Frontiers in protein-lipid interactions studied by native mass spectrometry

Kin Kuan, Hoi

Exeter College

This thesis is submitted for the degree of *Doctor of Philosophy,* 2018

Abstract

Coordination between membrane proteins and their surrounding lipid environment dictates the function of biological membranes. While the study of membrane protein-lipid interactions is challenging for well-established techniques due to their transient and dynamic nature, recent developments in native mass spectrometry (MS) enables interrogation of membrane protein-lipid interactions in a label-free and high-resolution manner. This thesis presents two major developments in the study of membrane protein-lipid interactions via native MS: (1) the development of suitable MS methods for the transfer and characterization of nanodisc assemblies, promising membrane mimetics alternative to classical detergent micelles; and (2) the characterization of the lipid interactions with class A GPCRs.

Nanodiscs were transmitted into the gas-phase and collisionally activated to produce mass spectra, which were highly complex. To interpret these spectra, a dual Fourier transform method (**Chapter 2**) and a mass defect analysis (**Chapter 3**) were developed, allowing the determination of the lipid composition and the stoichiometries of the lipid/protein components of the nanodiscs. Analysis of the spectral results revealed that nanodisc-based MS provides novel insights into annular lipid interactions. In addition, the dissociation behavior of nanodiscs upon collisional activation was explored to understand how best to characterize nanodiscs in the gas-phase.

In the second part of this thesis, the lipid interactions of class A G protein-coupled receptors (GPCRs) were carefully characterized by native MS and

coarse-grained molecular dynamics simulations. Class A GPCRs were shown to preferentially bind phosphatidylinositol bisphosphate (PIP_2) over related lipids and potential PIP_2-binding hotspots were identified. The effects of PIP_2 on GPCR-G protein coupling were then investigated employing a Gα protein mimetic. Together, this information leads to a hypothesis that PIP_2 binding modulates GPCR-G protein coupling and identifies a potential lipid-binding motif. Overall, therefore this thesis develops two novel methodologies, the native MS of nanodiscs and G protein coupling, both of which facilitate studies of protein-lipid interactions in challenging systems.

Acknowledgement

I have many people to thank for their training, intellectual inspiration and support during the course of my DPhil. First and foremost, I would like to thank Professor Carol Robinson for giving me the invaluable opportunity to work in her laboratory and for her continued guidance and encouragement. I am grateful for the mentorship of Dr. Michael Marty and Dr. Hsin-Yung Yen. Their generosity in time and effort was key in my scientific development and progression. Since they have expertise in different disciplines, it has been an enjoyable time working with them.

I have had the pleasure of working alongside many brilliant and supportive people, from the Robinson, Benesch, and Baldwin groups as well as external collaborators. I would like to thank Dr. Tim Allison and Dr. Di Wu for their advice, discussion and timely support. I owe thanks also to Dr. Mark Agasid, Dr. Lucia Geis-Asteggiante, and Anne Gough for their feedback and correction to this thesis. I would like to thank Todd, Vicky, and Helena for the maintenance of the lab and instruments. In particular, I would also like to thank Jingwen, Joe, Matteo, Max, Anna, Heidi, and Iva.

The work described in this thesis would not have been possible without the contributions from my collaborators. I would like to thank Professor Chris Tate and Professor Andreas Plückthun for providing expression plasmids and high-quality protein samples. I would like to thank Dr. Wanling Song and Dr. George Hedger from Professor Mark Sansom's laboratory for preforming molecular dynamics simulations.

Finally, I am immensely grateful to my family. With their love and support I started my journey here and approached closer to my dream. I also would like to thank my friends for the company and joy brought to me.

Table of Contents

Abstract i
Acknowledgement iii
Table of Contents v
Contributions vii
Abbreviation ix
Chapter 1. Mass spectrometry as a tool for the investigation of protein-lipid interactions 1
1.1. Membrane protein-lipid interactions 1
1.1.1. Membrane proteins are important biological and pharmaceutical targets 1
1.1.2. Biological membrane lipid environments are heterogeneous and spatiotemporally dynamic 4
1.1.3. Lipids modulate the structure and function of membrane proteins 8
1.2. G protein-coupled receptors 15
1.3. Mass spectrometry 20
1.3.1. Ionization 22
1.3.2. Mechanistic models of ion production generated by the ESI process 25
1.3.3. Gas-phase ion collision 28
1.3.4. Mass analysis 29
1.3.5. Detection 35
1.3.6. Interpretation of protein nESI mass spectra 36
1.3.7. Instrumentation 37
1.3.8. Application of native MS in the study of protein-lipid interactions 40
1.4. Membrane mimetics 44
1.4.1. Native MS requires use of membrane mimetics 44
1.4.2. Native MS studies of nanodiscs 48
1.5. Thesis objectives 52
Chapter 2. Characterization of heterogeneous-lipid nanodiscs using native MS 54
2.1. Introduction 54
2.2. Methods 57
2.2.1. Sample preparation and MS analysis 57
2.2.2. Data analysis 62

2.3. Results and discussion 67
2.3.1. Using lipids of similar masses to obtain resolved mass spectra 67
2.3.2. Determining the average lipid mass of nanodiscs by autocorrelation 70
2.3.3. Determining the average lipid mass of nanodiscs by dual Fourier transform 74
2.3.4. The dissociation mechanisms of nanodiscs in the gas-phase 82
2.4. Conclusion 88
Chapter 3. Gas-phase liberation of membrane proteins from nanodiscs reveals the bound lipid annulus 89
3.1. Introduction 89
3.2. Methods 92
3.2.1. Sample preparation and characterization 92
3.2.2. Data analysis 96
3.3. Results and discussion 100
3.3.1. Preparation of membrane protein nanodiscs 100
3.3.2. MS measurement of membrane protein nanodiscs 103
3.3.3. Charge state disambiguation and mass assignment 106
3.3.4. Composition assignments reveal membrane protein complexes with many bound lipids 109
3.3.5. MS of membrane protein nanodiscs informs different chemical shells of lipid interactions 116
3.4. Conclusion 124
Chapter 4. An investigation of GPCR-lipid interactions and their role in G protein signaling pathway by native MS 125
4.1. Introduction 125
4.2. Methods 128
4.2.1. Sample preparation 128
4.2.2. MS and simulation methods 132
Results and discussion 137
4.2.3. Characterization of class A GPCR interactions with different lipids 137
4.2.4. Investigating the PIP_2-binding motifs of class A GPCRs 145
4.2.5. PIP_2 interactions enhance GPCR-G protein coupling 153
4.3. Conclusion 163
Chapter 5. Conclusion 164
Reference 171

Contributions

This thesis is entirely the author's work unless otherwise stated. Source of figures, from the author or literatures, is denoted in the caption. **Chapter 2** is majorly based on the works in Hoi *et al*.[1]. The author received supervision from Dr. M. Marty in this work. Dr. Marty provided the MSP plasmid. The author suggested that peak distances in the spectra/mass distributions are related to the lipid composition of nanodiscs. Dr. Marty suggested the autocorrelation and dual Fourier transform methods and coded these functions in UniDec. The author did all the other works, including MSP and nanodisc production, MS spectra acquisition, and data analysis.

Chapter 3 summarizes works in Marty *et al*.[2]. The author received supervision from Dr. M. Marty in this work. Dr. Marty and Dr. Tim Allison provided the plasmids of MSPs, AmtB, and AqpZ. The author made the proteins and nanodiscs and acquired spectra. For data analysis, Dr. Marty suggested mass defect analysis and coded this function in UniDec. The author ran deconvolution and mass defect analysis and processed the results (see **Figure 3.6** and **Figure 3.7**). Dr. Marty averaged the mass defect distributions and lipid distributions (**Figure 3.9**), determined the mass defects (**Table 3.1**), and fit the mass distributions with Gaussians (**Figure 3.10**). Dr. Marty did all the works of the simulations.

Chapter 4 includes works published in Yen *et al*.[3]. The author received supervision from Dr. H. Yen in this work. As much collaboration is involved, this section focuses only the works with the author's contribution. Collaborators'

contribution is acknowledged in the relevant sections. For protein production, the author made the six different NTSR1 variants/mutants. The author generated the plasmids through mutagenesis using the plasmids of NTSR1 HTGH4-ΔIC3B and NTSR1 HTGH4 43–421 provided by Professor Andreas Plückthun. Dr. Yen designed the mutagenesis primers. Professor Plückthun and Professor Chris Tate provided all other proteins. In **Section 4.2.3**, lipidomics (see **Figure 4.3** E and F): Dr. Yen prepared the lipid extracts; Dr. Di Wu ran the LC-MS measurements; the author did the analysis and assignment. The author did the competitive binding experiments for PIs (**Figure 4.6**). In **Section 4.2.4**, comparison between different NTSR1 constructs (**Figure 4.9**) and identification of PIP_2-binding sites by mutagenesis (**Figure 4.10**) were the author's works. In **Section 4.2.5**, comparison of the PIP_2 affinity between active and inactive β_1AR (**Figure 4.12)**: Dr. Yen prepared the samples and acquired the spectra; the author analyzed the spectra. β_1AR-Nb6b9 coupling experiments (**Figure 4.13** B): Dr. Yen prepared the samples and analyzed the data; the author acquired the spectra. Coupling experiments of β_1AR-Nb6b9 and β_1AR-mini-$G\alpha_{i(s)}$ in the presence of PIP_2 (**Figure 4.15** B and C): Dr. Yen and Dr. Idlir Liko prepared the samples; the author acquired the spectra and analyzed the data.

Abbreviation

$A_{2A}R$	Adenosine A_{2A} receptor
AC	Alternating current voltage
AmtB	Ammonium channel
Apo-A1	Apolipoprotein A1
Aqp0	Aquaporin-0
AqpZ	Aquaporin Z
C8E4	Tetraethylene glycol monooctyl ether
CCS	Collisional cross section
CDL	Cardiolipin
CEM	Chain ejection model
CG	Coarse-grained
CHS	Cholestery hemisuccinate
CID	Collision-induced dissociation
CMC	Critical micelle concentration
CRM	Charged residue model
Cryo-EM	Cryo-electron microscopy
CTD	Cytoplasmic domain
Da	Dalton
DAG	Diacylglycerol
DC	Direct current voltage
DDM	n-Dodecyl-β-D-maltoside
DESI	Desorption electrospray ionization
ECD	Extracellular domain
EI	Electron ionization

ESI	Electrospray ionization
Fos-Coline-16	n-Hexadecylphosphocholine
FWHM	Full-width at half-maximum
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
GIRK	G protein-coupled inwardly-rectifying potassium channel
GM3	Monosialodihexosylganglioside
GPCR	G protein-coupled receptor
GRK	G protein-coupled receptor kinase
HCD	Higher-energy collisional dissociation
ICL	Intercellular loops
IEM	Ion evaporation model
IM	Ion mobility
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
Kir	Inward-rectifier potassium ion channel
LB	Luria Broth
LILBID	Laser-induced liquid beam ion desorption
LMNG	2,2-didecylpropane-1,3-bis-β-D-maltopyranoside
m/*z*	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionization
MAPK	Mitogen-activated protein kinases
MBP	Maltose binding protein

MCP	Microchannel plate
MD	Molecular dynamics
MS	Mass spectrometry
MSP	Membrane scaffold protein
MST	Microscale thermophoresis
MWCO	Molecular weight cut-off
Nb6B9	Nanobody
nESI	Nano-Electrospray ionization
NMR	Nuclear magnetic resonance
NTSR1	Neurotensin receptor 1
OG	n-Octyl-β-D-Glucoside
$P2Y_1R$	Purinergic receptor
PA	Phosphatidate
PC	Phosphatidylcholine
PDB	Protein Data Bank
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PIP	Phosphatidylinositol monophosphate
PIP_2	Phosphatidylinositol bisphosphate
PIP_3	Phosphatidylinositol trisphosphate
PMF	Potential mean force
PS	Phosphatidylserine
PTM	Post-translational modification
r-IMAC	Reverse immobilized metal affinity chromatography

SEC	Size exclusion chromatography
S-lens	Stacked ring ion guide
SMALP	Styrene maleic acid co-polymer lipid particle
SPR	Surface plasmon resonance
TEV	Tobacco etch virus
Th	Thomson
TM	Transmembrane
TMD	Transmembrane domain
ToF	Time-of-flight
UHV	Ultra-high vacuum
UniDec	Universal Deconvolution
β_1AR	β_1-adrenergic receptor
β_2AR	β_2-adrenergic receptor

Chapter 1. Mass spectrometry as a tool for the investigation of protein-lipid interactions

1.1. Membrane protein-lipid interactions

1.1.1. Membrane proteins are important biological and pharmaceutical targets

Biological membranes serve as a physical barrier to define the cellular environment as well as different cellular compartments. Many proteins reside transiently as peripheral membrane proteins or permanently as integral membrane proteins on the membrane and are so-called membrane proteins (**Figure 1.1**)[4]. They make up to 20–30% of the proteins encoded in the human genome and perform a variety of functions[5-7]. In addition to transportation performed by channels and transporters and communication by receptors across the membrane, membrane proteins are also responsible for cell-cell recognition, cell junctions, attachment to the cytoskeleton and extracellular matrix, enzymatic activities, and stability of the membrane[8]. These functions are essential to sustain the cellular functions and therefore relate to the pathogenesis of many diseases as well as pathogen infections[6,9-12]. Membrane proteins, therefore, are important biological and pharmaceutical targets.

Transmembrane proteins are the major integral membrane proteins that span the entire membrane lipid bilayer[4]. They have evolved with a diverse architectural repertoire to perform versatile functions necessary to living organisms. Transmembrane proteins comprise transmembrane domains (TMDs) joining soluble domains on the two sides of the membrane (e.g., extracellular and cytoplasmic domains (ECDs and CTDs) for membrane proteins on the

plasma membrane). This specialized architecture allows transmembrane proteins access to both sides of the membrane, enabling communication and transportation between the separated aqueous environments. The soluble and transmembrane domains are distinct in structure, size and topology for different protein classes[13]. For example, single-pass (bitopic) membrane proteins, such as cytochrome P450, span the membrane only once. Multi-pass membrane (polytopic) proteins span the membrane multiple times and form an α-helix bundle or β-barrel in bacteria, where the N-terminus and C-terminus may be located on the same or opposite sides of the membrane[13]. Some transmembrane proteins exist as monomers and their sizes are smaller than 50 kDa (Dalton) (e.g., Neurotensin receptor 1 (NTSR1)). Other transmembrane proteins may have multiple subunits, exceeding 100 MDa, for example, the nuclear pore complex[14].

Because of the remarkable nature of their structure and diverse function, transmembrane proteins constitute over 50% of the targets of currently approved drugs (**Figure 1.2**)[6,9]. Ion channels, for example, are associated with neurological/neuromuscular disorders, including epilepsy, hypokalemic periodic paralysis and myotonia[15]. Receptor tyrosine kinases are closely related to the activation and progression of cancer[16,17]. For microbial pathogenesis, outer membrane proteins are the key multidrug efflux components that facilitate multi-drug resistance of bacteria[18]. Transmembrane proteins are also involved in the infection pathway of various viruses[19,20].

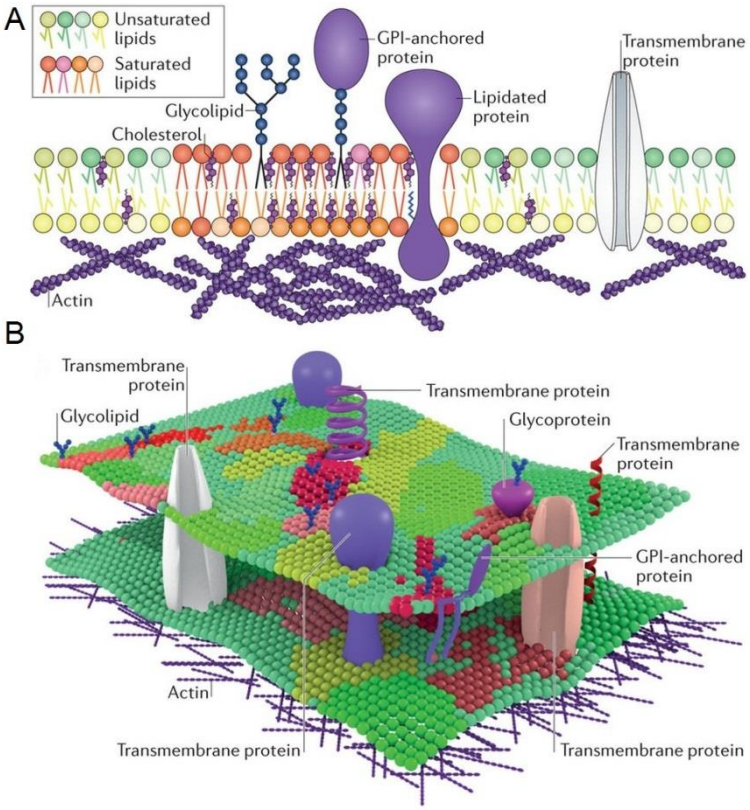

Figure 1.1 Cartoon representation of the organizational layout of the plasma membrane. (A) The plasma membrane consists of a variety of proteins and lipids that are dynamically interacting with each other. While protein and lipid components rapidly diffuse within the membrane due to the fluidity of the membrane, protein-protein, protein-lipid, and lipid-lipid interactions lead to asymmetry between the two leaflets and lateral heterogeneity of lipid distributions. (B) Membranes are thought to consist of various raft-like (more-ordered and less fluidic) and non-raft-like domains, which have distinct compositions and properties. Figure from Sezgin *et al*.[21].

