + \theta_{n+1}}{2\theta_s} \exp\left(i\frac{\pi}{\lambda}(\theta_n d_n + \theta_{n+1} d_{n+1})\right), \tag{9}$$

where d_n denotes the thickness of the layer n, and λ is the wavelength of the incidence beam. The relation between the amplitudes of the first layer and the last is

$$\begin{pmatrix} 1 \\ A_0^r \end{pmatrix} = \sum_{n=1}^{N} M_n \begin{pmatrix} A_n^i \\ 0 \end{pmatrix}, \tag{10}$$

where two boundary conditions are used: the amplitude of the incidence beam is assumed to be 1, and the reflection amplitude coming from the bottom layer is 0. This equation enables us to solve for all the other amplitudes for the layers in between.

The decay of the incidence beam and the excited fluorescence is characterized by a penetration depth at which the intensity decays to $1/e$ of the original. For the incidence wave at layer n, the penetration depth is

$$\Lambda_n^i = \frac{\lambda_i}{4\pi \operatorname{Im}(\theta_n)}, \tag{11}$$

where the superscript or subscript i refers to the excitation wave. For the fluorescence wave, the treatment can be greatly simplified as it has an incidence angle close to $\pi/2$, far larger than the critical angle, thus reflection can be ignored. The penetration depth of the fluorescence beam is

$$\Lambda_n^f = \frac{\lambda_f}{4\pi\beta_n^f}, \tag{12}$$

where the superscript or subscript f refers to the fluorescence wave.

Having calculated the amplitudes of the transmission and reflection wave at each layer, and the penetration depths, the intensity of the excited fluorescence for each layer at position z is

$$I_n(z) = \exp\left(-\sum_{m=0}^{n-1} \frac{d_m}{\Lambda_m^i}\right) \cdot \left|A_n^t \exp(ik\theta_n z) + A_n^r \exp(-ik\theta_n z)\right|^2 \cdot \exp\left(-\frac{z}{\Lambda_n^i}\right), \tag{13}$$

The fluorescence intensity originating from layer n is

$$F_n = C_0 c_n \int_{z_n}^{z_{n+1}} I_n(z) \exp\left(-\frac{z}{\Lambda_n^f}\right) dz, \tag{14}$$

where C_0 is the scale factor that characterizes the geometry of the setup and the quantum efficiency of the excitation, c_n the concentration of the excited element, and z_n the position at which layer n begins. The total fluorescence intensity is then calculated as a sum over all layers

$$F_0 = \sum_{n=1}^{N} F_n. \quad (15)$$

Note that the top layer (n=0) is ignored in the sum, because even if it were to contain any fluorescing agents, the x-ray would not be able to penetrate through this infinite layer.

6.5 Data analysis in *XeRay*

XeRay is an x-ray fluorescence analytical package with a graphical user interface (GUI) under the MATLAB environment. The MATLAB platform was chosen because of its popularity with science applications in academia and industry, cross-platform support, backward compatibility with earlier functions, and superb data manipulation capabilities. *XeRay* allows for individual spectral inspection of different elements, fits emission peaks to single or double Gaussian or Lorentzian line shapes, and fits the integrated fluorescence intensity over an incident angle range. If there exists a parallel pool of processors for MATLAB, which can be easily set up, computations in the fitting processes are efficiently parallelized into this pool of CPU cores. The program uses a combination of a brute force search and a trust region reflective (TRR) method [20] in the MATLAB *lsqnonlin* optimization function to search through the parameter space to obtain the fits and χ^2 values, defined as

$$\chi^2 = \sum_i \frac{(\hat{y}_i - y_i)^2}{\delta y_i}, \quad (16)$$

where $\hat{y}_i$ is the fitted fluorescence signal, y_i is the signal, and δy_i is the error of the signal.

XeRay uses the chemical formula for each layer to calculate the refractive index. If a chemical formula is not provided, the absorption term is assumed to be 0, and the dispersion term is approximated by

$$\delta = \frac{\lambda^2 r_e}{2\pi}\rho_e, \tag{17}$$

where ρ_e is the electron density. So using a chemical formula that matches the composition of a layer would improve the accuracy of fitting results.

The parameter space determining a fluorescence intensity *vs.* angle curve is composed of N+2 parameters for a system made of N (N $\geq$ 2) layers. *XeRay* assumes that the thickness of each layer is given; thickness information could come from hypotheses, design specifications, and physical measurements such as those from x-ray reflectivity and electron microscopy. The first three parameters are the aforementioned scale factor, a background noise, and an angle offset value. The purpose of fitting the angle offset is to account for the systematic error from the motor system moving to the desired angles. The background comes from the presence of the element of interest in places not considered by the modeled system, such as in the absorbers used to attenuate the beam, and in the x-ray collimator used before the detector. The background is assumed constant, because in our measurements the background is usually scattered around some mean value within the small energy window where the peak occurs, whether experimental or control experiments are being conducted. The background can either be fitted, or be fixed to the average fluorescence intensity from a control sample without the element of interest across the same range of angles, which is similar to the blank subtraction method used in this iron x-ray fluorescence study [26]. The remaining N-1 parameters are the concentrations of the element of interest in each layer excluding the top one. The concentrations should conform to the unit of the

user's choice, which could be in number or moles per unit volume, or any other forms. To narrow the search space for fitting, as much physical intuition about the system as possible should be used. For example, if a given layer is devoid of the element of interest, the concentration of that element should be set to 0.

The geometry of the system is accounted for with two important inputs: the detection length of the detector, and the slit size of the beam. The software assumes the beam to be centered around the detection length, which would maximize the signal in a real experimental setup. Depending on the relative magnitude of the detection length and the slit size of the incidence beam, the entire detection area could be fully or partially utilized, as illustrated in **Figure 6.2**. These two scenarios are also accounted for in the calculations as follows.

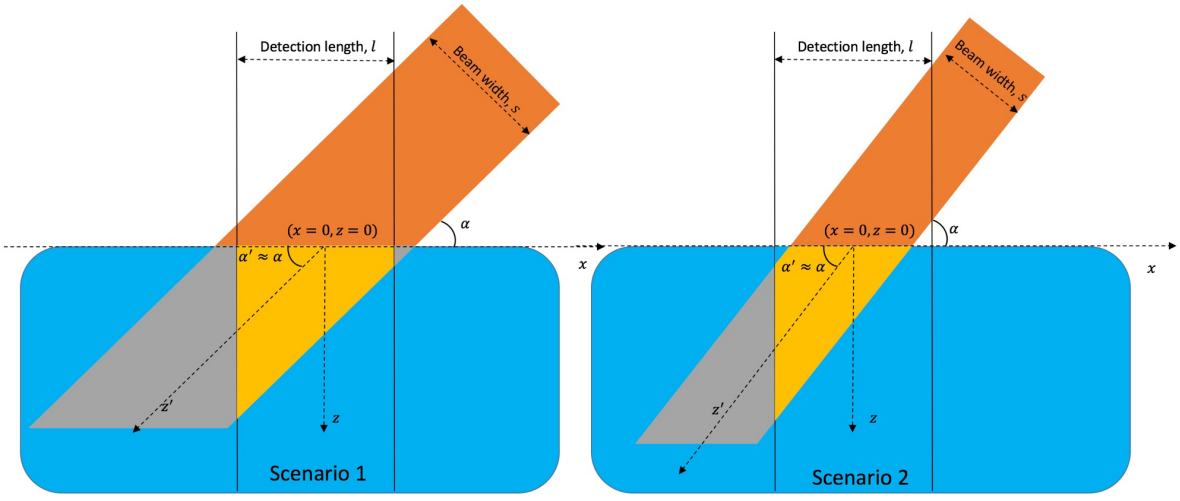

Figure 6.2: Two integration scenarios of the x-ray fluorescence signal. Left: The incidence beam footprint is larger than the detection length. Right: The incidence beam footprint is smaller than the detection length. The beam width is s, angle of incidence is α, and the detection length is l. The x-axis is assumed to be horizontal from left to right, z-axis perpendicular to the interface going vertically from up to down, and y-axis is out of plane, perpendicular to the x-z plane shown. Additionally, a z'-axis is assumed to follow the propagation direction of the transmitted beam. The yellow regions indicate the region from which fluorescence signal can be detected.