Despite their important physiological implications, elucidating membrane protein structure and function is a time consuming and challenging process. Firstly, membrane proteins have many members and their structures and functions are highly diverse. Secondly, membrane proteins are often naturally expressed at low levels, and hence it is a challenge to produce sufficient material for characterization[22-24]. Thirdly, the inherent plasticity of membrane proteins may

hinder high-resolution structure determination by crystallography and electron microscopy[24,25]. Finally, membrane proteins require an appropriate amphipathic environment to maintain their stability, structure, and function[4,26-30].

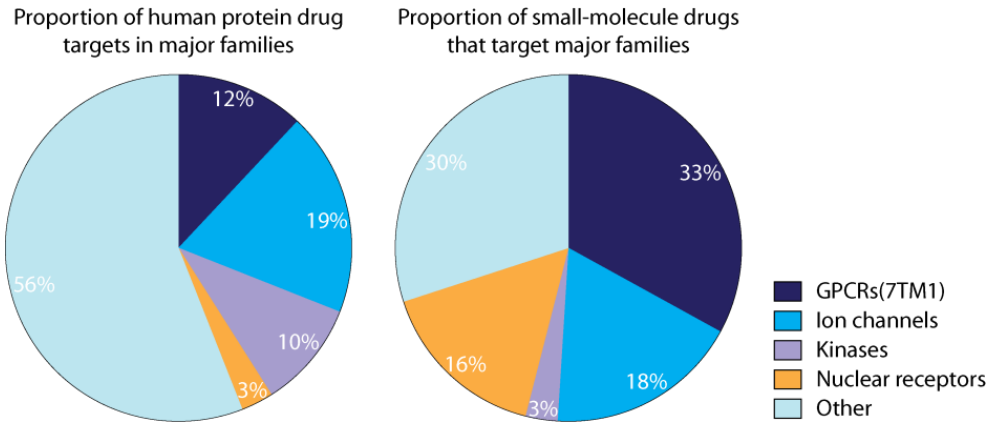

Figure 1.2 Human protein drug targets. The charts show the distribution of the human drug targets based on their gene families (*left*) and the proportion of small-molecule drugs acting on these families (*right*). Transmembrane proteins, G protein-coupled receptors and ion channels, account for over 50% targets of marketed drugs. The data is sourced from Santos *et al*.[9]. Figure drawn by the author.

1.1.2. Biological membrane lipid environments are heterogeneous and spatiotemporally dynamic

Biological membranes are composed of two building blocks: membrane proteins and lipids. Each of these building blocks accounts for ~50% of the bio-mass of biological membranes[4]. Lipids are amphipathic molecules that form bilayer structures in biological membranes, providing a hydrophobic barrier to separate the cytoplasmic and extracellular aqueous environments as well as an adaptable lipid matrix for membrane proteins to perform their function. About 5% of the genes in eukaryotic cells are used to synthesize lipids and they produce thousands of different lipids[31]. Major membrane lipid classes include

glycerophospholipids, sphingolipids, and sterols (**Figure 1.3**)[8,32]. Individual subclasses share a common chemical backbone with each class containing many members with different structures. Biological membranes thus contain a heterogeneous and complex mixture of lipids.

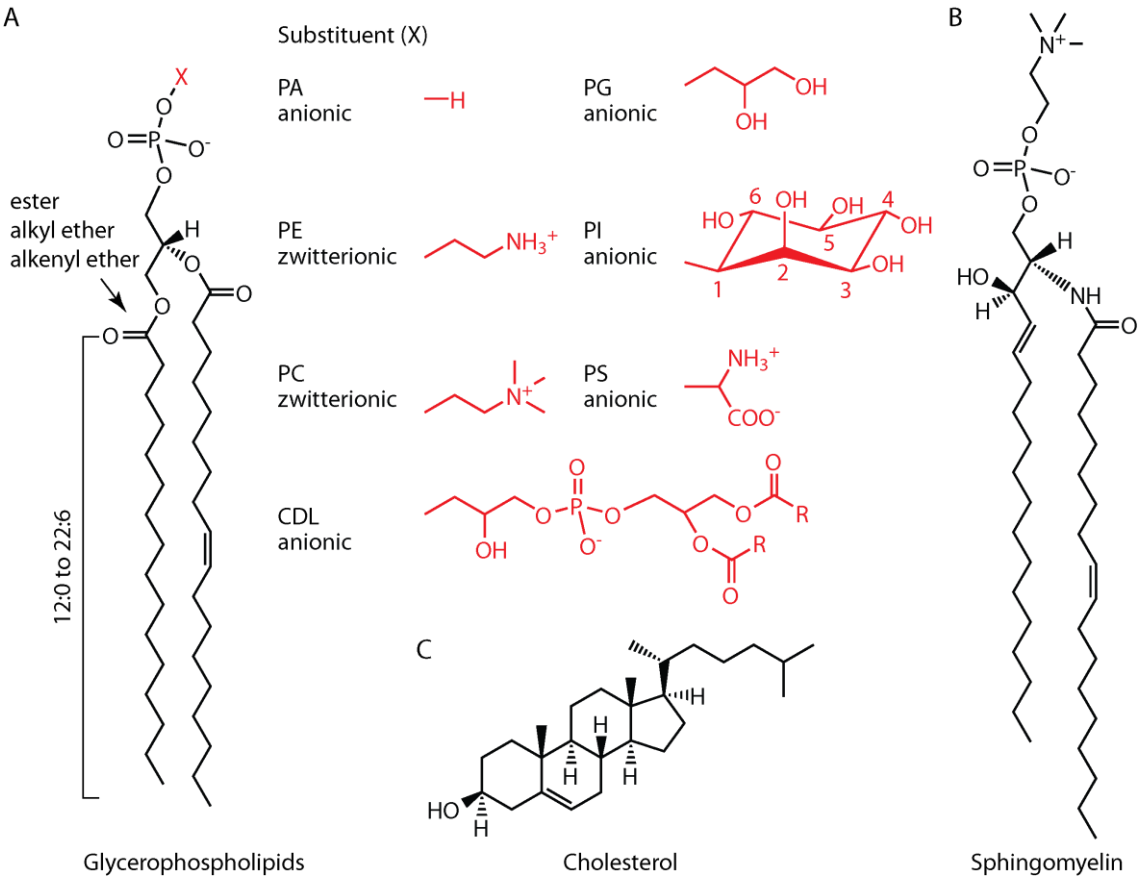

Figure 1.3 The major membrane lipids. The three major classes of membrane lipids are (A) glycerophospholipids, (B) sphingolipids (e.g., Sphingomyelin), and (C) sterols (e.g., cholesterol). Common head groups (X) of glycerophospholipids are listed. Figure drawn by the author.

The physical properties of membranes are defined by their lipid composition. Using glycerophospholipids as an example, these lipids contain two hydrophobic fatty acid tails (R groups) and a phosphate-based hydrophilic head group joined by a glycerol backbone. The phosphate group of phospholipids

can be modified with different chemical structures, resulting in phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylglycerol (PG), phosphatidate (PA), Phosphatidylinositol (PI), and cardiolipin (CDL) (**Figure 1.3** A). Due to the acidity of the phosphate group, most phospholipids are zwitterionic or anionic and result in a negatively charged cell membrane at physiological pH. The charge on the head group mediates the polar and electrostatic lipid-lipid and protein-lipid interactions[33,34]. The head group chemistry also affects how membranes response to changes in the aqueous environment (e.g., pH and ionic strength)[35]. Further, the size of the head groups contributes to the overall geometry of the lipids (inverted conical (e.g., polyphosphoinositides), cylindrical (e.g., PC), and conical (e.g., PA and CDL)) and hence their packing and the curvature of the lipid bilayer[36].

The lipid tails contribute to the lateral lipid distribution, permeability, fluidity, transition temperature, curvature, and lateral pressure profile of the membrane[36]. The acyl chain length of the R groups, commonly 12–24 carbons in length[4], determines the hydrophobic thickness of the lipid and therefore its hydrophobic interactions with other lipids and the TMD of membrane proteins. The lipid-lipid, lipid-protein, and protein-protein hydrophobic interactions drive the lateral heterogeneity and formation of micro-domains (rafts)[21]. Unsaturation in the R groups, typically 0–2 double bonds, leads to cis-conformation in the lipid tails, affecting the geometry of the lipid and its packing in the lipid bilayer. To maintain the fluidity of the membrane, glycerophospholipids are incorporated

which usually contain saturated acyl chains, unsaturated longer chains, or both saturated and unsaturated chains.

In addition to the glycerophospholipids, sphingolipids usually contain long-chain saturated acyl moieties[31,37]. These sphingolipids can form extensive inter- and intra-molecular hydrogen-bonding interactions with themselves[31,37]. Cholesterol has a characteristic planar and rigid four-fused-ring structure and preferentially interacts with sphingolipids through hydrogen-bonding and hydrophobic interactions[21,38,39]. These interactions allow sphingolipids and cholesterol to self-assemble in biological membranes, leading to relatively ordered micro-domains enriched with sphingolipids and cholesterol termed rafts[21,40].

The portfolio of membrane lipids varies according to the species and the cell type[32]. For example, PC lipids account ~40% of the phospholipids in mammal plasma membranes but ~10% in that of yeast[31]. Lipids are unevenly distributed in the plasma membrane and different organelles, e.g., most CDL in eukaryotic cells is confined to the mitochondrial membrane[31]. Even on the same membrane, lipid composition can be asymmetric between the two leaflets and distinct laterally (e.g., through the formation of lipid rafts)[4]. Furthermore, cellular membranes are dynamic, and as such, their lipid environments are constantly changing[41,42]. As discussed in the next section, membrane protein functions are modulated by their lipid environments. Because membrane proteins are diverse in structure and function, they require different membrane environments for proper functionality[43]. The spatiotemporal distribution of the huge variety of lipid species is important to sustain membrane performance[44].

1.1.3. Lipids modulate the structure and function of membrane proteins

It has long been recognized that the structure and function of proteins is dependent on their amino acid sequence as well as their post-translational modifications (PTMs) and environment[45-50]. Integral membrane proteins are essentially surrounded by, and interact with, many lipids. Early studies simply considered lipid bilayers as an adaptable matrix for anchoring membrane proteins, which is an essential role of the lipid bilayer[44,51]. However, lipids are now generally recognized as critical components of membrane protein structure and function[30,31,50-53]. Protein-lipid interactions affect membrane proteins in many ways, including their ligand affinity[54], activity[55,56], conformational dynamics[50,51,54,56,57], oligomerization status[58-61], stability[62,63], folding and topology[13,49], lateral distributions[21,64,65], and translocation within the cell[37,66]. Furthermore, some lipids act as substrates of membrane proteins and secondary messengers in cell signaling[67-70]. Therefore, protein-lipid interactions should be taken into account when considering membrane protein structure and function. From the protein perspective, membrane lipids are defined as bulk lipids, annular lipids, and non-annular lipids (**Figure 1.4**). These lipids interact and influence membrane proteins through different mechanisms.

Non-annular lipids: Non-annular lipids interact with membrane proteins at specific binding sites with high affinity, which is similar to the characteristics of ligand interactions[71]. Non-annular lipid interactions are best exemplified by the interactions of several ion channels and phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$)[56,57,72-74]. $PI(4,5)P_2$ is the primary agonist of inward-rectifier potassium ion channels (Kir2)[73,75]. It is a low-abundant lipid with three negative charges. Its

poly-anionic nature however allows $PI(4,5)P_2$ to interact with multiple basic residues in the CTD and TMD of the Kir2.2 channel. Binding of $PI(4,5)P_2$ induces a 6 Å upward movement of the CTD toward the TMD and insertion of the G-loop into the TMD. This leads to opening of the channel passage (**Figure 1.5**)[73]. Although other phospholipids, such as PA, can also bind to the lipid-binding site in the TMD, they cannot attract the CTD and open the channel. Additionally, it is striking that $PI(3,4)P_2$ cannot activate the Kir2.1 channel, indicating a high stereo-specificity at these lipid-binding sites[75]. This specific lipid-gated channel opening mechanism may have evolved first to keep the ion channel silent during trafficking and/or endocytosis, and later to use $PI(4,5)P_2$ as an auxiliary handle for activity regulation since PIP_2 can be readily manipulated by kinases and phospholipase activities[56,72].

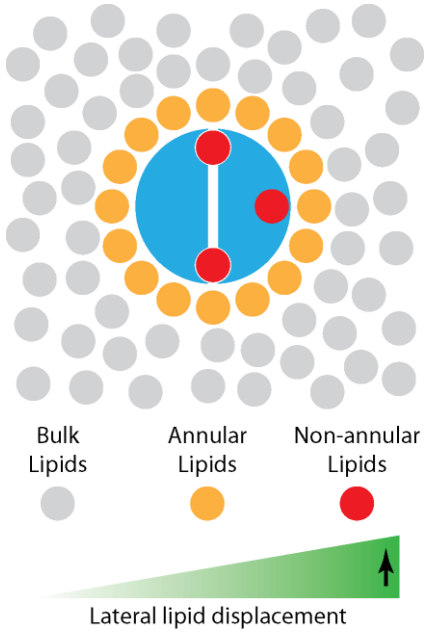

Figure 1.4 Membrane protein-lipid interactions within a cell membrane. Integral membrane proteins (*blue*) interact with bulk lipids (*grey*), annular lipids (*orange*), and non-annular lipids/lipid ligands (*red*). Inspired by Stangl *et al.*[76]. Figure drawn by the author.

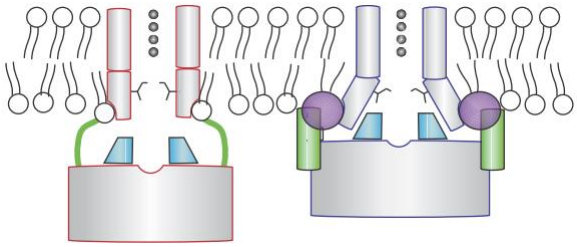

Figure 1.5 Proposed mechanism of Kir2.2 activations. Binding of PIP_2 (*purple*) leads to upward displacement of the CTD and widens the channel pore. Figure from Hansen *et al*.[73].

G protein-coupled inwardly-rectifying potassium channels (GIRKs) represent a further example and are distinct from other Kir channels as their activation relies on the synergistic effect of G proteins and PIP_2[77,78]. GIRKs have two functional gates: the inner helix gate in the TMD and the G loop gate in the CTD[57]. A R201A mutation of GIRK2 (Kir3.2) was used to mimic the effect of G protein binding, which opens the G loop gate[57,79]. A high-resolution structure of the R201A mutant bound with PIP_2 revealed that PIP_2 induced conformational changes of Kir3.2 and widened the channel pore of the inner helix gate. This example demonstrates the complex regulation of membrane protein function and structure by the coordination of both protein-protein and protein-lipid interactions.

Annular lipids: Integral membrane proteins are “solvated” by many lipids. Lipids in the first lipid shell that have direct contacts to the membrane protein surface are referred to as annular lipids or boundary lipids[27,71,80,81]. These lipids are essential elements for the structural integrity and activity of membrane proteins. As hydrophobic mismatches between the membrane protein TMD and lipid environment lead to energy penalties, membrane proteins are constantly

looking for lipid partners that provide an optimal hydrophobic thickness. Instead of specific binding motifs, the acyl chains of annular lipids adopt irregular but ordered conformations to accommodate the surface and hydrophobic thickness of the TMD[27,42,82]. Such interactions also ensure a tight seal around the membrane protein and lead to the impermeability of the membrane. Although annular lipids are in constant exchange with bulk lipids, their motion and diffusion are restricted by the protein-lipid interactions. The exchange rate between annular lipids and bulk lipids is typically between 1 and 2×10^7 s^{-1} at 37 °C, which is about half of the exchange rate between two bulk lipids[83,84].

Although a high-resolution structure of membrane proteins with a complete shell of annular lipids is not yet in existence, there is much experimental evidence to support the presence of annular lipids[82,85-90]. A prominent example was given by the high-resolution structures of aquaporin-0 (Aqp0) crystalized in DMPC and *E. coli* lipids, respectively[82,85]. In both structures, the Aqp0 subunits were virtually identical (**Figure 1.6** A) and conservatively co-crystalized with seven annular lipids. DMPC and *E. coli* polar lipids occupied similar positions. However, the binding modes were distinct between the two sets of seven lipids (**Figure 1.6** B), suggesting a lower specificity of annular lipid interactions than non-annular interactions.

Further, the activity of sarcoplasmic reticulum calcium-ATPase was studied in lipid bilayers containing a mixture of PC and PE[91]. This study discovered that the ATPase activity dropped when the PC content in the lipid bilayer was below 20%. It was proposed that 30 annular PC lipids were required to form the first

annulus for the calcium-ATPase. Deficiency of PC in the lipid bilayer leads to the reduction of calcium-ATPase activity[80,89,92]. Additionally, the calcium-ATPase is also sensitive to the acyl chain length of the PC in a lipid bilayer[93,94]. While optimal ATPase activity was found in C16–C18 chains, short (< C14) or long (> C24) chains were inhibitory. In another example, the structure of the heat- and capsaicin-activated ion channel, TRPV1, reconstituted in nanodiscs were resolved using cryogenic-electron microscopy (cryo-EM)[90]. Several annular lipids were observed to stabilize the open state of the channel-toxin complexes.