When the footprint $s/\sin\ \alpha$ is larger than the detection length l (**Scenario 1** of **Figure 2**):

I_B

$$= CN_B \iint I(z)dzdx$$

$$= CN_B I_0 T_i(\alpha)\left(\int_{-\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\int_0^{(x+\frac{l}{2})\tan\alpha} e^{-\frac{z}{\Lambda(\alpha)}}dzdx - \int_{\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\int_0^{(x-\frac{l}{2})\tan\alpha} e^{-\frac{z}{\Lambda(\alpha)}}dzdx\right)$$

$$= CN_B I_0 T_i(\alpha)\left(\int_{-\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\int_0^{(x+\frac{l}{2})\tan\alpha} e^{-\frac{z}{\Lambda(\alpha)}}dzdx - \int_{\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\int_0^{(x-\frac{l}{2})\tan\alpha} e^{-\frac{z}{\Lambda(\alpha)}}dzdx\right)$$

$$= CN_B I_0 T_i(\alpha)\left(\int_{-\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\Lambda(\alpha)\left(1-e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(x+\frac{l}{2}\right)}\right)dx - \int_{\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\Lambda(\alpha)\left(1-e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(x-\frac{l}{2}\right)}\right)dx\right)$$

$$= CN_B I_0 T_i(\alpha)\Lambda(\alpha)\left(\int_{-\frac{l}{2}}^{\frac{s}{2\sin\alpha}}\left(1-e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(x+\frac{l}{2}\right)}\right)dx + \int_{\frac{s}{2\sin\alpha}}^{\frac{l}{2}}\left(1-e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(x-\frac{l}{2}\right)}\right)dx\right)$$

$$= CN_B I_0 T_i(\alpha)\Lambda(\alpha)\left(l+\frac{\Lambda(\alpha)}{\tan\alpha}e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(\frac{s}{2\sin\alpha}+\frac{l}{2}\right)} - \frac{\Lambda(\alpha)}{\tan\alpha} + \frac{\Lambda(\alpha)}{\tan\alpha} - \frac{\Lambda(\alpha)}{\tan\alpha}e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(\frac{s}{2\sin\alpha}-\frac{l}{2}\right)}\right)$$

$$= CN_B I_0 T_i(\alpha)\Lambda(\alpha)\left(l-\frac{2\Lambda(\alpha)}{\tan\alpha}e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(\frac{s}{2\sin\alpha}\right)}\sinh\left(\frac{l}{2}\frac{\tan\alpha}{\Lambda(\alpha)}\right)\right)$$

I_S

$$= CN_S I_0 T_i(\alpha)\int_{-\frac{l}{2}}^{\frac{l}{2}}dx$$

$$= CN_S I_0 T_i(\alpha)l$$

When the footprint $s/\sin\ \alpha$ is smaller than the detection length l (**Scenario 2** of **Figure 2**):

I_B

$$= CN_B \iint I(z)dzdx$$

$$= CN_B I_0 T_i(\alpha) \int_{-\frac{s}{2\sin\alpha}}^{\frac{s}{2\sin\alpha}} \int_0^{\left(x+\frac{l}{2}\right)\tan\alpha} e^{-\frac{z}{\Lambda(\alpha)}} dzdx$$

$$= CN_B I_0 T_i(\alpha) \int_{-\frac{s}{2\sin\alpha}}^{\frac{s}{2\sin\alpha}} \int_0^{\left(x+\frac{l}{2}\right)\sin\alpha} e^{-\frac{z}{\Lambda(\alpha)}} dzdx$$

$$= CN_B I_0 T_i(\alpha) \int_{-\frac{s}{2\sin\alpha}}^{\frac{s}{2\sin\alpha}} \Lambda(\alpha)\left(1 - e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(x+\frac{l}{2}\right)}\right) dx$$

$$= CN_B I_0 T_i(\alpha)\Lambda(\alpha)\left(l - \frac{2\Lambda(\alpha)}{\tan\alpha} e^{-\frac{\tan\alpha}{\Lambda(\alpha)}\left(\frac{l}{2}\right)} \sinh\left(\frac{s}{2\sin\alpha}\frac{\tan\alpha}{\Lambda(\alpha)}\right)\right)$$

I_S

$$= CN_S I_0 T_i(\alpha) \int_{-\frac{s}{2\sin\alpha}}^{\frac{s}{2\sin\alpha}} dx$$

$$= \frac{CN_S I_0 T_i(\alpha) s}{\sin\alpha}$$

XeRay uses a "likelihood" approach to estimate errors of the fitting parameters, instead of the reduced χ^2 to estimate fitting uncertainties [30]. The likelihood of a set of parameter values being real is equivalent to the Bayesian probability of the fitting curve they produce given the observed data. This likelihood is closely linked to χ^2, wherein increasing χ^2 values are associated with diminishing likelihoods. In fact, the likelihood of a fit being the true fit is proportional to $\exp(-\chi^2/2)$ under two key assumptions: (1) the model is appropriate, and (2) the fitting error behaves as a normal distribution. Both assumptions are proper in our context of gathering data by counting photons, and fitting data to properly derived equations. Thus

$$Likelihood \propto \exp(-\chi^2/2). \quad (18)$$

The program generally runs three fitting procedures for a given set of parameters. The **first procedure** runs an optimization over all parameters being fitted using the TRR method, which is a computationally inexpensive optimization algorithm. Intelligent starting values greatly increase the odds of TRR finding the true global minimum, though this is not always a guaranteed outcome. In order to obtain error estimates of fitted parameters, the program in the **second procedure** uses a brute force approach to fit each parameter systematically. Specifically, the program steps through a grid of one parameter while optimizing the others by TRR to find the minimized χ^2 value for each step. Then the program calculates the normalized likelihood values over the grid. Under most circumstances, a plot of the likelihoods over the grid produces a bell curve that is easily fit to a Gaussian function to yield the mean and estimated standard deviation of the fitted parameter. This second procedure is repeated for every parameter. Assuming all the likelihoods behave in this Gaussian fashion, one standard deviation of a parameter would represent a 0.6826^n error window, where n is the number of parameters being fitted. Therefore, standard deviations would need to be adjusted to reflect a true 0.6826 error window. Aside from the variance of a single parameter, we might be interested in looking at the correlation of two parameters in order to identify highly correlated scenarios. In the **third procedure**, the program steps through a grid of each pair of two parameters and optimizes the χ^2 via TRR, calculates the joint likelihoods over the grid, and enables the user to visualize the χ^2 or likelihood distribution of each pair.

In a typical fitting process, the user first builds up the layered system with the GUI and enters the starting values for the parameters while visualizing the calculated curve with the experimental data. Then the user tries to pick a number of parameters for *XeRay* to fit while holding the other parameters constant, plots the likelihood curve, and refines the fitting parameters and ranges for further rounds of fitting.

The *XeRay* package is published as a public depository on github.com. Users can obtain the most up-to-date package at https://github.com/ZhiliangGong/xeray, and add the package including all the subfolders to the MATLAB search path to enable it. Note that at the time of writing, the software runs on MATLAB 2015a or later versions. There are two versions of the software, both contained in the package. The first version is fired up with command 'XeRayGUI' and analyzes datasets with spectra, giving the user the ability to visualize, compare, and integrate peaks in the software. The second version is fired up with command 'XeRayGUI(2)', and is for datasets with integrated intensity.

6.6 Case Study 1 - Ca^{2+} accumulation at an SOPA monolayer

Enter 'XeRayGUI' at the MATLAB command line to launch the version for the dataset with spectra and resize the GUI for proper display. The GUI contains three panels: the left for listing data sets, the middle with two plot regions, and the right with inputs, outputs, and other manipulation tools. This version of the software takes in spectral data files of a standardized text format with extension 'xfluo', which specifies the incidence angle range, energy in keV, and x-ray fluorescence intensity and error for each corresponding angle. Two examples of 'xfluo' files

can be found in the ‘example’ folder. If manually formatting data files becomes too tedious, the user might consider writing a script to automate the formatting process, as was proven necessary for the authors working at the ChemMatCARS facility.

As a demo, we will utilize TXRF data collected from two $CaCl_2$ systems. The first one ‘bulk.xfluo’ is from an aqueous solution of 50 mM $CaCl_2$ with 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5; the second one ‘surf.xfluo’ is from an aqueous solution of 1 mM $CaCl_2$ with 10 mM HEPES, pH 7.5 with an SOPA monolayer at 30 mN/m. Click ‘load’ to load both data files from the ‘example’ folder. After loading, the first spectrum from the ‘bulk’ data set appears. The program might take a significantly longer time to load for its first launch in a MATLAB environment. The user can now carry out some data visualization manipulations, such as comparing the spectra of multiple scans at multiple angles, zooming in for a specific element by selecting the one from the drop-down menu, fitting the peak of that element with Gaussian or Lorentzian line shapes, subtracting a linear background from the fitted peaks, etc. If the displayed energy range for an element needs to be changed, or an element of interest is not listed, it can be added by clicking the ‘Add or Modify’ option in the element drop-down menu.

For this demo, choose Ca from the element drop-down list, check ‘Start Fitting’, and then select all of the listed angles. Now plotted at the lower figure is the integrated fluorescence intensities and the fits for all angles; the fits can look slightly different depending on the selected Gaussian or Lorentzian lineshape. Now the GUI should look like **Figure 6.3**. Since there are two Gaussian

peaks, each with a different energy and thus a different penetration depth in solution, the program chooses the peak with the higher intensity for subsequent data fitting.

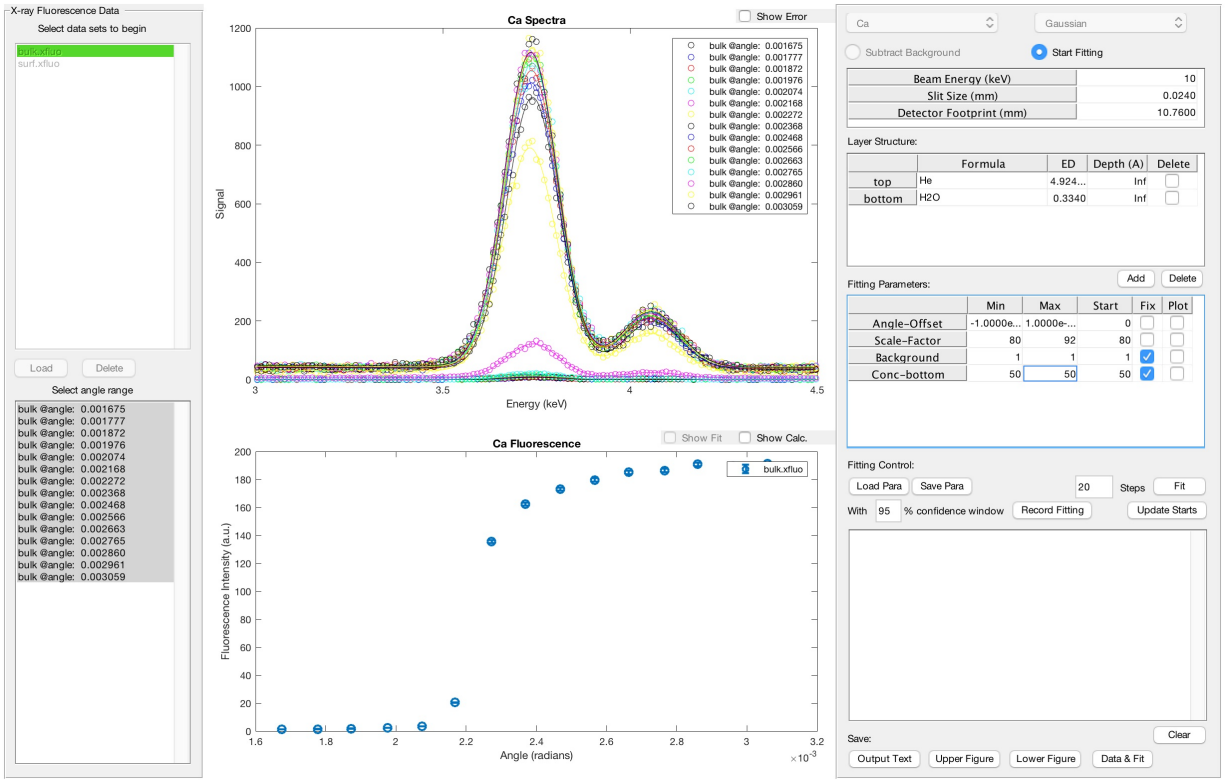

Figure 6.3: **Fluorescence spectra and integrated signal from the bulk.xfluo data file.** Scale the GUI window to show all graphical elements in the GUI properly. Input and check the following parameters: beam energy in keV, the slit size for the incidence beam in mm, the detection length of the detector in mm. Build the layer structure by use of the 'Layer Structure' table, in which a chemical formula describes the overall composition of a layer, and the electron density of the layers as calculated by the user, or obtained otherwise. Note that the chemical formula should only contain element abbreviations and stoichiometry numbers, where the first letter of an element abbreviation is in uppercase and the following letter, if applicable, in lowercase. Any element without an associated stoichiometry number is implied to be one. No spaces and other special symbols are allowed in the formula. Only the relative ratio of the atoms matter, so CH_2 is equivalent to C_2H_4. For example, a solution of 1 M $CaCl_2$ can be represented as $H_2OCa0.018Cl0.036$. Note that the bulk concentration is fixed at 50 mM, surface concentration at 0 nm^{-2}, and background at 0 a.u. for this 'bulk' data set. The two varying parameters are angle offset and scale factor.

A few fields describing the experimental system need to be filled in, as discussed in the caption for **Figure 6.3**. The electron density of each layer can be obtained either by calculation given the composition and density of the layer, or by other means, such as x-ray reflectivity studies. Below the parameter inputs is a table containing the fitting parameters. The angle offset parameter accounts for the small instrumental misalignment. As previously mentioned, the concentration in the top layer should be 0. Here the user can specify the lower bound, upper bound, and starting value for each parameter that the program fits. The 'Min' and 'Max' columns correspond to the lower and upper bounds. A parameter can be held constant throughout the fitting by checking 'Fix'.