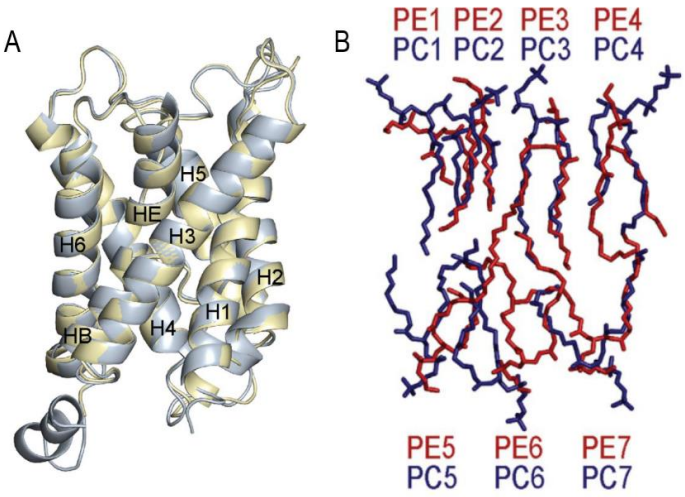

Figure 1.6 Comparisons of Aqp0 structures co-crystalized with annular lipids. (A) The structures of the Aqp0 co-crystalized with DMPC (*light blue*) and *E. coli* polar lipids (*gold*) are virtually identical. (B) Bound DMPC (*blue*) and *E. coli* polar lipids (*red*) are overlaid for comparison. Figure from Hite *et al*.[85].

Bulk lipids: Bulk lipids have a low residence time at the protein-lipid interface and diffuse rapidly within the lipid bilayer (typical diffusion coefficients of DOPC in a liquid disordered phase are ~7 $\times 10^{-12}$ m^2/sec)[71,95]. The effects of the bulk

lipids on membrane proteins depend on the generic properties of the membrane. As mentioned above, activation of Kir2 channels requires binding of PIP_2 at a specific binding pocket. There is a second lipid-binding site in the channel interacting with bulk anionic lipids. These interactions increase the sensitivity of the channel to PIP_2 by 10 to 100-fold[77,96,97]. The apparent membrane surface potential is a more decisive factor than the charge of individual lipid for these interactions. Other membrane physical properties, such as curvature and hydrophobic thickness affect the tension experienced by mechanosensitive channels, leading to opening/gating of the channel[50,98-100].

Overall, the repertoire of protein-lipid interactions and their effects are diverse. Cells manipulate membrane protein function through their protein-lipid interactions. As such, lipid-interaction sites may present promising target sites for drug and therapy development. Several recent studies have demonstrated drug/lipid-binding sites at the TMD of ion channels and G protein-coupled receptors (GPCRs)[101-105], highlighting the valuable information that can be learned from studying the interactions at protein-lipid interfaces. Further, the ability of lipids to mediate membrane protein functions also leads to the concept of membrane-lipid therapy, which aims to modify membrane lipid structures and micro-domains for pharmacological intervention of human diseases[106-108].

Nevertheless, membrane protein-lipid interactions are difficult to study and may lead to difficulties similar to those encountered when studying membrane proteins (**Section 1.1.1**). Additionally, cell membranes are extremely heterogeneous. Although the distribution and structure of many lipids are

established, the lipidome of even simple organisms such as *E. coli* remains incomplete[109]. Moreover, the dynamic and transient nature of protein-lipid interactions makes direct observation of these interactions challenging, which is further compounded by the different lipid interaction modes (e.g., the stoichiometry and binding site).

1.2. G protein-coupled receptors

GPCRs detect and respond to hormones, neurotransmitters, ions, photons, and other stimuli, triggering a diverse array of signaling cascades[110]. As such, GPCRs play a fundamental role in physiology and are associated with over 30 different human diseases of diverse etiology, including epilepsy, asthma, obesity, and heart failure[111-113]. Therefore, these receptors constitute ~30% of current drug targets and represent attractive therapeutic targets (**Figure 1.2**). GPCRs comprise the largest protein superfamily in the human genome[114]. Estimated to have more than 800 family members, GPCRs are classified into class A–F based on their sequence homology and functional similarity. Class A–F corresponds to rhodopsin-like receptors, secretin receptors, metabotropic glutamate receptors, fungal mating pheromone receptors, cAMP receptors, and frizzled/smoothened receptors, respectively[115,116]. Of all classes, class A GPCRs account for ~80% of all GPCRs and are the most frequently studied.

Despite their functional versatility, GPCRs share a remarkably conserved 7-TM topology and common mechanisms of signal transduction. In general, activation of GPCRs initiates from the binding of agonists. Upon activation, GPCRs transduce signals through G protein-dependent pathways by heterotrimeric G proteins and G protein-independent pathways by G protein-coupled receptor kinase (GRK)-mediated phosphorylation and arrestin coupling (**Figure 1.7**)[110,117,118]. In the G protein-dependent pathways, GPCRs recruit heterotrimeric G proteins (G_s, $G_{i/o}$, $G_{q/11}$, and $G_{12/13}$; there are 21 Gα, 6 Gβ, and 12 Gγ subunits identified in human)[119]. GPCRs act as guanine nucleotide exchange factors (GEFs), facilitating GDP/GTP exchange of the Gα subunit and

the subsequent conformational changes of the G protein that lead to dissociation of the G protein into the Gα and Gβγ subunits, which in turn activate their corresponding downstream effectors. Gα has an intrinsic GTPase activity. Once GTP-bound Gα hydrolyzes its GTP, GDP-bound Gα then re-associates with Gβγ to form an inactive heterotrimer. On the other hand, activated GPCRs can be phosphorylated at the C-terminal by GRKs, triggering arrestin coupling and activation. While arrestin coupling desensitizes the G protein-dependent pathways and promotes internalization, activated arrestins transduce signals through interactions with different effector proteins. The internalized receptor may be degraded intracellularly or recycled to the plasma membrane. Importantly, although it was assumed that internalized GPCRs were inactive[120], it is possible to form super-complexes of GPCRs, G proteins, and arrestins in the endosome, causing sustained G protein signaling[121,122].

The diverse signaling pathways of GPCRs rely on their dynamic participation in a range of functionally distinct conformations[123,124]. Binding of ligands alters the conformational equilibrium of the receptor in the inactive state. This biases conformations favoring the binding of G proteins, arrestins, and/or other signaling proteins at the intracellular interface of the receptor. Since there is a range of ligands with different efficacy and binding modes[125,126], they dictate the degree of the overall response of the receptor (e.g., full agonists and partial agonists) as well as the selectivity toward the diverse downstream pathways (e.g., biased ligands)[105,127-129].

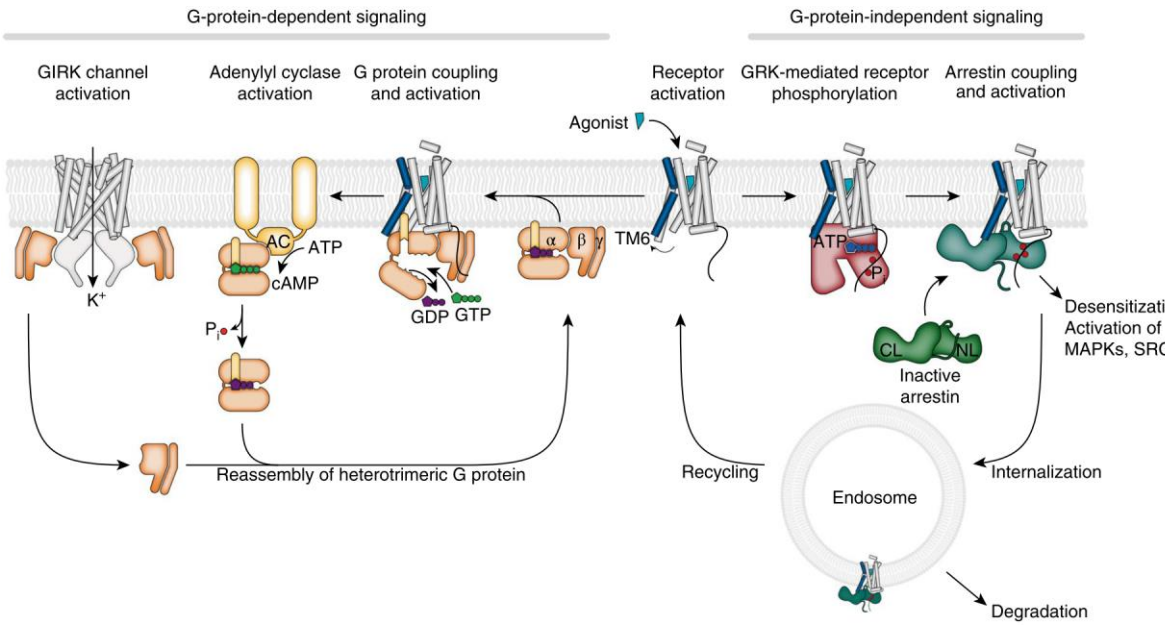

Figure 1.7 The signal transduction pathways of GPCRs. Agonist binding activates the GPCR and induces conformational changes that involve an outward movement of TM6 (*blue*). This allows the receptor to engage with a range of heterotrimeric G proteins (*orange*). The receptor facilitates nucleotide exchange and dissociation of the G protein into the Gα and Gβγ subunits. The dissociated subunits then mediate downstream signaling through different effector proteins, such as adenylyl cyclase (*yellow*) and GIRK (*gray*). G protein-mediated signaling terminates following GTP hydrolysis and re-association of the Gα and Gβγ into an inactive G protein. Alternatively, activated GPCR can be phosphorylated by GRK (*red*) and then couples to an arrestin (*green*). While arrestin coupling desensitizes the receptor and promotes internalization, the activated arrestin regulates different downstream effector proteins, such as mitogen-activated protein kinases (MAPKs) and Src kinases. The internalized GPCR may be degraded intracellularly or recycled to the plasma membrane. Figure from Hilger *et al*.[117].

Although the exact mechanism of GPCR signal transduction remains elusive, many studies have illustrated the relationships among ligand interactions, receptor conformations, and the downstream signaling in the last decade. One of the most exciting discoveries is the signature of GPCR activation. Binding of agonists and G proteins stabilizes the receptor in a fully active conformation that

is distinguished by outward movement of transmembrane helix 6 (TM6) (~14 Å) and is accompanied by rearrangement of the TM bundle[130-134]. This conformational change is essential to create a cleft to accommodate downstream transducers. Moreover, ligand efficacies are normally interpreted as how ligands affect the conformational sampling of the receptor, which alters the efficiency and kinetics of G protein coupling. One alternative mechanism was demonstrated whereby ligands allosterically modulate the conformation of receptor-coupled G proteins, influencing the turnover rate of the G protein upon GTP binding and the kinetics of downstream effector activation[135]. In a second mechanism, G protein coupling (nucleotide-free state) allosterically modulates the ligand affinity at the orthosteric binding site by stabilizing the GPCR in a "closed" active conformation[132,136]. Receptors in this "closed" conformation orientate amino acid side chains to shield the orthosteric binding site. These side chains restrict the dissociation and binding of ligands, leading to increased affinity in the ligand-bound receptor but decreased affinity in the ligand-free receptor.

Similar to other membrane proteins, lipids play a role in regulating GPCR signal transduction. In particular, the effect of cholesterol on GPCR signaling has been explored in many studies, invoked by the structure of β_2-adrenergic receptor (β_2AR) with co-crystalized cholesterols[137,138]. Cholesterol was reported to influence GPCR dimerization, stability, ligand affinity, and activation[139-141]. Modulatory effects were also observed for phospholipids. Different phospholipids were shown to have different impacts on the agonist affinity of β_2AR and the kinetics of receptor activation[54]. In addition to directly modulating

the receptor, lipid environments also influence the coupling efficiency of G proteins to GPCRs and the partition of G proteins in the membrane (e.g., enriched in lipid rafts)[138,142-144].

Overall, elucidating the molecular insight of GPCR signal transduction represents a multifactorial problem. It is necessary to understand the efficacy and binding mode of different ligands and connect these to the sampling of the conformational space of the receptor. This connection is further complicated by their potential effect on the selectivity of divergent signaling pathways and the behavior of transducers. Further, the loose allosteric coupling between the orthosteric ligand-binding site and the intracellular transducer-binding site suggests another dimension of regulation in GPCR signaling. Lastly, although this remains poorly understood, the function of GPCRs is closely related to the lipid environment. More studies are required to fully decipher GPCR signal transduction at the molecular level. While capturing these non-covalent interactions using crystallography has proved challenging, investigations by spectroscopic methods may be feasible.

1.3. Mass spectrometry

Many methods have been developed to study protein-lipid interactions. For example, X-ray crystallography, nuclear magnetic resonance (NMR), isothermal titration calorimetry (ITC), microscale thermophoresis (MST), and surface plasmon resonance (SPR), which are classical methods used to study biomolecular interactions, are adapted to study protein-lipid interactions[30,145]. More recently, cryogenic-electron microscopy (cryo-EM) and methods that rely on labeling, chemical modification/crosslinking, and protein/lipid arrays (e.g., lipid pull-down and lipid-overlay assay) have also been applied[30,90,145,146].

X-ray crystallography, cryo-EM, and NMR have provided important mechanistic insights into protein-lipid interactions, such as lipid binding stoichiometries, binding sites, and structural effects on the proteins. However, these structural methods are low throughput and require a large amount of sample[30]. It is difficult to unambiguously assign electron density of lipids from high-resolution structures[147]. For example, the unknown electron density of the X-ray structure of V-type ATPase was assigned to 20 PG lipids[148] but was proved later as 10 CDL molecules[149]. SPR is a highly sensitive technique to measure physical constants of binding kinetics and affinities. However, SPR requires immobilization of one of the binding partners on a surface, which could be challenging. MST is a highly sensitive technique that measures the mobility of molecules, which is sensitive to the changes in conformation, charge, and size resulted from binding events[150]. MST can be used to measure the affinity constant of protein-lipid interactions. However, MST typically requires fluorescent labeling. Although the intrinsic fluorescence from Trp/Tyr residues

can be exploited to achieve labeling-free experiments[151], this depends of the abundance of these residues in the protein. ITC is a label-free method to study the thermodynamics of protein-lipid interactions, providing information such as dissociation constant and stoichiometry. However, this method requires a large amount of samples to gain a measurable amount of heat. In addition, MST and IST are useful to determine thermodynamic parameters of lipid interactions but they do not provide kinetic information. Further, SPR, MST, and ITC measure the average property of all the populations in the solution and could not provide information for individual lipid-bound states or oligomeric states. These methods are also less useful to identify unknown endogenous lipid ligands co-purified with proteins. While lipid pull-down and lipid-overlay assays allow read out of protein-lipid interactions, labeling and modification of lipids or proteins may compromise the nature of the interactions. Array-type assays allow systematic study of protein-lipid interactions for different proteins and lipids. However, these methods require a full array of the lipidome or proteome, which is currently incomplete[145].

Mass spectrometry (MS) has proved invaluable in the fields of proteomics, lipidomics, metabolomics, and more recently, in the biophysical characterization of proteins and their interactions[152-154]. Advances in instrumentation and methodology have enabled the analysis of intact protein complexes (both soluble and membrane proteins) and preservation of their intra- and intermolecular non-covalent interactions in the gas-phase[155-157]. This technique, so called non-covalent or native MS, thus provides an additional and complimentary tool for structural biology[153,158-160].

Native MS is able to provide accurate mass determination of protein complexes, from which elucidation of the stoichiometry of the protein subunits and associated adducts is possible (e.g., lipids)[161,162]. In addition, native MS can resolve and therefore allow studies of different populations in the solution, such as different lipid-bound states or oligomeric states. As this technique measure the mass of the molecule, no labeling is required. In combination with different gas-phase (e.g., collision-induced dissociation (CID) and ion mobility (IM)) and solution-phase experiments (e.g., titration assays and mutagenesis), a wealth of information of the protein becomes accessible: ligand identity, kinetics, interaction affinity, binding sites, PTM profile, protein stability, protein conformation, and subunit connectivity[153,160,163,164]. In the following sections, basic principle and applications of native MS will be discussed.

1.3.1. Ionization

MS analysis initiates from ionization and transfer of analytes into the gas-phase, during which analytes become charged because of protonation, deprotonation, or forming adducts with charged ions. Electrospray ionization (ESI) is the most common ionization source applied in protein MS[159,165]. It is a soft ionization technique, which applies little residual energy onto the analyte molecules and therefore results in little fragmentation during ionization[166,167]. Folded protein ions can be produced by ESI[168], whereas matrix-assisted laser desorption ionization (MALDI) in general denatures proteins[169]. Laser-induced liquid beam ion desorption (LILBID)[170] and desorption electrospray ionization (DESI)[171] show great promise but are not yet routinely applied to membrane proteins. Soft ionization is distinct from hard ionization, e.g., electron ionization (EI)[172], in

which ions are largely fragmented due to excess activation. The amount of energy deposited on the analyte in ESI is lower than in EI, maintaining the integrity of protein ions and non-covalent interactions during the ionization process[159]. Because protein ions gain charges from protonation, deprotonation, or by forming adducts with other ions, protein ions generated by ESI typically carry multiple charge states. This avoids producing high mass-to-charge ratio (*m*/*z*) signals exceeding the detection range of many mass analyzers (e.g., singly/doubly charged ions produced by MALDI)[158]. Therefore, the advent of soft ionization techniques has formed the basis of the later application of MS in biophysical studies.

Protein ions are introduced into the gas-phase via application of an electric field to a protein-containing solution. The sample solution is loaded into a conductive capillary and experiences a strong electric field generated by the potential gradient (1.4–1.8 kV) between the capillary and the downstream cone (**Figure 1.8**)[173]. The electric field causes polarization of the solution and accumulation of charges at the surface of the tip[174]. Under the simultaneous effects of the liquid's surface tension and electric field, the liquid surface is distorted into a Taylor cone[175]. When the electric field overcomes the liquid's surface tension, a mist of charged droplets is emitted from the tip.

As the droplets travel downstream, the solvent continues to evaporate. Evaporation shrinks the droplet and increases the Coulombic repulsion on the droplet surface until reaching a minimum droplet radius (Rayleigh limit) with balanced Coulombic repulsion and surface tension force. This limit is defined by

the Rayleigh equation:

$$Z_{Ry} \cdot e = 8\pi(\varepsilon_0 \gamma R^3)^{\frac{1}{2}} \qquad \textbf{Equation 1.1}$$

where the Z_{Ry} is the charge number, e is the elementary charge, R is the droplet diameter, ε_0 is the electrical permittivity of vacuum, and γ is the surface tension of the solvent. Beyond the charge limit of a given droplet diameter, the droplet fissures into smaller droplets and achieves further desolvation. The smaller droplets may take up with ~2% in mass and ~15% of the charge from the parental droplet[176]. Successive rounds of desolvation and fission eventually produce desolvated, charged ions in the gas-phase. To improve desolvation and ionization efficiency, ESI usually employs volatile buffers such as ammonium acetate.

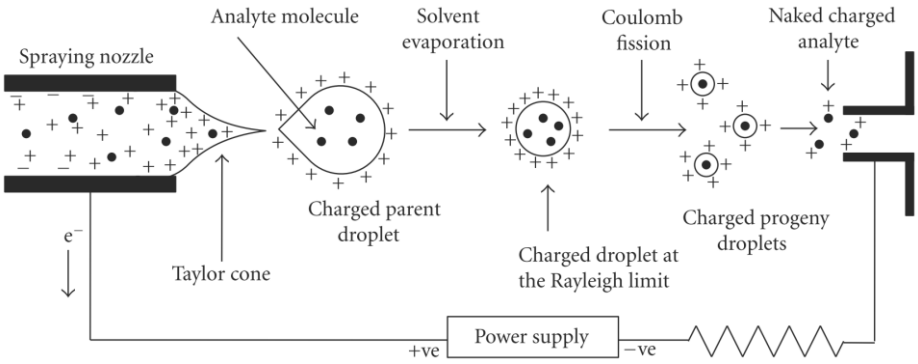

Figure 1.8 A schematic of electrospray ionization process. The sample solution is loaded into a conductive capillary and experiences an electric gradient across the capillary and the downstream cone. Accumulation of charges at the capillary tip overcomes the liquid's surface tension, leading to a mist of droplets. Evaporation reduces the droplet size, increasing Columbic repulsion on the droplet surface that subsequently triggers droplet fission upon Rayleigh limit. Repeated droplet fission and evaporation events eventually give rise to desolvated, charged ions. Nebulizing gases, such as N_2, are often used to confine the mist and accelerate solvent evaporation. Figure from Banerjee *et al.*[177].

Nano-electrospray ionization (nESI) is a sub-type of ESI. Although nESI works effectively by the same mechanism as ESI, it is achieved with a smaller needle orifice (1–5 μm) and much slower flow rates (10–50 nL/min)[159,178], nESI has many advantages over ESI for studying protein complexes. Due to a smaller capillary diameter, initial droplets emitted by nESI (150–200 nm) are an order of magnitude smaller than those emitted by traditional ESI[158,159]. This reduces the number and energy of collisions required in the desolvation process and thus provides a gentler ionization condition[159]. As desolvation becomes more efficient, it is less likely that incomplete desolvation will occur, which leads to peak broadening and lower resolution. Since offspring droplets carry salts from the parental droplet, smaller initial droplet size reduces the number of salt adducts in the offspring droplets and improves the spectral resolution[158,159]. Further, the slower flow rate reduces sample consumption. While ESI sample volume is in the milliliter range, nESI requires as few as a couple microliters of samples[159]. Spectra can be produced from a few picomoles of a protein complex. This advantage is particularly useful for membrane proteins given their low expression level.

1.3.2. Mechanistic models of ion production generated by the ESI process

Although the exact mechanism of protein ionization from ESI remains controversial, multiple models have been established to explain how ions are formed from charged droplets[179]. Two generally accepted theories include the ion evaporation model (IEM)[180] and charged residue model (CRM) (**Figure 1.9** *left* and *middle*)[167]. The IEM proposes that solvent evaporation leads to a higher surface charge density in the droplet containing the analyte. As the Coulombic

repulsion overcomes the solvation forces, charged particles and proteins are emitted from the droplet, resulting in protein ions. While the IEM provides an explanation for the charge state distribution of small ions, it appears to underestimate the charge state of macromolecules[181]. On the other hand, the second model, CRM, proposes that repeating evaporation and fission cycles generate nearly-desolvated droplets containing a single analyte. Further evaporation removes solvent molecules as neutral, depositing the charges from the droplet surface to the protein analyte. According to the CRM, the charge state of the protein ion is directly correlated to the droplet diameter and Rayleigh limit. This correlation was observed for different proteins regardless of their surface chemistry[182,183], thus, the CRM appears to better model protein ionization[159].

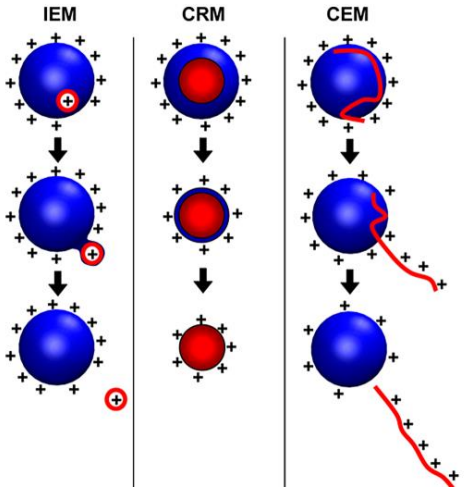

Figure 1.9 Proposed mechanisms of ESI. The IEM (*left*) shows that the charged analyte is ejected from a droplet as the Coulombic repulsion overcomes the solvation forces. The CRM (*middle*) proposes that a protein ion is formed by evaporation of a near-desolvated droplet containing a single analyte. The CEM (*right*) model proposes that an unfolded protein is gradually ejected from a charged droplet due to the unfavorable exposure of hydrophobic

residues to the solvent. Figure from Konermann *et al*.[179].

In addition to these two models, there is a third model, chain ejection model (CEM), describing the ionization of unfolded proteins (**Figure 1.9** *right*)[184]. In the CEM, protein unfolding causes the protein's hydrophobic residues to become solvent-accessible. Since this is unfavorable in free energy, the unfolded protein is gradually ejected from the droplet until completely expelled. While the CRM provides a mechanism for folded proteins, the CEM has been adapted to account for the ESI behavior of unfolded proteins. The CEM suggests that unfolded proteins carry a higher average charge and a broader charge state distribution than that of a folded protein as suggested by CRM. Such charge state dependencies upon folding state have been demonstrated experimentally[158]. It is well established that the observed charge state distribution of a protein serves as an indication of the protein-folding state (**Figure 1.10**).

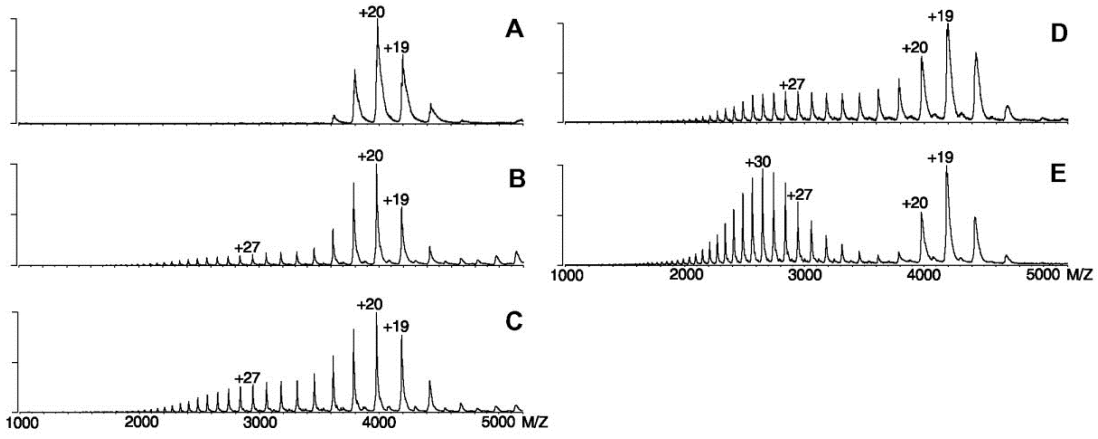

Figure 1.10 Representative nESI spectra of human serum transferrin acquired at different solution pHs. Spectra were acquired from solutions at pH (A) 7.5, (B) 5.5, (C) 4.0, (D) 3.5, and (E) 3.0. Figure from Gumerov *et al*.[185].

1.3.3. Gas-phase ion collision

Ions in the gas-phase are subjected to many collisions with gas molecules (e.g., air, N_2, and Argon). These collisions have many applications in MS, including improved transmission of ions into the mass spectrometer; collisional activation of the ions leading to unfolding, dissociation, or fragmentation; and collisional cross sections (CCS) measurement by ion mobility-mass spectrometry (IM-MS). Because large analyte ions generated by ESI obtain a large amount of kinetic energy from the potential and pressure gradient along the ion path, they are often difficult to focus and transmit. To address this issue, elevated pressures are applied in the source region[159,186]. At higher pressures, ions undergo more collisions with the neutral gas. Through collisions, the kinetic energy of ions is converted into internal vibrational energy of the ions and deposited into the colliding gas molecule. As the kinetic energy of the ions is reduced, ions become easier to focus. This approach is called collision focusing or collision cooling[159,186]. In addition, increasing the number of collisions in the source region improves desolvation and ionization efficiency, leading to a better resolution in the spectra.

Another important application of ion collision is CID. Gas-phase collision converts the kinetic energy of an ion into its internal energy, leading to fragmentation or disruption of non-covalent interactions. In native MS, membrane proteins are typically solubilized in detergent micelles and ionized as a membrane protein-detergent complex. Detergents are stripped by CID, carrying some internal energy with them and thereby stabilizing the folded protein structure in a similar fashion as evaporative cooling[187]. Although

detergents are critical in preserving the protein structure during ionization, they need to be removed post-ionization to produce a resolved mass spectrum[188]. Membrane protein-detergent complexes are typically activated by CID in a collision cell to liberate the encapsulated membrane protein. In many cases, membrane proteins are co-purified with lipids after purification. The vast majority of these lipids are weakly bound to the protein and removed by CID similar to detergents. Lipid adducts that have relatively strong interactions with the protein may be retained after collisional activation and observed in mass spectra. An excess of collisional activation may promote protein unfolding, subunit dissociation and/or adduct dissociation, CID therefore needs to be carefully controlled to preserve the target intact protein-lipid complex.

Gas-phase ion collision is also used in IM-MS[189]. The basic concept of IM separation involves injecting ions into a drift tube filled with neutral gas. Ions are directed through the drift tube by an electric field. Because small ions have a smaller CCS, hence a lower probability to collide with the gas molecules, they travel faster than larger ions in the drift tube. This allows ions to be separated based on their CCS values. The time that an ion has spent in the drift tube is measured and used to determine the CCS of the ion. As the CCS correlates to the size and geometry of the ion, IM-MS provides additional information of the protein structure[164,190].

1.3.4. Mass analysis

The core ability of MS is to determine the *m*/*z* of the sample ions. A range of mass analyzers employing different mechanisms has been developed. They

provide different resolution, mass accuracy, and *m*/*z* range to accommodate various experimental requirements.

Quadrupole: A quadrupole mass analyzer contains four parallel rods, either cylindrical or hyperbolic (**Figure 1.11**). The same voltage is applied between diagonal rods but opposite polarity is applied between neighboring rods. The voltage is composed of a direct current voltage (DC) of magnitude U and an alternating current voltage (AC) of magnitude V operating at a radiofrequency (RF) f, which is expressed as $\pm(U + Vcos(2\pi ft))$[191]. The potential gives rise to an electrical quadrupole field between the rods. The field deflects ions injected into the quadrupole in the *x* and *y* dimensions, causing them to travel in helical trajectories described by the Mathieu equation[191]. Only ions harboring stable oscillations with an amplitude smaller than the inner radius between the diagonal rods (r_0) traverse the quadrupole. For a given quadrupole with a predefined r_0 and angular frequency ($2\pi f$), the stability of the trajectory depends on the *m*/*z* of the ion and the U and V according to the Mathieu equation. In the RF-scanning mode, a particular range of *m*/*z* is selectively transmitted by applying appropriate DC and AC voltages[191]. Mass filtering enables selection of target ions for detection or tandem MS and is widely applied in different mass spectrometers.

Beside acting as a mass filter, quadrupoles may function as ion guides to focus the ion beam by switching to an RF-only mode (i.e., $U = 0$)[191]. In this mode, ions of a wide range of *m*/*z* are transmitted with stable trajectories. However, ions of high *m*/*z* suffer from poorer focusing and the maximum m/z

transmittable is defined by the V,f, and r_0[192]:

$$\left(\frac{m}{z}\right)_{max} = \frac{4\times10^6 V}{f^2 r_0^2} \qquad \textbf{Equation 1.2}$$

In principle, increasing the V or reducing the r_0 are useful to increase the transmittable *m*/*z* range. However, higher RF amplitudes and closer rod distance may lead to electric arc. Therefore, quadrupoles operating at lower frequencies have been developed to improve transmission of high *m*/*z* ions[159]. Furthermore, a quadrupole can be modified into a linear ion trap by placing two end electrodes at the source and exit regions. By applying a potential on the end electrodes, ions can be confined within the quadrupole[193].

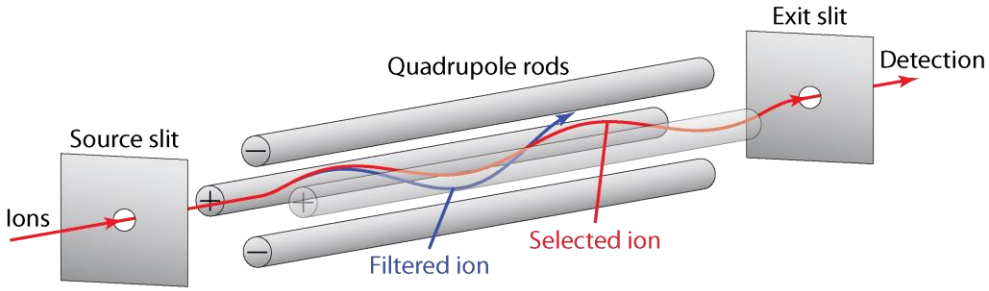

Figure 1.11 A schematic of a quadrupole analyzer. A quadrupole contains four parallel rods. Diagonal rods are connected with the same AC and DC voltages but the voltage sign is opposite to that of the consecutive rods. With a given AC and DC voltage, the quadrupole selectively transmits ions of a particular range of *m*/*z*. Figure drawn by the author.

ToF: ToF analyzers are another common type of mass analyzers[194]. The concept of ToF uses the simple speed-distance-time relationship to measure the *m*/*z* value of ions. The analyzer consists of two regions, a pusher, or source region, accelerating the in-coming ions and a field-free region terminating at a detector. In the idealized ToF model, ions at rest are accelerated from the source region by an accelerating potential (V), leading to a velocity (v):

$$v = \left(2V \cdot \frac{z}{m}\right)^{\frac{1}{2}}$$ **Equation 1.3**

The ions traverse the field-free region of length, D, and their arrival times to the detector (t) are measured. **Equation 1.4** describes the relation between the *m*/*z* of the ion and the arrival time as:

$$t = \frac{D}{v} = \frac{D}{\sqrt{2V}} \cdot \sqrt{\frac{m}{z}}$$ **Equation 1.4**

Due to the proportionality between the arrival time and $\sqrt{m/z}$, the time spectra recorded by the detector are translated into conventional mass spectra. Given the principle of ToF analyzers, they do not have a theoretical upper mass limit and can achieve fast detection with high sensitivity and resolution[195].

In reality, ions arriving at the pusher are introduced into the drift tube with a distribution of kinetic energies. This leads to different arrival times among the isobaric ions (i.e., with the same *m*/*z*) and thus reduces the measurement resolution[196]. This issue can be addressed by accelerating the ions to orthogonal direction and/or using a reflectron (**Figure 1.12**)[197]. In an orthogonal reflectron ToF mass analyzer, which combines both methods, ions are deflected toward the reflectron orthogonal to the incoming direction. In such a configuration, the ions gain an orthogonal velocity independent of their initial velocity in the incoming direction. The reflectron contains a series of lenses with a linear electrostatic field, deflecting the ions into a mirror direction. Since ions with higher kinetic energy penetrate further into the reflection, difference in velocity in the orthogonal direction is compensated by a longer travel distance.

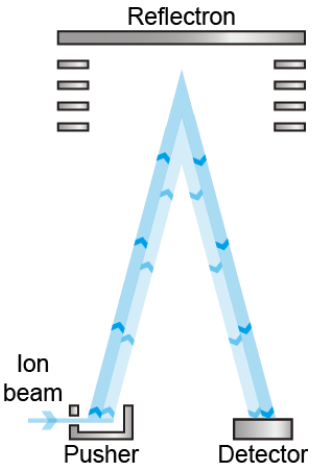

Figure 1.12 A schematic of a reflectron ToF mass analyzer. Ions in the ToF mass analyzer are deflected in a direction orthogonal to their incoming direction. The arrival time of the ions to reach the detector is used to calculate the *m*/*z* ratio of the ions. Difference in the initial kinetic energy between ions is compensated by using a reflectron. The arrows denote the trajectories of two ions of the same *m*/*z* but different initial kinetic energy. Figure drawn by the author.