The user can click 'Load Para' to load 'bulk.xfluopara' to first analyze the bulk fluorescence signal. Inputs and data in the table can be manually edited, or can be loaded by clicking the 'Load Para' button, which loads a specially formatted text file with extension 'xfluopara'. The user can click 'Save Para' to save the current inputs and table data in this format, or manually create 'xfluopara' files based on the existing format. The purpose here is to fit the scale factor C_0.

After loading the parameters, the user can check the 'Show Calc.' box above the lower figure to show the calculated intensity for the given parameters, and perform a fitting procedure by clicking 'Fit'. If a parallel pool is set up in MATLAB, an indicator should light up, and the fitting would proceed in a parallelized fashion.

After the fitting procedure is completed, the fit is plotted on top of the data in the lower figure, and the fitting results are logged in the output box with a time stamp. Now the user can choose to visualize the fitting quality based on likelihood, χ^2, and their joint distributions between two parameters. To do so, one simply checks the 'Plot' option in the table for the fitted parameters. The user can also change the reported confidence window, save the text from the output window, save the lower figure and upper figure, save the data the GUI is using, or save the data and fit as a text file. The user can try checking 'Plot' for angle offset, and the GUI should look something like **Figure 6.4**.

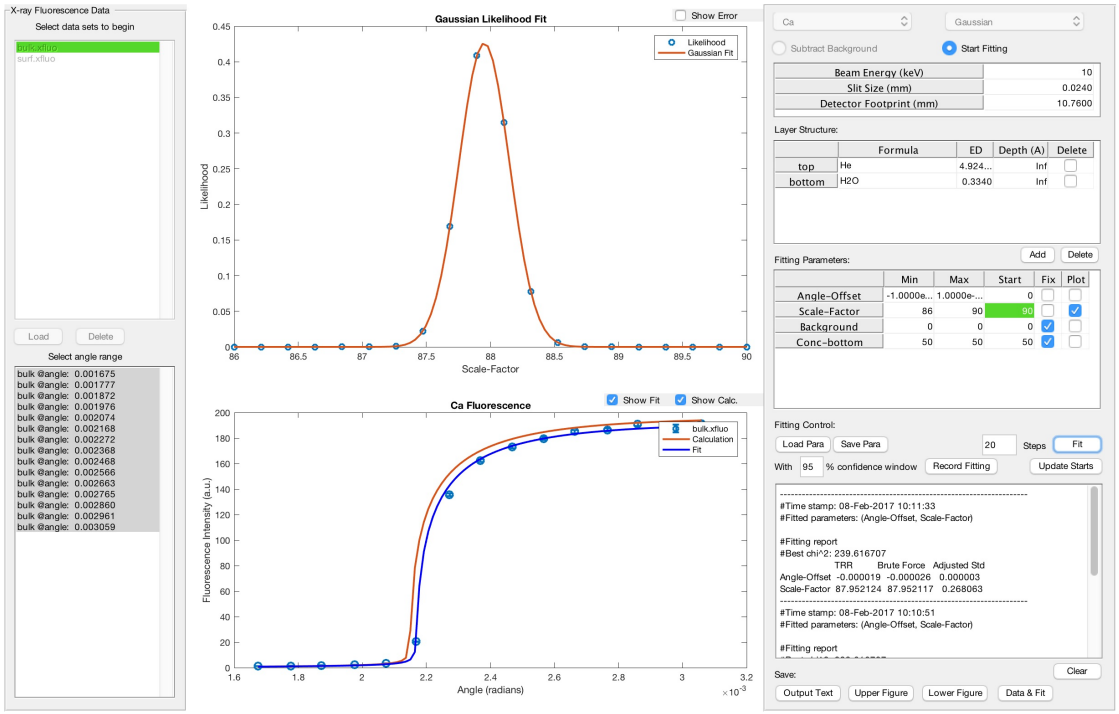

Figure 6.4. Fitted bulk.xfluo data set, and visualization of the likelihood distribution of the angle offset. Fitting results are logged in the output box. Note that TRR in the output box means fitting is performed using the trust region reflective method, and 'Adjusted Std' would represent a real 0.6826 error window for each parameter based on a brute force search. The purpose of fitting the 'bulk' data is to obtain a scale factor, C_0, which is found to be 88.0 ± 0.27, a very tight error window. The upper plot shows the likelihood distribution of the scale factor and its Gaussian fit.

The user should carefully inspect the likelihood and χ^2 distributions of the fitting parameters to ensure that all of them are nicely centered Gaussian or Gaussian-like distributions. **Figure 6.5** shows an improper range for the scale factor. A typical fitting process would require multiple trials toward an eventual narrowing down of the errors to a tight range for each fitting parameter.

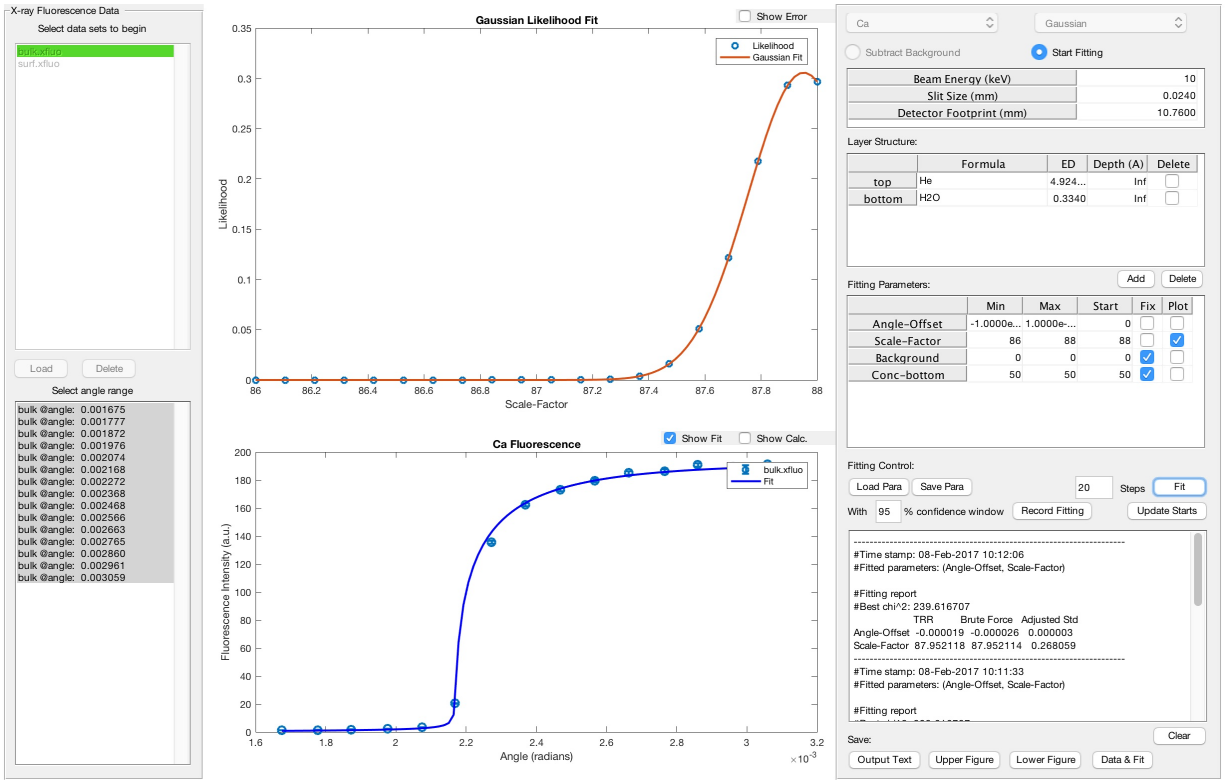

Figure 6.5. Example of improperly set fitting range. An example of improperly set fitting range evident from both the fact that the likelihood distribution of this parameter is not a bell curve, and that the fit is significantly lower than the data in the lower figure. Here the upper bound of the scale factor is too small, and needs to be increased.