Orbitrap: The Orbitrap is effectively an ion trap that can measure *m*/*z* with high accuracy and resolution (**Figure 1.13**)[198,199]. The precursor of Orbitrap is the Kingdon trap[200], a device that traps ions using a static electric field, in contrast to a RF field or magnetic field found in other mass analyzers. The Kingdon trap consists of three components: a thin-wire central electrode, a coaxial outer cylindrical electrode and two endcap electrodes. A DC potential (U) is applied between the central and outer electrodes, leading to a radial electrostatic field that is described by:

$$U(r) = Aln(r) + B \qquad \textbf{Equation 1.5}$$

where r is the radial coordinate, and A and B are constants. The radial field causes ions with appropriate radial velocity to move in a stable orbit around the central electrode.

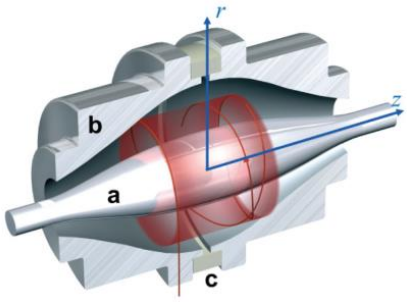

Figure 1.13 A schematic of an Orbitrap mass analyzer. Ions are injected into the Orbitrap perpendicular to the *z*-axis and then trapped by the electric field. The ions orbit around (a) the central electrode and oscillate in between (b) two isolated outer electrodes split by (c) an insulating ring. Image current received by the outer electrodes is Fourier transformed into the oscillating frequencies of the ions to produce an *m*/*z* spectrum. Figure from Scigelova *et al*.[201].

In the Orbitrap, the central electrode becomes a spindle-shaped electrode and the outer electrode is replaced by two complementary-shaped electrodes that are electrically isolated[198]. The electrostatic potential is described as:

$$U(r,z) = \frac{kz^2}{2} - \frac{kr^2}{4} + \frac{k}{2}(R_m)^2 ln\left[\frac{r}{R_m}\right] + C \qquad \textbf{Equation 1.6}$$

where k is the field curvature, R_m is the characteristic radius, and C is a constant. Ion packets are introduced into the trap perpendicular to the axial direction and with a displacement from the center. Because there are no cross-terms in r and z, the motion of the ions in the field is described as two independent motions: a radial orbital motion around the center electrode and an axial harmonic oscillation along the *z*-axis. The frequency of axial oscillation (ω) are measured and converted to the m/z of the ion based on the following relationship:

$$\omega = \sqrt{k(z/m)} \qquad \textbf{Equation 1.7}$$

As ions are focused in the C-trap and injected orthogonally into the Orbitrap, the effects of kinetic energy and spatial spread of the ions on the measurement of

the axial oscillation are minimized. Therefore, the Orbitrap provides high-resolution and high-accuracy mass analysis[202,203]. In addition, the Orbitrap is small and relatively inexpensive. These advantages make the Orbitrap an attractive analyzer in MS.

The Orbitrap integrates the function of both mass analyzer and detector[202]. As an ion oscillates axially in the Orbitrap, it generates an image current oscillating between the two outer electrodes. The coherent oscillations of a packet of ions with different *m*/*z* lead to a superimposed wave current. This time-domain signal is detected and Fourier transformed into a *m*/*z* spectrum, where the current is decomposed into individual frequencies to determine the *m*/*z* of each ion (**Equation 1.7**). Since the ions are trapped in the Orbitrap, detection time of the oscillating frequencies can be extended to improve the mass accuracy and sensitivity.

1.3.5. Detection

Detectors receive the electric current induced by the physical properties of ions (e.g., trajectory and oscillation frequency) and convert them into useful signals. Unlike the Orbitrap, which integrates the function of both mass analyzer and detector, many analyzers, including the quadrupole and ToF analyzer, require a separate detector. The most common stand-alone detector is the microchannel plate (MCP) detector, a continuous dynode electron multiplier containing many parallel cylindrical channels (**Figure 1.14**)[204]. Ions are deposited on the detector coated with a high-resistance semiconductor and their kinetic energy triggers the emissions of secondary electrons. The emitted secondary electrons collide

within the channels and release more electrons, leading to signal amplification. A potential (typically ~1 kV) is applied to the MCP to decouple the discharging and recharging of the plate due to the electron cascade.

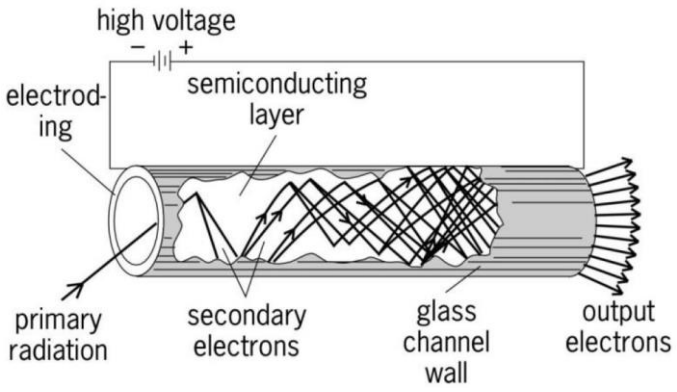

Figure 1.14 A schematic of the MCP working principle. A MCP consists of many channels. Collisions by particles lead to a cascade of secondary electron emission that amplifies the signal and produces a current. Figure from Ladislas Wiza *et al*.[205].

Although the detector response is linear to the number of ions arriving at the detector, the spectral intensity does not reflect the abundance of the corresponding species in solution. This is mainly due to the non-equivalent ESI ionization efficiency between different ion species. Therefore, MS is semi-quantitative rather than quantitative. Internal standards and/or isotopic labelling are commonly used to obtain quantitative MS measurements[206].

1.3.6. Interpretation of protein nESI mass spectra

Folded protein ions generated by nESI follow an envelope (typically Gaussian[207]) of consecutive charge states (**Figure 1.10**). These charge states are affected by many determinants, such as the geometry and fold-state of the protein. For

membrane proteins, their charge states are also dependent on their detergent/lipid environment. For example, membrane proteins solubilized in tetraethylene glycol monooctyl ether (C8E4) may show a lower average charge state than in n-dodecyl-β-D-maltoside (DDM). This is because when C8E4 dissociates from the protein-detergent complex, they can carry positive charges with them, leading to a charge reducing effect[208].

The charge state distribution of a protein ion leads to a series of signals with different *m*/*z*. In the positive polarity, a protonated protein ion with a protein mass of M and a charge of z has a mass of $M + z$ due to protonation. It produces a spectral signal at:

$$x_z = \frac{M+z}{z} \qquad \textbf{Equation 1.8}$$

Its neighboring charge state ($z + 1$) produces a signal at:

$$x_{z+1} = \frac{M+(z+1)}{z+1} \qquad \textbf{Equation 1.9}$$

Solving **Equation 1.8** and **Equation 1.9** results in:

$$M = \frac{(x_z-1)(x_{z+1}-1)}{x_z-x_{z+1}} \qquad \textbf{Equation 1.10}$$

According to **Equation 1.10**, the mass of a protein complex can be calculated from the *m*/*z* of two adjacent charge states.

1.3.7. Instrumentation

Orbitrap mass spectrometer: This thesis used exclusively the Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Q Exactive) (**Figure 1.15**) for all MS experiments with the exception of lipidomics analysis and tandem MS experiments. In the Q Exactive mass spectrometer,

ions are generated by a nESI source and directed to the cone. They are then guided and focused by the S-lens (stacked ring ion guide) and the bent flatapole operating in the RF-only mode. The configuration of the bent flatapole prevents transmission of neutral species. Ions arriving at the quadrupole are transmitted to the C-trap or mass-selected for tandem MS analysis. The C-trap acts as a pivot to focus and direct the ion packet to the higher-energy collisional dissociation (HCD) cell for collisional activation and then to the Orbitrap for mass detection. The instrument has several modifications in the electronic and HCD gas inlet as described by Gault *et al*.[209]. The overall outcome of these modifications enables the mass transmission of the quadrupole to be extended to 20,000 *m*/*z* and a higher pressure achievable in the HCD cell for better ion focusing/trapping. Pressure of the ultra-high vacuum (UHV) changes from the standard 10^{-11}–10^{-10} mbar operating pressures to ~10^{-9} mbar.

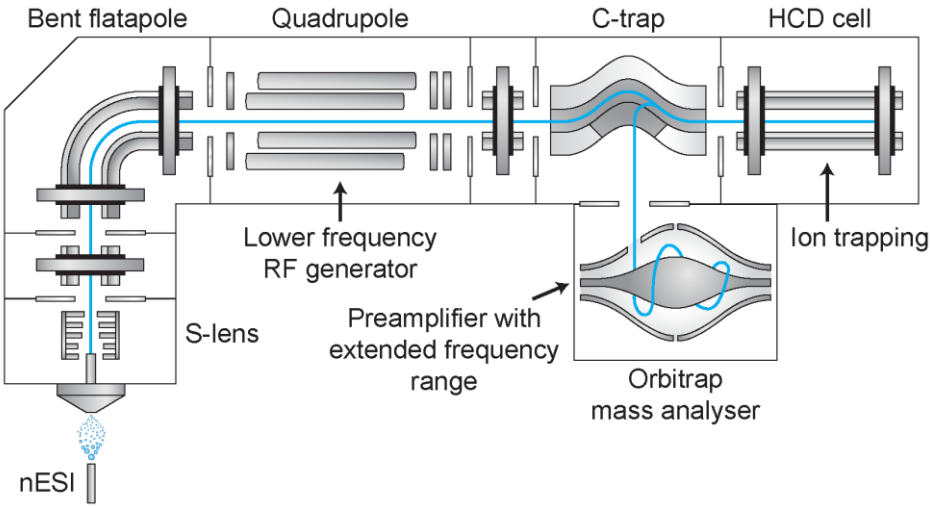

Figure 1.15 A schematic of the Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer. Ions are produced from the nESI source and transmitted to the HCD cell for collisional activation, followed by mass detection in the Orbitrap. The arrows denote modifications. Figure drawn by the author.

Linear trap quadrupole: A LTQ Orbitrap XL hybrid Ion Trap-Orbitrap Mass Spectrometer (LTQ Orbitrap XL) was used for lipidomics analysis. The architecture of the LTQ Orbitrap XL is similar to the Q Exactive. However, the quadrupole is proceeded by a linear quadrupole ion trap, where the mass selection and fragmentation by CID occur.

Synapt: A modified QqToF (Q and q indicate a quadrupole operating in the normal mode and RF-only mode, respectively) mass spectrometer (Synapt G2) (**Figure 1.16**) was used for tandem mass experiments of protein-lipid complexes. Ions are generated from a z-spray ion source and transmitted by a set of ion guides optimized for ion transmission and focusing. The z-spray configuration reduces the transmission of neutral species. Pumping efficiency at the front end of the instrument is reduced to elevate the pressure (~5 mbar[192]) for softer ionization conditions and collision focusing. Ions that arrive to the quadrupole are mass-selected or transmitted to the next quadrupole (i.e., trap) that serves as a collision cell. The RF of the quadrupole is reduced to 300 kHz to achieve an upper *m*/*z* transmission range to 32,000 *m*/*z*[192]. The trap operates in the RF-only mode and is charged with buffer gas, typically N_2 or Ar, for CID. The precursor and product ions are then sent to the ToF coupled with a MCP-type detector for mass analysis and detection.

Although not used in this thesis, the Synapt is equipped with an IM cell, which is used in IM-MS experiments. IM-MS experiments provide compelling evidence to support preservation of folded protein structure by native MS[210]. For example, CCS determined by IM-MS can be compared with the theoretical value

predicted from atomic coordinates[211,212]. CCS measurements of proteins in their native states can agree well with their conformation and oligomeric state[190,210].

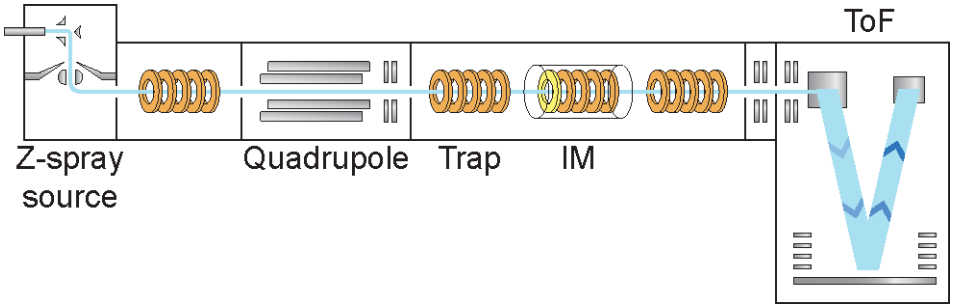

Figure 1.16 A schematic of the Synapt G2. Ions are generated from a z-spray source and transmitted to the quadrupole. At this point, the quadrupole can operate in a scanning mode or in a mass-selection mode (See **Figure 4.2**). Ions are then transmitted to the second quadrupole (trap) for CID. Finally, the precursor and product ions are sent to the ToF coupled with a MCP-type detector for mass analysis and detection. Figure drawn by the author.

1.3.8. Application of native MS in the study of protein-lipid interactions

As a biophysical tool, native MS represents an integration of high-speed, high-sensitivity and high-resolution separation techniques for the investigation of transient protein-lipid complexes in the context of complex mixtures[211]. Firstly, MS is a high throughput technique; typical MS experiments take a few seconds to a couple of minutes to obtain high-quality mass spectra. Secondly, the high sensitivity of MS affords MS measurements with low sample consumption[159]. This is particularly advantageous for membrane proteins due to their low expression levels and purification yields. Thirdly, while many techniques characterize protein interactions using exogenous lipids, MS can capture interactions of both exogenous and endogenous co-purified lipids[30,213]. This ability is useful to identify specific lipid interactions and establish their

relationship to protein function[63,214,215]. Another important advantage of MS is its ability to separate heterogeneous species in the *m*/*z* space, allowing interrogation of different sub-populations in complex mixtures. While biological samples are generally heterogeneous and often problematic to many structural biology approaches (e.g., cryo-EM), MS is compatible with versatile biological samples with complex compositions[216].

Given its unique abilities, native MS has revealed many important insights into membrane protein-lipid interactions. For example, native MS was used to study the selectivity of different lipids to membrane proteins[209]. Despite the small mass differences, CDLs bound to the sugar exporter semiSWEET were differentiated by their chain length using a high-resolution Orbitrap mass spectrometer. Comparison between the average chain length of the CDLs found in *E. coli* and observed in MS revealed a preference of semiSWEET to bind long chain CDLs.

In addition to comparing the affinity of different lipid interactions, the functional and structural effect of lipid interactions can be studied using MS. This is exemplified by a previous MS study of the ATP-binding cassette transporter, TmrAB. This study developed a controlled delipidation protocol and discovered that a large number of bound lipids were retained on TmrAB even after a prolonged delipidation process[87]. Interestingly, these lipids may be correlated to the unassigned density around the TMD of TmrAB as observed in the cryo-EM structure[217]. Quantitative lipidomics and native MS revealed the tightly bound lipids as annular PG and Lipid A derivatives at the substrate-binding site.

Further MS experiments demonstrated that the ATPase activity of TmrAB resulted in depletion of the Lipid A, indicating TmrAB may function as an ATP-dependent lipid flippase[87].

MS was also used to study the interactions of lipids of interest by addition of exogenous lipids. A previous study used IM-MS to study the effect of lipid-binding to the stability of different membrane proteins, which have distinct topologies, oligomeric states, and propensities for lipid binding[63]. In this study, proteins were collisionally activated, leading to unfolding and larger CCS value. The extent of unfolding was estimated by the CCS and used as an indicator of the protein stability. Protein stability was quantitatively compared between different lipid adducts and different lipid-bound states. It was discovered that the lipids significantly stabilized the protein structures correlated with their specific functional and conformational effect.

Complementary information obtained from native MS is useful to assign unknown density observed in X-ray crystallography and cryo-EM. High-resolution structure determination provides significant detail of protein structure and conformation. Lipids, due to their dynamic interactions, often yield less defined density and are difficult to assign. In the X-ray structure of the membrane rotor of the V-type ATPase from *E. hirae*, lipid density was attributed to 20 PG lipids, given that PG is the most abundant phospholipid in the organism[148]. However, this electron density was subsequently revealed as 10 CDL molecules located in the orifice of the K ring by MS[149]. The combination of

MS and high-resolution structure determination allows a more in-depth understanding of the molecular insight of protein-lipid interactions.

1.4. Membrane mimetics

1.4.1. Native MS requires use of membrane mimetics

Native MS is intrinsically an *in vitro* analytical technique, which transfers the samples from the solution-phase to the gas-phase for analysis. As membrane proteins have low solubility, they need to be solubilized prior to MS analysis. Detergents are conventionally employed to extract and solubilize membrane proteins, allowing the proteins to be ionized with nESI. The most common native MS method for membrane proteins is to electrospray them in detergent micelles (two critical micelle concentration (CMC) is typically used) and remove the detergents subsequently by CID[163,188]. This method enables preservation of non-covalent interactions and intact oligomers in the gas-phase. Importantly, membrane proteins liberated from micelles can retain their bound lipids[213], both co-purified endogenous lipids and exogenous lipids. This feature facilitates investigations of exogenous protein-lipid interactions as well as screening of different protein-lipid interactions[63,173,209].