After fitting the data of the 50 mM $CaCl_2$ bulk signal, we have obtained the scale factor of 88.0 ± 0.27, and can now proceed to fit the surface data. The user can uncheck 'Start Fitting', select 'surf.xfluo' in the file panel, select the whole angle range, check 'Start Fitting', load the

'surf.xfluopara' for the surface data, and click 'Fit'. Refer to **Figure 6.6** for the fitting parameters used. Fitting this data set would give a surface Ca^{2+} concentration of 4.93 ± 0.14 M, which for the layer thickness of 8 nm as measured from x-ray reflectivity, corresponds to 2.38 ± 0.06 ions/nm^2. The surface concentration of SOPA is 2.45 ± 0.03 molecules/nm^2, according to the surface pressure *vs*. area per molecule isotherm of SOPA at the experimental temperature of 23 ℃ (**Figure 6.7**) on a buffer with 10 mM HEPES and 1 mM Ca^{2+}, pH 7.5. Therefore, the stoichiometry between SOPA headgroups and Ca^{2+} on the surface is close to 1:1. In fact, two earlier studies also revealed a 1:1 stoichiometry between Ca^{2+} and 1,2-dimyristoyl-*sn*-glycero-3-phosphate (DMPA) [31], and 1,2-dioleoyl-*sn*-glycero-3-phosphate (DOPA) [32] respectively. It is worth noting that metal ions heavier than calcium generally have much higher excitation efficiency and thus produce much more accurate measurements.

6.7 Case Study 2 – using *XeRay* to analyze x-ray fluorescence data at the dodecane/surfactant/water interface

In addition to testing the software from an air/liquid interface data set, we further checked the capability of *XeRay* to analyze total reflection fluorescence from the liquid/liquid interface. Here we used an x-ray fluorescence dataset on Sr^{2+}, a model heavy metal ion, from a dodecane/surfactant/water system as reported by Bu *et al.* [33]. Here the top liquid phase is the organic solvent dodecane, and the bottom liquid phase is an aqueous solution of Sr^{2+} at pH 5.7. The calibration sample data are collected from a dodecane/aqueous solution sample with 50 mM Sr^{2+} in the aqueous phase, and no surfactants, and thus no interfacial enrichment of Sr^{2+}. The interface sample data are collected from a dodecane/surfactant/water system, where the

surfactant is di-hexadecyl phosphate (DHP), and the aqueous solution contains 0.01mM Sr^{2+}. The 'examples' folder contains the two data files as 'sr-calibration', and 're-interface', respectively. It is worth noting that in liquid/liquid interface studies, x-ray comes from the side of the top phase before entering the system, rather than from the top, incurring little scattering upon entering the system. As a result, it is proper to assume the top organic phase to be infinite in thickness.

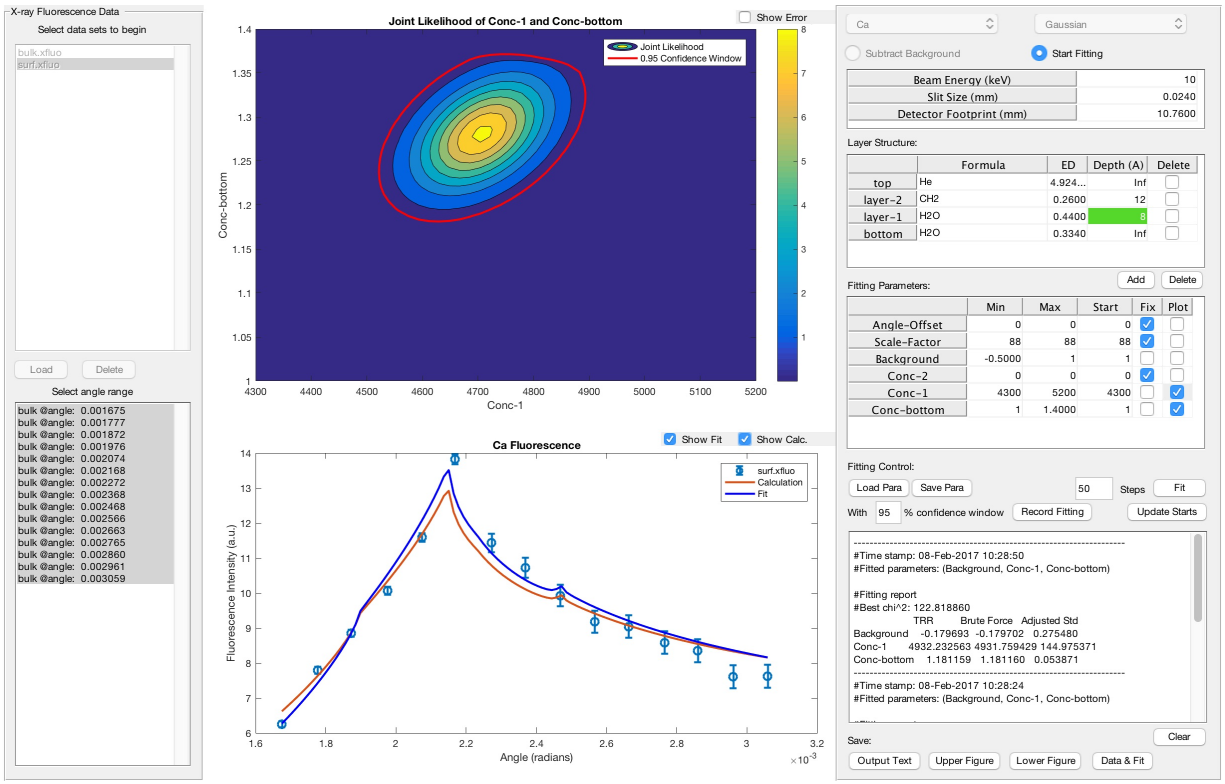

Figure 6.6. Fitting results for 1 mM $CaCl_2$ with a SOPA film yield a concentration of 4.93 ± 0.14 M, corresponding to a surface concentration of 2.38 ± 0.06 ions/nm^2. The fitted results indicate a 1:1 stoichiometry between SOPA headgroups and Ca^{2+} ions. Note changes in the 'Fix' column of the table, where the scale factor is now fixed to be 88. Here a constant negative background needs to be added to the signal from the insoluble Langmuir monolayer system, likely because the sensitive surface combined with our assumption of a perfectly coherent incidence beam leads to an overestimation of the signal.