As discussed in **Section 1.3.3**, membrane proteins are typically sprayed in MS compatible detergents that can be easily removed by CID, such as zwitterionic and non-ionic detergents. In the gas-phase, ionic interactions, hydrogen bonding, and polar interactions become stronger due to the absence of polar solvent molecules, whereas hydrophobic interactions become less important for intermolecular interactions. In the gas-phase, zwitterionic and non-ionic detergents form weaker interactions with membrane proteins than ionic detergents and therefore are generally easier to be removed by CID. Some non-ionic saccharide detergents, such as DDM, require relatively high CID

energies to dissociate, because they can form multiple hydrogen bonds with membrane proteins. For some membrane proteins, their solution compatible detergents may be different from those compatible with MS. Typically, buffer exchange with MS compatible detergents is carried out immediately prior to MS analysis. Alternatively, a detergent screen is performed to find a suitable detergent composition.

There are many different detergents and they are useful tools to study membrane proteins. Detergents form micelles and provide a suitable environment to maintain the membrane protein structure and function. Different detergents interact differently with membrane proteins, which may affect their stability, structure, function, and interactions with lipids. However, the choice of detergents may not necessarily influence the experimental observation. For example, many GPCRs were studied with different detergents and their structures remained reasonably folded; from these structures, many structural features of GPCR were learned[218]. Although detergents are widely used, cautions should be taken because micelle environments cannot fully replicate the native lipid bilayer environments.

In addition to detergents, there are a variety of membrane mimetics (**Figure 1.17**), each with its own advantages and disadvantages for application in MS. The application of these mimetics to native MS has been reviewed[219]. Detergents and amphipols (amphipathic copolymers of hydrophobic and hydrophilic monomer units) are both non-bilayer membrane mimetics. Both mimetics were useful in preserving the structure of membrane proteins and

protein-lipid interactions in the gas-phase[220-222]. However, since the structures of these synthetic molecules are distinct from lipids, they do not necessarily provide an environment comparable to the native biological membrane. Following reconstitution in these mimetics, membrane proteins might struggle to preserve their native structure, function, and lipid interactions[220].

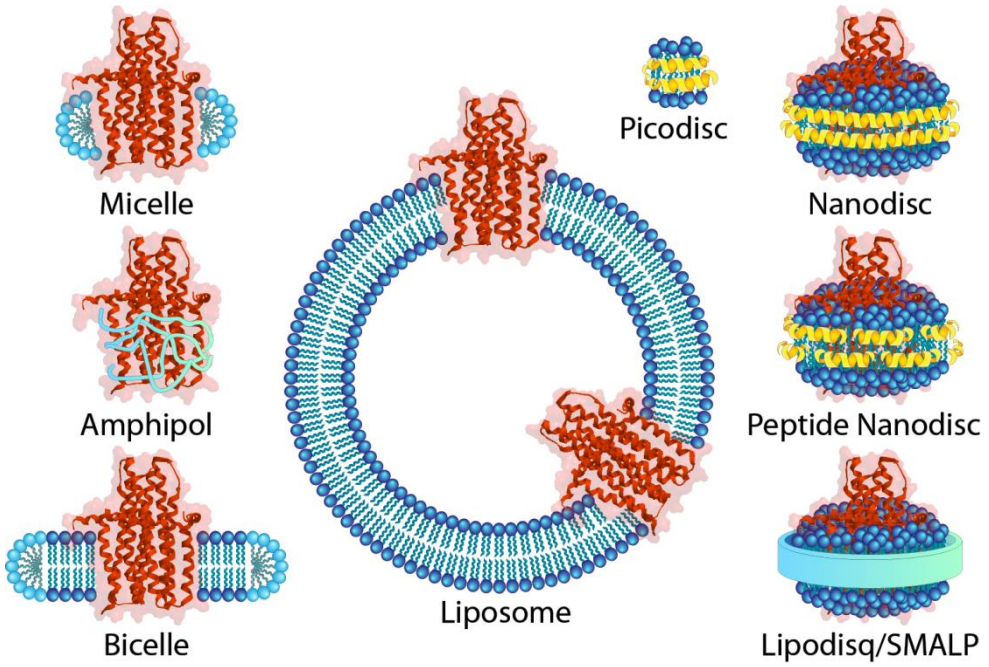

Figure 1.17 A Schematic of various membrane mimetics used for membrane proteins. Structures of proteins (*red*), scaffold protein/peptides (*yellow*), lipid bilayers (*blue*), and synthetic chemicals (detergents, amphipods, and styrene-malic acid copolymers) (*cyan*) are indicated by different colors. Figure adapted from Marty *et al.*[219] and drawn by the author.

In contrast to detergents and amphipols, liposomes, bicelles, nanodiscs, and styrene maleic acid co-polymer lipid particles (SMALPs) offer their reconstituted membrane protein a lipid bilayer that better models the native membrane environment. These mimetics allow membrane proteins and protein-lipid interactions to be studied in lipid bilayer environments. Liposomes are large

vesicles, which form compartments similar to organelles and plasma membranes. However, they are difficult to study with native MS due to their large size and polydispersity. Bicelles are discoidal lipid bilayers capped with detergents[223]. Although having a bilayer structure, bicelles are polydisperse and not detergent-free. SMALPs are a relatively new mimetic system that has a discoidal lipid bilayer stabilized by SMA[224,225]. SMA can directly interact and extract membrane proteins together with their native membrane environment from the membrane. Although SMALPs provide an opportunity to study protein-lipid interactions from the native environment, MS methods to analyze SMALPs have been difficult to establish.

Nanodiscs are a discoidal lipid bilayer enclosed by two amphipathic membrane scaffold proteins (MSPs) in a "belt" conformation[226-230]. MSPs are engineered apolipoprotein A1 (Apo-A1) that can solubilize and stabilize a patch of lipids. Nanodiscs are formed by a self-assembly process in which MSPs, lipids, and membrane proteins are transiently solubilized in detergents. By gradually depleting the detergents, proteins and lipids self-assemble into nanodiscs. Because MSPs have a fixed length and hydrophobic surface area, nanodiscs tend to be formed with a specific lipid/MSP ratio to maximize the hydrophobic interactions between the MSPs and lipids[227]. The polydispersity of nanodiscs is around ± 5 lipids per complex[231]. This renders nanodiscs lower polydispersity than other mimetics. Besides, the lipid composition of nanodiscs can be precisely controlled, allowing investigation of protein-lipid interactions in various lipid environments[232]. Therefore, while nanodiscs provide a detergent-free and native-like lipid bilayer with a controllable lipid composition, they are relatively

homogeneous and monodisperse. Nanodiscs are an appealing mimetic for different membrane proteins and a broad range of spectroscopy techniques, including MS[233,234].

1.4.2. Native MS studies of nanodiscs

Nanodiscs provide a promising platform to study membrane proteins in native MS. An early work applying native MS to examine membrane protein complexes reconstituted in different membrane mimetics revealed that more membrane protein oligomers were preserved in the gas-phase using nanodiscs than detergents and amphipols[220]. Membrane proteins liberated from nanodiscs had a lower charge state than when in detergents, implying that the proteins were in a more compacted and folded state. Further, nanodiscs were employed to study the interactions between peripheral membrane proteins and glycolipids by a catch-and-release experiment[235-237]. Here, glycolipid nanodiscs were incubated with the soluble protein of interest and ionized by nESI. Protein-nanodisc complexes were activated to liberate the proteins with bound glycolipids, allowing determination of the stoichiometry and affinity of the lipid interactions. The protein-glycolipid complexes can be isolated and further activated to release the individual glycolipids for identification.

In contrast to the two studies above, intact nanodiscs without membrane proteins were preserved in the gas-phase and characterized by native MS[231,238]. Their spectra were complex and had a board range of *m*/*z* values (**Figure 1.18** A). This was because nanodiscs when ionized produced multiple charge states. Each charge state contained an ensemble of nanodiscs that differed by the

number of lipids per complex (e.g., complexes with 140–160 lipids) and gave rise to regularly spaced peaks in the spectra (**Figure 1.18** B). Adjacent charge states overlap effectively at around integer multiples of the mass of the lipid (**Figure 1.18** C; see also **Figure 3.4**). For example, ions from a nanodisc with two MSPs (44,088 Da) and 165 DMPC (678 Da) at 23+ charge state has a *m*/*z* of 6,781 (an integer multiple of 678) which is equal to the *m*/*z* of other ions at the adjacent charge state 22+ with 155 DMPC (peak 1). Whereas ions with 166 DMPC at 23+ charge and ions with 156 DMPC at 22+ charge appear at 6,810 *m*/*z* and 6,812 *m*/*z*, respectively (peak 2). The overlap between different charge states leads to a characteristic “resonance” in the nanodisc spectra.

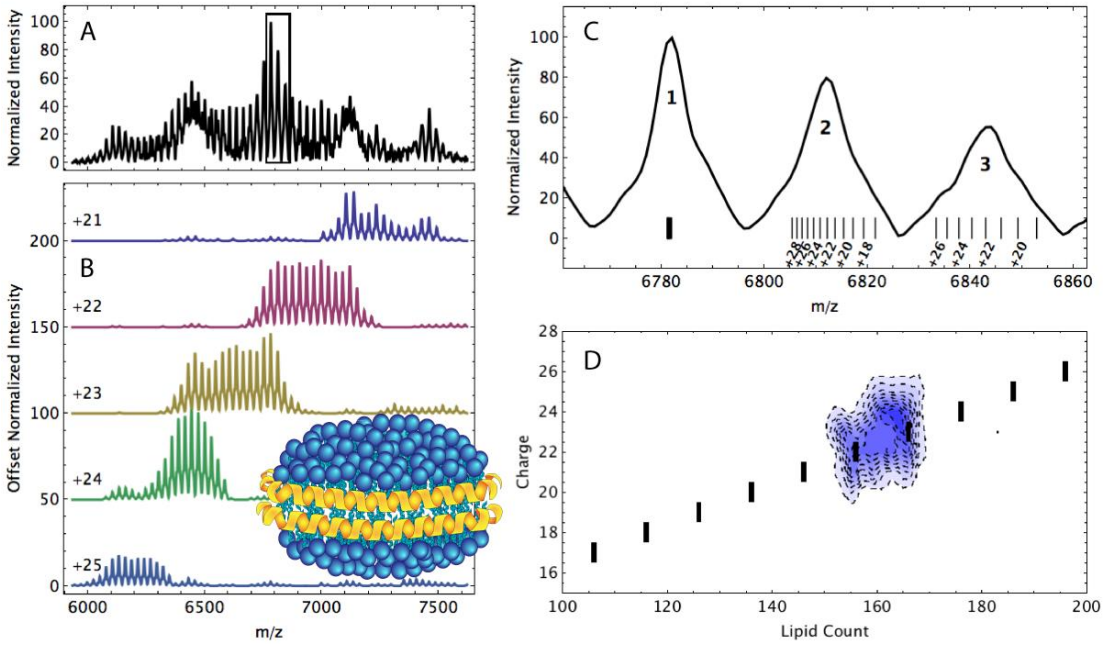

Figure 1.18 Illustration of the charge state overlap in nanodisc spectra. (A) Spectra of nanodiscs contain (B) multiple overlapping charge states. (C) The boxed region in (A) is expanded, showing that the overlap between charge states is more significant at peak 1, which is at the integer multiple of the mass of the lipid. Potential charge states of peak 1, 2, and 3 are denoted below the peaks. (D) Deconvolution reveals the lipid counts and charge states in (C)

with the intensity shown by contours. The potential combinations of charge states and lipid counts that overlap at peak 1 are denoted by rectangles. Figure from Marty *et al*.[219].

Due to the overlap between charge states, mass spectra of nanodiscs cannot be interpreted using conventional methods for mass assignment (**Equation 1.10**). Universal Deconvolution (UniDec), which is a Bayesian deconvolution algorithm based on the Richardson-Lucy algorithm, was employed to deconvolve the overlapping charge states[231,239]. UniDec describes mass spectra as a convolution between a peak shape and a set of delta functions in a *m*/*z* versus *z* matrix (**Figure 1.19**). It first generates a Bayesian prior (a *m*/*z* versus *z* matrix) with an equal probability for all charge states. The prior is then processed with successive iterations of three steps to determine the contribution of each *m*/*z* and charge state to the overall spectrum.

During the first step, charge distributions in the prior are smoothed by a filter. This step is based on the assumption that ions produced by ESI follow a charge state envelope which is typically Gaussian[207]. The abundance of a complex at charge state *z* is related to the abundance at $z - 1$ and $z + 1$. Similarly, as nanodiscs have a distribution of the number of lipids per complex, a second filter can be applied on the *m*/*z* dimension based on the mass of the lipid. In the second step, the prior is summed along the *z*-axis and convolved with a peak shape to project a simulated *m*/*z* spectrum. Finally, each element in the prior is multiplied with the intensity ratio of the spectral data over the projected spectrum at its specific *m*/*z* value. Following iterations, the algorithm converges when the projected spectrum matches the subject spectrum. The corresponding prior obtained after iterations defines the contribution of each charge state at

each *m*/*z*. This *m*/*z* versus *z* matrix can be converted to a *m* versus *z* matrix for simplicity. Overall, UniDec can separate the overlapped charge states in nanodisc spectra in a two-dimension matrix, allowing interpretation of the spectra.

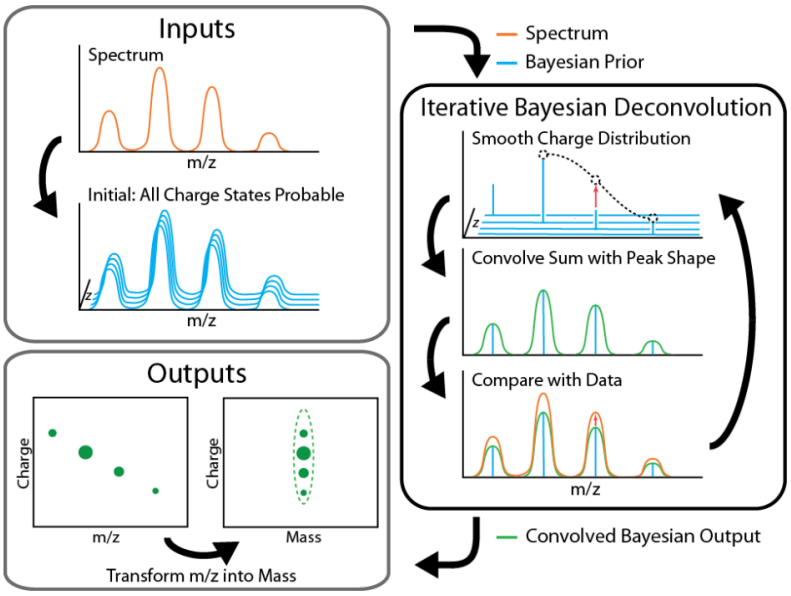

Figure 1.19 Illustration of UniDec deconvolution steps. The algorithm first generates a prior assuming equal probability for all charge states. The iterative Bayesian deconvolution proceeds with three steps. First, the prior is smoothed by a charge state filter and/or an oligomeric state filter. Second, the prior is summed along the *z*-axis and convolved with a peak shape to produce a projected spectrum. Third, the simulated spectrum is compared with the acquired spectrum and the intensity of the prior is adjusted prior to next iteration. Once converged, the prior is convolved with the peak shape to generate the *m*/*z* distribution for individual charge states. The *m*/*z* versus *z* matrix can be transformed into a *m* versus *z* matrix. Figure from Marty *et al*.[239].

1.5. Thesis objectives

Biological membranes are heterogeneous and dynamic systems governed by numerous transient non-covalent interactions. The unique abilities of native MS in the preservation of non-covalent interactions and separation of heterogeneous complexes enable direct interrogation of these non-covalent interactions, in particular, the membrane protein-lipid interactions. With the development in instrumentation, deconvolution strategy, membrane mimetics, and methodology, native MS has been applied to study many membrane proteins alongside their lipid interactions.

Although most developments of native MS for membrane proteins were based on detergent-solubilized proteins, nanodiscs present a promising alternative to detergents as a membrane mimetic. In addition to a more native-like environment, nanodiscs provide an additional benefit as their lipid environments can be precisely controlled. Although this is a particularly useful feature for characterizing the effect of lipid environment to membrane proteins, nanodiscs containing a lipid mixture are heterogeneous and challenging to analyze. Therefore, the initial aim was to establish a framework to enable native MS analysis of heterogeneous-lipid nanodiscs and their lipid compositions.

Following the study on nanodiscs without membrane proteins, nanodiscs of two model membrane proteins were studied using native MS. Despite sharing some similarity, the spectra of nanodiscs with membrane proteins are significantly more complex than nanodiscs devoid of embedded membrane proteins. Therefore, the second objective of this thesis was to establish a framework to

analyze these spectra for interrogation of the protein-lipid interactions. In addition, the dissociation behavior of nanodiscs was explored.

In parallel to nanodisc-based MS methods, recent developments in detergent-based methods have established detergent and MS conditions to enable, for the first time, the study of a class A GPCR by native MS[240]. Using these conditions, three different class A GPCRs were studied. By combining native MS with a variety of techniques and strategies, such as mutagenesis and multiplexing strategy, this thesis aims to provide a comprehensive characterization of the lipid interactions of class A GPCRs and GPCR-transducer complexes. Based on these data, a novel PIP_2-binding motif was found and its role in GPCR-G protein coupling was explored.

In summary, the thesis adopts native MS in combination with a range of techniques and methodologies to extend its scope in the study of membrane protein-lipid interactions, which reveals novel molecular insights into protein-lipid interactions.

Chapter 2. Characterization of heterogeneous-lipid nanodiscs using native MS

2.1. Introduction

Transient and weak lipid interactions are gaining attention for their ability to tune the biological function of membrane proteins[81,145]. Previous studies have shown that non-covalent (or native) mass spectrometry (MS) can provide insight into protein-lipid interactions and their structural implications for membrane proteins[30]. Briefly, membrane protein complexes solubilized in detergents are introduced to the gas-phase and liberated from detergent micelles by collisional activation[188]. By carefully optimizing the introduction and liberation of protein complexes, a wide range of non-covalent interactions, including protein-lipid interactions and drug binding, can be preserved[173]. Although MS analysis uses detergents almost exclusively to deliver membrane protein complexes into the gas-phase, alternative membrane mimetics present unique opportunities to investigate lipid interactions in lipid bilayer systems[219,220,222].

Nanodiscs offer a promising alternative to detergents for native MS studies. Nanodiscs are nanoscale lipid bilayers encircled by two membrane scaffold proteins (MSPs), which can range from 8 to 16 nm in diameter[230,241]. This structure offers the encapsulated membrane proteins a native-like lipid bilayer environment for proper folding and function[230,242]. Therefore, nanodiscs allow characterization of membrane protein structures and interactions in the context of the lipid bilayer. The initial nanodisc-based MS method was developed by Marty *et al*.[238], who established a Bayesian deconvolution algorithm (UniDec) to analyze spectra of intact nanodiscs without membrane proteins and determine

the stoichiometry of lipids per nanodisc[231,239]. A pioneering MS study demonstrated that membrane protein complexes in nanodiscs could be released by collisional activation for characterization[220]. These studies indicated nanodiscs are a compatible and promising membrane memetic for studying membrane proteins and their interactions with lipids.

Although prior studies commonly use nanodiscs with a homogeneous lipid composition, nanodiscs can be formed by heterogeneous lipid mixtures at precisely defined lipid compositions[232]. This feature provides an opportunity to assess the effect of lipids on membrane protein function, such as the effect of lipid head group chemistry and acyl chain length. However, the application of heterogeneous-nanodiscs brings additional analytical challenges to MS studies. First, heterogeneous lipids increase the polydispersity of nanodiscs, which complicate mass spectra, as shown in a study where nanodiscs with heterogeneous glycolipids were used to investigate the selectivity and binding kinetics of soluble protein-glycolipid interactions using MS[237]. While the liberated protein-glycolipid complexes were well resolved, the ions corresponding to the heterogeneous nanodiscs appeared as an unresolved peak that was impossible to analyze. Secondly, spectra are more difficult to interpret and assign, as mass alone may not be adequate to determine the lipid composition in nanodisc complexes. Lastly, lipid composition may change the gas-phase behavior of nanodiscs, especially how nanodiscs ionize and dissociate.

Here, a framework was presented to overcome the challenges in analyzing heterogeneous-lipid nanodiscs. This framework relies on the determination of the average lipid mass of nanodiscs that serves as a reporter of the lipid composition. Two novel methods were applied to measure the average lipid mass and their results were compared. Using this framework, nanodiscs with nine different lipid compositions were analyzed, revealing the relationship between the dissociation mechanisms of nanodiscs, lipid compositions, and instrumental polarities.

2.2. Methods

2.2.1. Sample preparation and MS analysis

Material: In this thesis, all materials and reagents were purchased from Sigma Aldrich (St. Louis, MO) unless otherwise stated. Detergents, n-Dodecyl-β-D-maltoside (DDM), n-Octyl-β-D-Glucoside (OG), 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (LMNG), and n-Hexadecylphosphocholine (Fos-Coline-16) were purchased from Anatrace (Maumee, OH). The critical micelle concentrations (CMC) of the detergents are provided in **Table 2.1**. All chromatography columns were purchased from GE Healthcare (Piscataway, NJ). Concentration was performed using ultrafiltration devices (Amicon) of appropriate molecular weight cut-off (MWCO): 10 kDa for MSPs, 50 kDa for NTSR1, 100 kDa for the AmtB, AqpZ, and all nanodiscs. All protein concentrations were determined by absorbance at 280 nm. Filtration steps used syringe filters (Millex) or spin filters (Spin-X) of 0.22 µm pore size. Centrifugal buffer exchange were performed prior to MS measurements using a Micro Bio-Spin^{TM} 6 column (for nanodiscs; Bio-Rad, Hercules, CA) or Zeba^{TM} Spin Desalting column 40k MWCO resin (for GPCRs; Thermo Fisher Scientific Inc., Rockford, IL, USA).

Table 2.1 Critical micelle concentrations of detergents. Table made by the author.

Detergent	CMC (w/v)	Detergent	CMC (w/v)
DDM	0.01%	LMNG	0.001%
OG	0.53%	Fos-Coline-16	0.00053%

Lipid: Powdered lipids of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (POPG), and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine (POPS) were products from Avanti Polar Lipids (Alabaster, AL). Lipid stocks were prepared in either chloroform (for POPC and DMPC) or 1:5 (v/v) methanol:chloroform (for POPG and POPS) and stored at −20 °C. Lipid concentrations were measured by phosphate analysis[243]. Dissolved lipid stocks were prepared as dried lipid films by nitrogen flow and overnight storage in a vacuum desiccator. The lipid films were re-solubilized in 0.1 M sodium cholate solution prior to reconstitution into nanodiscs. The final concentration of lipid was 0.05 M.

MSP expression and purification: Constructs of MSP1D1 and MSP1E3D1 were obtained from Dr. Michael Marty[2,227,241] (see Denisov *et al*. for detail[227]). Briefly, the constructs were generated by cloning a sequence of MSP preceded by a 6× histidine affinity tag and a tobacco etch virus (TEV) protease recognition site into a pET-28a vector. The constructs were transformed into *E. coli* BL21-Gold (DE3) strain (Bio Labs New England) by heat shock for protein expression.

All cells, proteins and nanodiscs were handled and purified at 4 °C, unless otherwise stated. Membrane scaffold proteins, MSP1D1 and MSP1E3D1 were expressed and purified following well-established protocols from the Sligar laboratory[241]. *E. coli* transformed with the MSP plasmid grew in Terrific Broth medium supplemented with resistant antibiotic (30 μg/mL kanamycin) at 37 °C,

shaking at 200 rpm (Innova incubator shaker). At OD_{600} of 0.6, protein expression was induced by adding 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG; 1 M IPTG stocks). After 3 h expression, cells were harvested by centrifugation (5,000 ×*g*, 10 min), re-suspended in buffer (20 mM Tris pH 7.4, 150 mM NaCl; 25 mL per liter of cell culture) and lysed using a microfluidizer M-110P operating at 22,000 PSI (Microfluidics Corp.). Cell debris in the lysate was removed by centrifugation (20,000 ×*g*, 20 min). The supernatant was collected.

MSPs were purified from the supernatant by immobilized metal affinity chromatography (IMAC) using a HisPrep FF 16/10 column equilibrated with loading buffer (40 mM Tris pH 7.4, 300 mM NaCl). Immobilized MSPs were washed with three buffers (loading buffer with 1% (v/v) Triton™ X-100; loading buffer with 50 mM sodium cholate; and loading buffer with 25 mM imidazole; 250 mL for each buffer), followed by elution using loading buffer with 0.5 M imidazole. Eluent was pooled and dialyzed in loading buffer overnight to remove imidazole (10 kDa MWCO Slide-A-Lyzer™ Dialysis Cassettes; Thermo Scientific). The MSPs were incubated with 1:100 w:w TEV protease with a histidine tag (Sigma) overnight and purified by reverse immobilized metal affinity chromatography (r-IMAC) to remove the polyhistidine tag. Tag removed MSPs were loaded on a HisPrep FF 16/10 column equilibrated with loading buffer. Flow through of r-IMAC containing the MSP was pooled together. Tag removed MSPs are referred to as MSP1D1(−) and MSP1E3D1(−). Cleaved MSP was buffer exchanged in 20 mM Tris pH 7.4 and 150 mM NaCl by dialysis (10 kDa MWCO) overnight. The MSP was concentrated to ~400 μM, snap freeze, and storage at −80 °C.

Nanodisc reconstitution and purification: Nanodiscs used here were formed using MSP1D1(−) with POPC-POPG or POPC-POPS at compositions ranging from 0 to 100% molar ratio of POPG or POPS at 25% intervals (**Figure 2.3**). They were prepared according to well-established protocols[2,241].

Cholate-solubilized lipids were prepared as described above and mixed at appropriate ratios prior to reconstitutions. For each sample preparation, 1 mL of purified MSP1D1(−) (400 µM stock) was used. Lipids (0.05 M) were added to MSP1D1(−) at a molar ratio of 65:1. The mixtures were allowed to incubate on ice or at room temperature (nanodiscs containing PS only) for 20 min, followed by overnight rotation with 1:1 (v/v) Amberlite XAD-2 hydrophobic beads to initiate nanodisc self-assembly.

The mixtures were filtered to recover the supernatant containing the nanodiscs. Nanodiscs were concentrated and injected on a Superose 6 Increase 10/300 size exclusion chromatography (SEC) column equilibrated with 0.2 M ammonium acetate (pH 7.4) (**Figure 2.1**). Peak fractions were pooled and concentrated to 5–25 µM. Purified nanodiscs were directly used for MS measurements or flash frozen and stored at −80 °C. Occasionally, nanodiscs needed an additional centrifugal buffer exchange (0.2 M ammonium acetate pH 7.4) to reduce sodium adducts observed in the mass spectra.

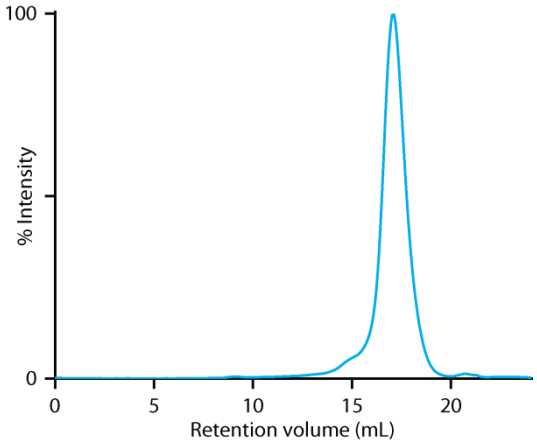

Figure 2.1epresentative SEC profile (Superose 6 Increase 10/300 column) of empty nanodiscs. Figure drawn by the author.

Mass spectrometry: All spectra were obtained from three biological replicates to establish reproducibility. MS analysis of nanodiscs was performed using a modified Q-Exactive™ hybrid quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific Inc., Bremen, Germany) optimized for the analysis of high *m*/*z* ions (**Figure 1.15**) (detail description can be found in refeneces[203,209,244,245]). Briefly, the instrument applied a lower RF frequency to the quadrupole and used a gas line supply to achieve higher pressures in the HCD cell. Electronics were also modified to extend the frequency detection range of the Orbitrap.

Nanodiscs (purified in 0.2 M ammonium acetate pH 7.4) were loaded on gold capillary needles prepared in-house and introduced to the mass spectrometer through a static nano-spray source[246]. Voltage applied on the capillary was 1.4 kV in both positive and negatively polarity. Both polarities used the same parameters except switching the relevant voltages to the respective polarity. The temperature of the ion transfer capillary was set to 200 °C. The S-lens RF

potential was set to maximum (200 V) to bias transmission of high *m*/*z* ions. The quadrupole was operated with a selection window of 2,500 to 20,000 *m*/*z*. Ions were trapped in the HCD cell operated in native mode for collisional activation. The HCD cell was filled with Argon and voltages were applied from 0 to 190 V in 10 V steps with a two-minute acquisition time at each step. Pressure reading of UHV was 1.05×10^{-9} mbar. The C-trap entrance lens was set to 6 V. The automatic gain control was set to 10^6 with a maximum injection time of 20 ms. Ion signals were recorded between 2,500 to 20,000 *m*/*z* at a resolution of 17,500 at *m*/*z* = 200 (a transient time of 64 ms) and noise level parameter of 3. Each spectral scan was an average of 10 micro-scans. Mass spectral data were acquired as Thermo raw files. All spectra were summed over a two-minute collection window without additional processing. The mass spectrometer was calibrated using CsI clusters.

2.2.2. Data analysis

Deconvolution: Mass spectral data were converted into mzML format by ProteoWizard and then text files by pymzml[247,248]. Mass spectra were deconvolved by UniDec software similar to previous studies of homogeneous-lipid nanodiscs (available for download at http://unidec.chem.ox.ac.uk)[2,239]. Prior to deconvolution, spectra were smoothed with a Gaussian (σ = 5 data points) and subjected to a curved baseline subtraction with a width of 100[249]. No linearization or further processing was performed. Deconvolution used a mass range of 4–280 kDa, a charge range of 1–30, a Gaussian peak shape with a 10 Th (Thomson) full-width at half-maximum (FWHM), and an oligomer mass (a parameter used in the software that defines a “guess” of the average lipid mass,

m_l) determined by dual Fourier transform as discussed below. The Bayesian priors were smoothed in the charge dimension using a Gaussian filter with a width of σ = 0.5 and in the mass dimension using a log mean filter with a width of 3 × m_l. By deconvolution, MS signals overlapped in the *m*/*z* space were resolved in a mass versus charge space (**Figure 2.2** A and B), where the mass-axis was linearized at 10 Da intervals. Deconvolution affords a simplified view of the mass spectra. Summing the mass versus charge space along the x-axis and y-axis gave the distribution of charge and mass observed in the mass spectra (**Figure 2.2** C and D). These distributions were monitored over the range of collision voltages applied in the measurements (**Figure 2.2** E and F).

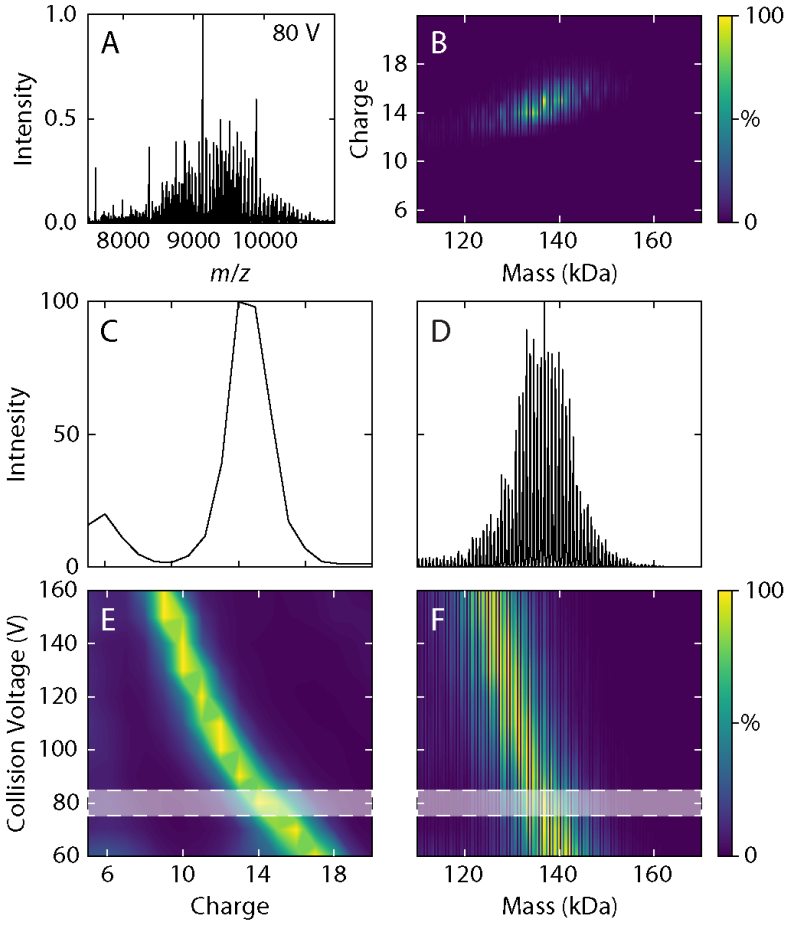

Figure 2.2 Analysis of nanodisc mass spectra by deconvolution. (A) The raw mass spectra of a POPC nanodisc collected at 80 V is deconvolved into (B) an array of mass versus charge. Summing the intensities along the x-axis and the y-axis gives (C) the charge distribution and (D) the mass distribution, respectively. (E) and (F) show the charge distribution and mass distribution when voltages are applied from 60 to 160 V. Boxed areas correspond to the distributions in (C) and (D). Figure from Hoi *et al*.[1] and drawn by the author.