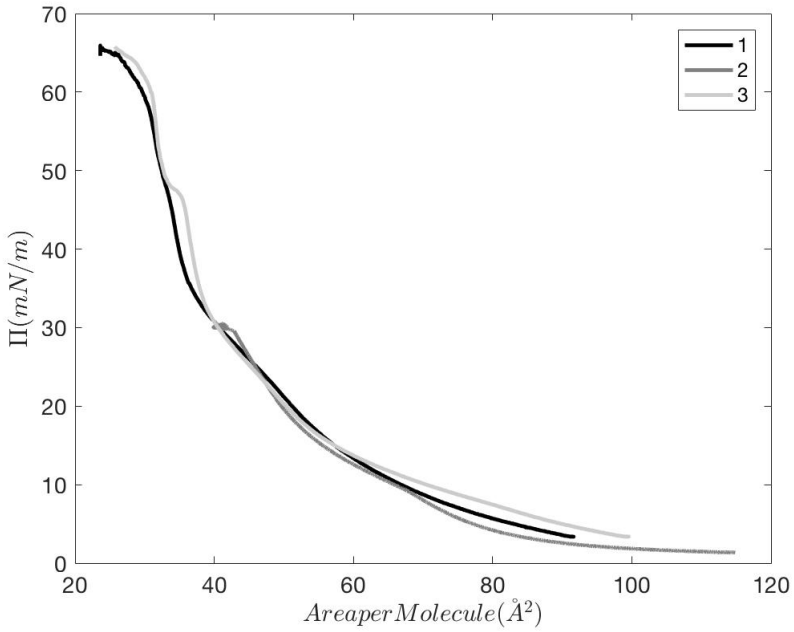

Figure 6.7: SOPA isotherms at 23 °C with subphase 10 mM HEPES and 1 mM Ca^{2+}, pH 7.5. Three repeats of the isotherm yielded area per molecule at 30 mN/m to be 40.73, 40.56, and 41.00 $Å^2$, respectively.

Here the fluorescence spectra are already integrated to give the total fluorescence signal, in which case we boot up *XeRay* by typing 'XeRayGUI(2)' in the MATLAB console. Except for the fact that there are no spectra visualization components in this new user interface (see **Figure 6.8** and **Figure 6.9**), the analysis procedure closely resembles that of the air/liquid case.

A scale factor of 2.49 ± 0.01 is first obtained from analyzing the calibration sample, which is then applied to fitting the surface enrichment of Sr^{2+} at the headgroup layer of the interface sample. The layer structure of the interface sample is built according to the result of fitting x-ray reflectivity data [33]. Specifically, there are five layers in total, the top infinite layer of dodecane, a small disordered layer attributed to the end methyl group of the surfactant, the hydrophobic tailgroup layer of the surfactant, the hydrophilic headgroup layer of the surfactant, and the

bottom infinite layer of water. Sr^{2+} ions would predominantly accumulate at the headgroup layer, and the fitting results indicate as much as 5.3 ± 0.2 M Sr^{2+} is present in the headgroup layer, more than 5 orders of magnitude than its bulk concentration. This layer concentration translates to an area of 68^{+6}_{-5} Å^2 (95% confidence) per Sr^{2+} ion, overlapping with the result from the previously reported range of 86^{+20}_{-15}Å^2 (95% confidence). One possible contribution to the discrepancy between the mean values of the two results is the possible curvature of the liquid/liquid interface accounted for by the earlier report [33], but not considered in *XeRay*.

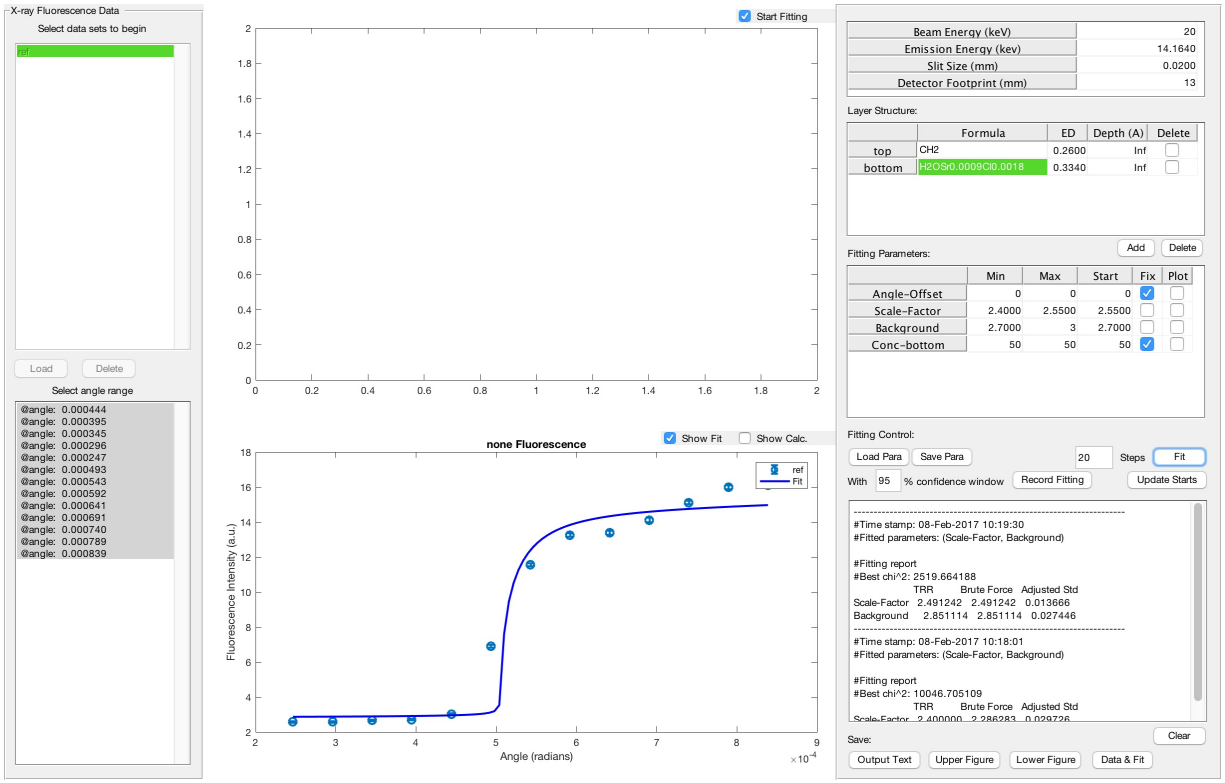

Figure 6.8. Fitting x-ray fluorescence from the dodecane/water calibration sample yields a scale factor of 2.49 ± 0.01. Here the thickness of dodecane is about 1.1 mm. Since the incidence wave enters the dodecane phase from the side, it would scatter minimally when entering the system, and would have traveled a significant length before reaching the interface, so the top phase is assumed to be infinite in thickness. Note that 'CH2' was used to represent the chemical formula of dodecane, which is to suggest that slight deviations of the stoichiometry ratios in a chemical formula are not critical.

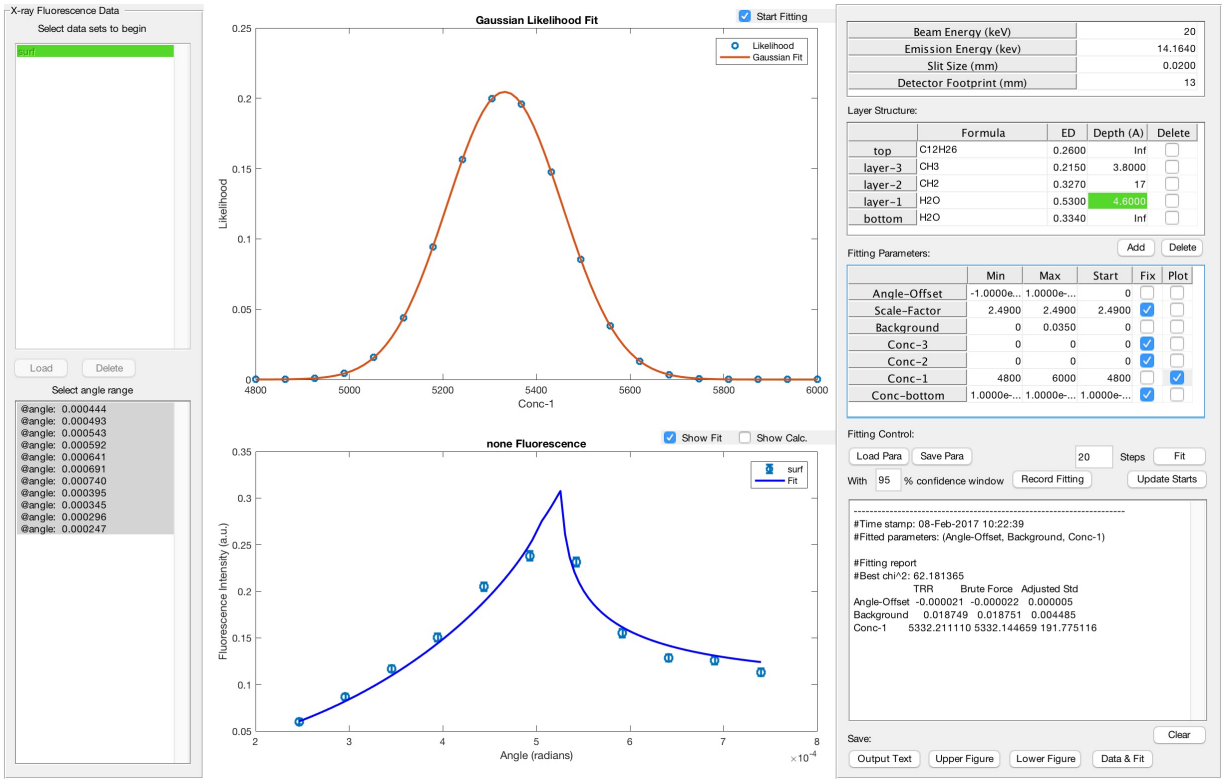

Figure 6.9. Fitting Sr^{2+} fluorescence from a system of dodecane/surfactant/water. With only 0.01 mM Sr^{2+} in the bulk of the aqueous phase, the signal coming from the aqueous phase is comparably small so that the signal coming from the interface dominates. The fitted surface concentration of 5.3 $\pm$ 0.2 M represents more than 5 orders of magnitude enrichment of Sr^{2+} by the presence of the surfactant at the interface.

6.8 Conclusion

Using *XeRay* we have been able to accurately analyze TXRF data from two distinct model systems: Ca^{2+} accumulation at the air/liquid interface, and Sr^{2+} extraction at the liquid/liquid interface. With the air/liquid dataset, we established a 1:1 stoichiometry between SOPA headgroups and Ca^{2+} ions at the air/water interface. In the liquid/liquid dataset, we have been able to reproduce the fitting result of Sr^{2+} enrichment at the dodecane/surfactant/water interface

from a published report [17]. We reasonably believe that *XeRay* could be readily used in studies involving solids or any other phases.

6.9 References

1. Lemmon, M. Membrane recognition by phospholipid-binding domains. *Nature Reviews Molecular Cell Biology*. 9, 2 (2008), 99–111.
2. Clapham, D. Calcium Signaling. *Cell*. 131, 6 (2007), 1047–1058.
3. Decher G. and Schlenoff JB. Ed. Multilayer thin films: sequential assembly of nanocomposite materials. *Wiley-VCH*, Weinheim, Germany, 2012.
4. Kato, T. Polymer/Calcium Carbonate Layered Thin Film Composites. *Advanced Materials*. 12, 20 (2000), 1543–1546.
5. Beckhoff, B., Fliegauf, R., Kolbe, M., Müller, M., Weser, J., and Ulm, G. Reference-Free Total Reflection X-ray Fluorescence Analysis of Semiconductor Surfaces with Synchrotron Radiation. *Analytical Chemistry*. 79, 20 (2007), 7873–7882.
6. Moyer, B. A., Ed. Ion Exchange and Solvent Extraction: A Series of Advances. *CRC Press: Boca Raton, FL*, 2010; Vol. 19.
7. Somorjai GA. and Li Y. Introduction to surface chemistry and catalysis. 2nd Edition. *John Wiley & Sons, Inc.* Hoboken, NJ, 2010.
8. Wobrauschek, P. Total reflection x-ray fluorescence analysis - a review. *X-ray Spectroscopy.* 36, 5 (2007), 289–300.
9. Klockenkämper R. and von Bohlen A. Total-reflection x-ray fluorescence analysis and related methods. 2nd Edition. *John Wiley & Sons, Inc.*, Hoboken, NJ, 2014.
10. Yoneda, Y. and Horiuchi, T. Optical Flats for Use in X Ray Spectrochemical Microanalysis. *Review of Scientific Instruments*. 42, 7 (1971), 1069–1070.
11. Bu, W., Flores, K., Pleasants, J., and Vaknin, D. Preferential affinity of calcium ions to charged phosphatidic acid surface from a mixed calcium/barium solution: X-ray reflectivity and fluorescence studies. *Langmuir.* 25, 2 (2009), 1068–73.

12. Novikova, NN., Zheludeva, SI., and Konovalov, OV. Total reflection X-ray fluorescence study of Langmuir monolayers on water surface. *Journal of Applied Crystallography*. 36 (2003), 727.
13. Graber, ZT., Wang, W., Singh, G., Kuzmenko, I., and Vaknin, D. Competitive cation binding to phosphatidylinositol-4, 5-bisphosphate domains revealed by X-ray fluorescence. *RSC Advances*. 5 (2015), 106536.
14. Szoboszlai, N., Polgári, Z., Mihucz, VG., and Záray, G. Recent trends in total reflection X-ray fluorescence spectrometry for biological applications. *Analytica Chimica Acta*. 633, 1 (2009), 1-18.
15. Wang, W., Kuzmenko, I., and Vaknin, D. Iron near absorption edge X-ray spectroscopy at aqueous-membrane interfaces. *Physical Chemistry Chemical Physics.* 16 (2014), 13517.
16. Körner, A., Abuillan, W., Deichmann, C., Rossetti, F., Köhler, A., Konovalov, O., Wedlich, D. and Tanaka, M. Quantitative Determination of Lateral Concentration and Depth Profile of Histidine-Tagged Recombinant Proteins Probed by Grazing Incidence X-ray Fluorescence. *The Journal of Physical Chemistry B*. 117, 17 (2013), 5002-5008.
17. Bu, W., Mihaylov, M., Amoanu, D., Lin, B., Meron, M., Kuzmenko, I., Soderholm L., and Schlossman, M. X-ray Studies of Interfacial Strontium-Extractant Complexes in a Model Solvent Extraction System. *The Journal of Physical Chemistry B*. 118, 43 (2014), 12486–12500.
18. Geraki, K., and Farquharson, MJ. X-ray fluorescence and energy dispersive x-ray diffraction for the quantification of elemental concentrations in breast tissue. *Physics in Medicine and Biology.* 49 (2004), 99-110.
19. Marguí, E., Queralt, I., and Hidalgo, M. Application of X-ray fluorescence spectrometry to determination and quantitation of metals in vegetal material. *Trends in Analytical Chemistry*. 28, 3 (2009), 362.
20. Borgese, L., Zacco, A., and Bontempi, E. Total reflection of x-ray fluorescence (TXRF): a mature technique for environmental chemical nanoscale metrology. *Measurement Science and Technology.* 20 (2009), 084027.
21. Mantler, M. and Schreiner, M. X-ray fluorescence spectrometry in art and archaeology. *X-ray Spectrometry*. 29 (2000), 3-17.

22. Lin, B., Meron, M., Gebhardt, J., Graber, T., Schlossman, M. and Viccaro, PJ. The liquid surface/interface spectrometer at ChemMatCARS synchrotron facility at the Advanced Photon Source. *Physica B: Condensed Matter*. 336, 1-2 (2003), 75–80.
23. Von Bohlen, A. Total reflection X-ray fluorescence and grazing incidence X-ray spectrometry—tools for micro-and surface analysis. A review. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 64 (2009), 821-832.
24. Pianetta, P., Takaura, N., and Brennan, S. Total reflection x-ray fluorescence spectroscopy using synchrotron radiation for wafer surface trace impurity analysis. *Review of Scientific Instruments*. 66 (1995), 1293.
25. Rehr, JJ. and Ankudinov, AL. Progress in the theory and interpretation of XANES. *Coordination Chemistry Reviews*. 249, 1-2 (2005), 131-140.
26. Wang, W., Bu, W., Wang, L., Palo, P., Mallapragada, S., Nilsen-Hamilton, M. and Vaknin, D. Interfacial Properties and Iron Binding to Bacterial Proteins That Promote the Growth of Magnetite Nanocrystals: X-ray Reflectivity and Surface Spectroscopy Studies. *Langmuir*. 28, 9 (2012), 4274–4282.
27. http://henke.lbl.gov/optical_constants/getdb2.html
28. Vidal, B. and Vincent, P. Metallic multilayers for x rays using classical thin-film theory. *Applied Optics*. 23, 11 (1984), 1794.
29. Weisbrod U., Gutschke, R., Knoth J., and Schwenke H. Total reflection X-ray fluorescence spectrometry for quantitative surface and layer analysis. *Applied Physics A: Solids and Surfaces*. 53, 5 (1991), 449–456.
30. Andrae, R. Error estimation in astronomy: A guide. *arXiv preprint arXiv:1009.2755*. (2010).
31. Garidel, P. and Blume, A. Calcium induced nonideal mixing in liquid-crystalline phosphatidylcholine- phosphatidic acid bilayer membranes. *Langmuir*. 16, 4 (2000), 1662–1667.
32. Takahashi H., Yasue Y., Ohki K., and Hatta I. Structural and thermotropic properties of calcium-dimyristoylphosphatidic acid complexes at acidic and neutral pH conditions. *Biophysical Journal*. 69, 4 (1995), 1464–1472.
33. Bu, W., Hou, B., Mihaylov, M., Kuzmenko, I., Lin, B., Meron, M., Soderholm, L, Luo, G. and Schlossman, M.L. X-ray fluorescence from a model liquid/liquid solvent extraction system. *Journal of Applied Physics.* 110 (2011), 102214.

CHAPTER 7

CONCLUSIONS

This thesis presented a study of the interaction between PS and four proteins, namely Tim1, Tim3, Tim4, and MFG-E8, with a focus on Tim3 and MFG-E8. Using molecular dynamics simulations and x-ray reflectivity methods, molecular interaction models of Tim3 and MFG-E8 with PS-containing membranes are first proposed. Due to major rearrangements in protein conformation upon contacting the membrane, the crystal structure (Tim3), or the solution NMR structure (MFG-E8) could not fit the x-ray data obtained when the proteins are bound to the lipid membrane, but provided a starting point for MD simulations to obtain a lipid-equilibrated protein structure. The tandem use of computational and experimental methods proved necessary and successful in resolving the membrane bound states of Tim3 and MFG-E8, a clear demonstration of the value of combining MD simulations with molecular resolution synchrotron x-ray measurements.

Informed by the molecular interaction picture, the physicochemical parameter space for these four proteins were explored. In this exploration, a framework of decomposable lipid-protein interaction components, such as chemical coordination, electrostatic and hydrophobic forces, and mechanical work, was developed. Guided by this framework of interaction components, new interaction modulators in addition to PS were discovered, shedding light on the mechanisms of the functioning of these proteins. Of particular significance is the remarkable sensitivity of Tim3 to phosphatidic acid (PA) and membrane packing density. This framework of interaction

components between protein and lipid could be generally applicable to not only PS-binding proteins, but also to all other membrane binding proteins.

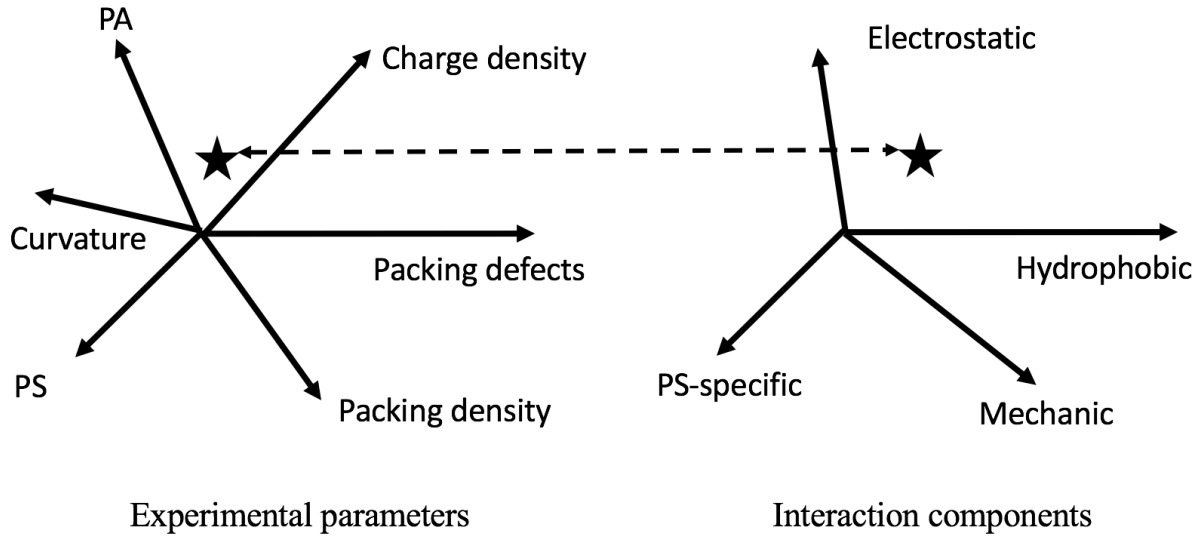

Figure 7.1: Connection between experimental parameter space and interaction components space.

By varying the key physical and chemical parameters, it became clear that these four proteins all share common electrostatic, hydrophobic, and mechanical interactions with the lipid membrane, in addition to their PS-specific binding sites. The four modes of interaction – PS-specific, electrostatic, hydrophobic, and mechanical – while conceptually independent, are nevertheless intertwined, and could be linked with experimental parameters designed to probe them, as illustrated in **Figure 7.1**. The spatial arrangement of the interaction motifs on the protein dictates the concurrence of these four types of interaction whenever the binding event occurs. The interplay of these four modes of interaction gives the four PS-binding proteins a large parameter space to differentiate among their PS-binding peers even though all four only have a single PS binding site, and thus might explain the diversity of their roles in physiological pathways.

Biological membranes are incredibly complex systems with hundreds of lipid species and thousands of protein species. This thesis has focused on a small portion of the physicochemical parameter space, and yet some findings genuinely surprised the author. In the march for deeper understanding the interactions involving membranes and proteins, the author advocates the structure-informed approach adopted in these studies, and the framework of interaction components developed in this thesis.