Determination of the average lipid mass (m_l): The average masses of the lipids (m_l) in nanodiscs were measured using autocorrelation and dual Fourier transform (Dr. Marty suggested the methods and created tools in UniDec; the author carried out the analysis). For the first approach, mass spectra were deconvolved with UniDec using the oligomer mass ranging from 744 to 767 Da at 0.1 Da intervals. 744 to 767 Da was used since the theoretical minimum and maximum of the average lipid mass are 749 (POPG) and 762 Da (POPS). Mass distributions were calculated from the deconvolved *m*/*z* versus *z* matrixes and linearized at 0.1 Da intervals. The resulting mass distributions (**Figure 2.4** A in **Section 2.3.2**) were processed by autocorrelation (The autocorrelation function is integrated into UniDec). Autocorrelation computed the correlations of a shifted mass distribution with its original as a function of the mass shift (τ) (**Figure 2.4** B), which is expressed as:

$$R(\tau) = \int_{-\infty}^{\infty} f(u)f(u+\tau)du \qquad \textbf{Equation 2.1}$$

where $R(\tau)$ is the autocorrelation at a mass shift τ and $f(u)$ represents the intensity at u Da in the mass distribution. The correlations were calculated for mass shifts at 0.1 Da steps. Because the theoretical average lipid mass was within 749 to 762 Da, the m_l was determined as the maximum in the autocorrelation spectra between 740 to 770 Da. The measured m_l were plotted

as a function of the oligomer mass used in the deconvolution (a “guess” of m_l) (**Figure 2.4** C). The experimental average lipid mass was determined by finding the measured m_l that is equal to the “guess” (i.e., a coordinate with equal ordinate and abscissa).

For the second method, mass spectral data were linearized (to create linear approximation) by interpolating the spectrum at 0.01 Th steps. Nanodiscs signals were isolated by truncating the *m*/*z* region below 6,500 *m*/*z* (**Figure 2.5** A). Centroids of the spectra were estimated as the peak position of the spectrum smoothed with a broad Gaussian (σ = 1,000 Th). Spectra were multiplied by a Gaussian window (σ = 1,500 Th) at the centroid (**Figure 2.5** *orange*). A dual Fourier transform approach was then performed with two sequential fast Fourier transforms (FFT) on the processed mass spectra (**Figure 2.5** B and C). The FFT employed the FFTPack in Scipy[250,251]. Experimental m_l was measured as the maximum between 740 to 770 Da in the second FFT (The dual Fourier transform method is a built-in function of UniDec).

For both autocorrelation and dual Fourier transform methods, the average lipid masses of nanodiscs were determined by averaging the m_l measured from spectra acquired at 60 to 100V CID over three replicates (**Figure 2.4** D and **Figure 2.5** E). These results were analyzed by linear regression. The lipid compositions of the nanodiscs were interpolated from the m_l and theoretical masses of POPC, POPG, and POPS (**Figure 2.5** F).

The m_I of local m/z region were analyzed by applying a sliding Gaussian window prior to the dual Fourier transform (**Figure 2.6** E to G). The resulting second Fourier transforms were plotted as a function of the centers of the Gaussian window. The ion intensity in the second Fourier transform is normalized by the respective intensity in the spectra smoothed with a broad Gaussian (σ = 1,000 Th).

2.3. Results and discussion

2.3.1. Using lipids of similar masses to obtain resolved mass spectra

This work used MS to study nanodiscs in a similar way to a previous study[231]. Briefly, nanodisc samples were produced following well-established protocols (the author received supervision and training from Dr. Michael Marty)[241]. Nanodiscs were introduced into a high-resolution Orbitrap mass spectrometer by nESI[203,209] (Note that the previous study used a Fourier-transform ion cyclotron resonance mass spectrometer). Ions were activated by collisions with argon in the HCD cell by applying voltages ranging from 60 to 190 V at 10 V steps. The resulting product ions were detected by the Orbitrap mass analyzer. Acquired mass spectra were analyzed as stated below.

One of the major challenges in MS analysis of nanodiscs is highlighted by the polydispersity of the system. Because the self-assembly of nanodiscs is a stochastic process, nanodiscs present a distribution in the number of lipids per complex. A previous MS study of homogeneous-lipid nanodiscs has demonstrated that the polydispersity of the lipid stoichiometry is around ± 5 lipids per complex. Heterogeneous-lipid nanodiscs have an additional dimension of polydispersity because the resulting complexes of a given lipid count exist in multiple sub-states of discrete lipid compositions. For instance, heterogeneous-lipid nanodiscs prepared with 50% POPC-POPG could have multiple states (64:61, 63:62, 62:63, 61:64, and so forth) that each gives a total of 125 lipids in the complex. This additional dimension of polydispersity leads to an increase in the number of possible ion species. Thus, mass spectra may become more complex and difficult to resolve.

Although heterogeneous-lipid nanodiscs are difficult to analyze, it is possible to analyze heterogeneous-lipid nanodiscs that contain lipids of similar masses. This study used nanodiscs reconstituted with POPC mixed with 0 to 100% POPG or POPS at 25% intervals (**Figure 2.3** A). The mass difference of POPC to POPG and POPS are 11 and 2 Da, respectively. Due to the small difference in the lipid mass, compositional discrete sub-states of the heterogeneous-lipid nanodiscs merged into a single peak that defined all nanodiscs with the same lipid count (**Figure 2.3** B). Mass spectra of heterogeneous-lipid nanodiscs thus resembled those of homogeneous-lipid nanodiscs and the peaks were resolved at a similar extent (**Figure 2.3** B). This approach compresses the discrete sub-states, assuring mass spectra with manageable resolution.

Mass spectra of nanodiscs are complicated due to overlapping charge states (**Figure 1.18**)[231]. Hence, manual assignments are often impractical. A previous study developed a Bayesian deconvolution algorithm, UniDec, and used it to unravel the overlapping charge states in the spectra of homogeneous-lipid nanodiscs (see **Section 1.4.2** for review of the analysis homogeneous-lipid nanodiscs by deconvolution)[239]. Similarly, UniDec was used here to separate the overlapping charge states, allowing mass assignments for each *m*/*z*. The program returned a matrix of mass against charge as shown in **Figure 2.2** A and B. The matrix can be reconstructed as a charge distribution or mass distribution (zero-charge mass spectrum) that shows the charges or masses of all the ions detected in the mass spectra (**Figure 2.2** C and D).

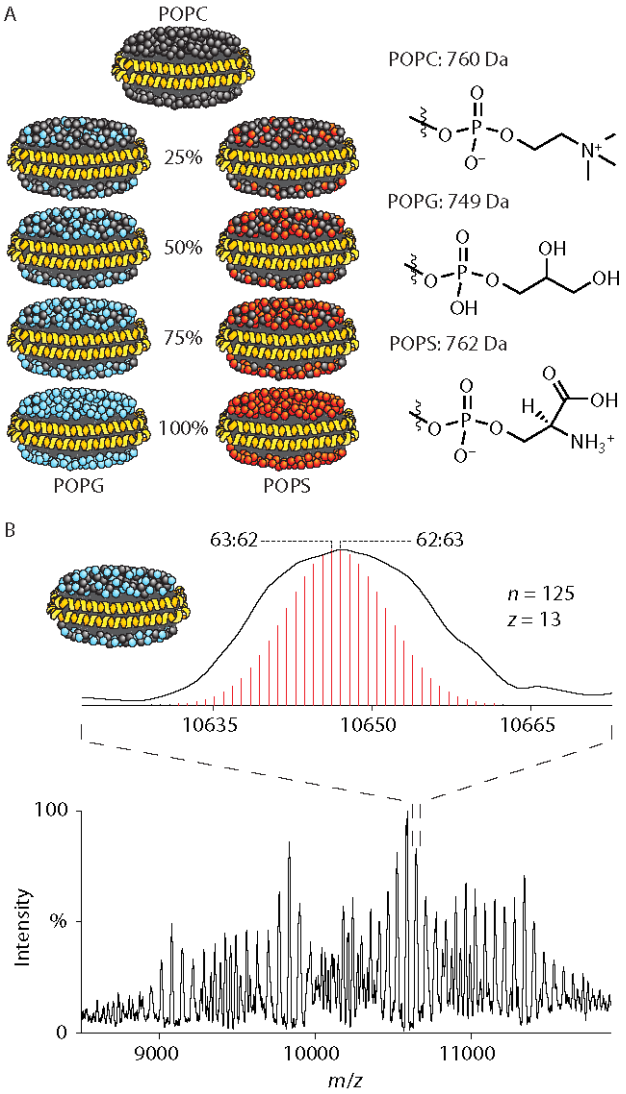

Figure 2.3 Schematic of heterogeneous-lipid nanodiscs and illustration of the sub-states with the same total lipid counts but different lipid compositions. (A) MSP1D1(−) (*yellow*) nanodiscs were made using pure or two component combinations of POPC (*grey*), POPG (*blue*), and POPS (*red*). The head group structures and masses of the lipids are shown on the right. The head groups are protonated/deprotonated to give a zero net charge as expected for lipids in nanodisc ions. (B) The raw mass spectrum of 50% POPC-POPG nanodisc was acquired at 80 V CID. The zoom-in region shows the peak primarily resulting from the nanodiscs of 125 lipids (*n*) per complex at a charge state of +13 (*z*). The red bars simulate the theoretical binomial distribution of sub-states that have the same total number of lipids but different combinations of POPG and POPC (e.g., 63:62 and 62:63). Figure from Hoi *et al.*[1] and drawn by the author.

2.3.2. Determining the average lipid mass of nanodiscs by autocorrelation

Although a set of masses was obtained from the spectra, unambiguous interpretation and composition assignment require the determination of the stoichiometries of the MSP and individual lipid constituents. The masses of the detected complexes approximated to the expected mass of an intact complex. This indicated that nanodiscs remain intact and retain both MSP scaffolds, which is consistent with what was observed in previous studies[231,238]. In contrast, it is less straightforward to determine the loading of each lipid in the complexes. Considering a most extreme case where lipids of the same mass are used, the total lipid stoichiometry can be determined from the detected mass subtracted from the mass of two MSPs. However, it would be impossible to evaluate the fraction of each lipid by the mass alone. Taking into account that the mass difference of POPG and POPS relative to POPC is < 11 Da, it would require a higher resolution than is currently available to determine the stoichiometries of individual lipids.

Although the lipid composition could not be obtained directly from the mass, the distance between peaks with adjacent lipid counts illuminates an alternative (**Figure 2.4** A). As nanodisc self-assembly is a stochastic process, it could be assumed that the probability of lipids being incorporated into nanodiscs is equal. Thus, for the heterogeneous-lipid nanodiscs, the compositional discrete sub-states of a given lipid count are predicted to populate according to a binomial distribution (**Figure 2.3** B) and their average lipid composition is identical to that of the entire system. Consequently, peaks corresponding to nanodiscs with different total lipid counts have the same average lipid composition and

resulting average lipid mass (m_l). The average lipid mass is equivalent to the distance between adjacent peaks (e.g., between complexes with 123 and 124 lipids) in the zero-charge mass spectra (**Figure 2.4** A), which is a measurable parameter. If the average lipid mass is provided, the lipid composition can be calculated using the masses of the lipid constituents.

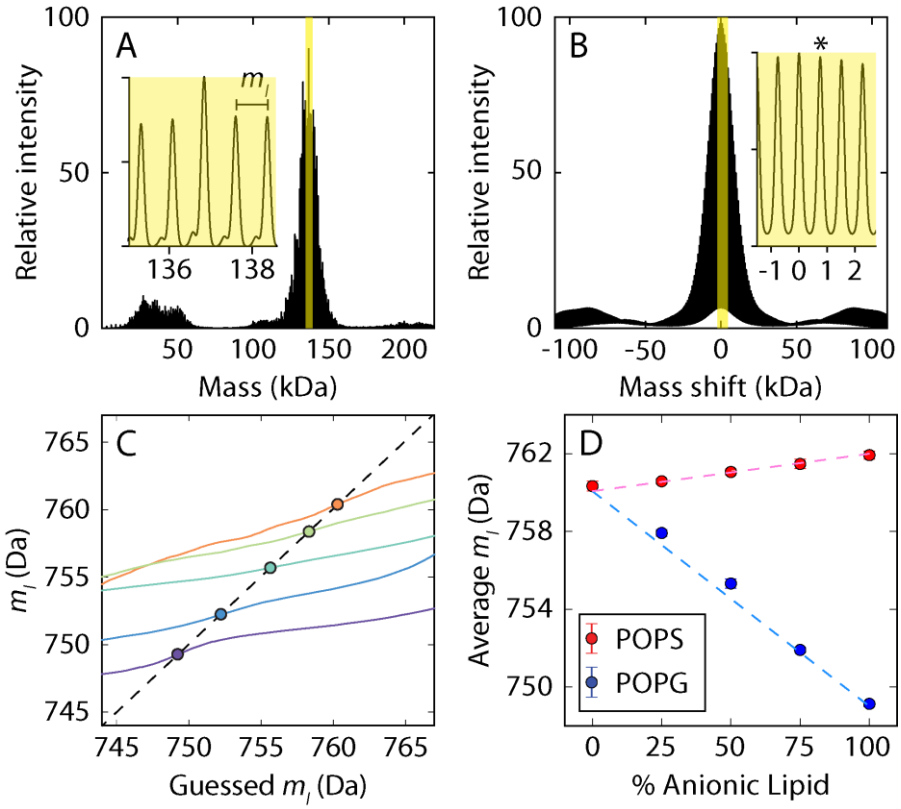

Figure 2.4 Measuring the average lipid mass (m_l) of nanodiscs by autocorrelation. (A) The zero-charge mass spectrum (mass distribution) was deconvolved from the mass spectrum of the POPC nanodisc acquired at 80 V CID by UniDec. The inset expands the highlighted region that shows the regularly spaced peaks that differ from each other by m_l. (B) Autocorrelation of (A) at 0.1 Da steps yields the autocorrelation spectrum. The inset denotes the highlighted region in the spectrum. The m_l is determined as the maximum (*asterisk*) between 740 and 770 Da. (C) Deconvolutions require a "guess" of the average lipid mass. The measured m_l is plotted as a function of the "guess" of the m_l used in deconvolution of the mass spectrum. The experimental

average lipid mass is determined by the coordinate (*dot*) intercept to the $x = y$ function (*dashed*). Each solid line corresponds to a POPC nanodisc with a different POPG content: *orange* to *purple* denotes a 0 to 100% POPG content. (D) The measured m_l of POPC-POPG (*blue*) and POPC-POPS (*red*) nanodiscs averaged from the mass spectra acquired at 60 to 100 V over three replicates is compared to the expected average lipid mass based on the lipid input in nanodisc reconstitution (*dashed*). Figure drawn by the author.

Two approaches were tested to measure the m_l. The initial approach directly measured the mass difference between peaks in the zero-charge mass spectra. However, the mass difference measured between adjacent peaks could vary by a couple of daltons. Thus, this approach was abandoned due to lack of precision. Subsequently, a more sophisticated autocorrelation approach was applied (The author suggested the correlation between the lipid composition and the peak distances in spectra/mass distributions; Dr. M. Marty suggested the approach and created a tool in UniDec; the author carried the analysis), taking advantage of the regularly repeating pattern created by the distribution of the lipid count per complex (**Figure 2.4** A). Mathematically, autocorrelation is commonly used to find the fundamental frequency of periodic signals. It is therefore useful to estimate the m_l that defines the frequency of the repeating patterns in the zero-charge mass spectra. Autocorrelation returns the correlation of a signal and a lagged copy of itself against the time lag (see **Equation 2.1**, **Figure 2.4** B), where time lag here represents the mass shift of the lagged copy. In the autocorrelation spectrum, peaks appeared at the lags when strong correlation was observed, which indicated the spacing between peaks in the zero-charge mass spectra. Because the lightest and heaviest lipids used are POPG (749 Da) and POPS (762 Da), respectively, the m_l was

determined by the maximum between 740 to 770 Da in the autocorrelation spectra (**Figure 2.4** B *asterisk*). Using autocorrelation, spacing between peaks can be measured with better precision (see below).

A prerequisite for using the autocorrelation method is to achieve an optimal deconvolution. Since autocorrelation was applied on the deconvolved zero-charge mass spectra, the measurement of the m_l could be sensitive to the deconvolution. An important parameter in deconvolution is the initial "guess" of the mass of the lipid inside the possible mass range (749 Da to 762 Da)[239]. Typically, the influence of this parameter to deconvolutions is minor, attributing to the robustness of UniDec[239]. Since previous studies measured the stoichiometry of lipids in nanodiscs, adequate mass accuracy can be achieved by deconvolutions bearing a "guess" with an error well above 10 Da. However, the difference between the average lipid masses of different nanodiscs can be as small as 0.5 Da. Higher mass accuracy in zero-charge mass spectra is required for the measurement of the m_l. An inaccurate "guess" may dramatically change the measured lipid composition. Consequently, it is necessary to find an appropriate "guess".

The appropriate "guess" of the lipid mass is defined as the "guess" that will lead to the actual m_l of the same value by autocorrelation. This definition is reasonable since the actual average lipid mass is expected to give a correct deconvolution that will return the actual average lipid mass by autocorrelation. It means that the appropriate "guess" and the resulting m_l are mutually identical. Finding such a "guess" provides the m_l and vice versa. To obtain the actual m_l,

deconvolutions were performed with “guesses” ranging from 744 to 767 Da at 0.1 Da intervals. The autocorrelation results of the deconvolutions were plotted as a function of the “guess” of the lipid mass (**Figure 2.4** C). The actual m_l was determined from the intersection with a function $x = y$. The intersection has a coordinate representing the linear trend when the “guess” and m_l have an equal value.

The autocorrelation approach was applied to measure the average lipid mass of the POPC-POPG and POPC-POPS nanodiscs. The m_l was determined by averaging the results from mass spectra acquired from 60 to 100 V CID over three independent replicates (**Figure 2.4** D). The results were compared to the expected average lipid mass based on the lipid composition used in nanodisc reconstitutions. Although some m_l values were significantly different from the expected values, the differences were less than 0.8 Da. The standard deviation was less than 0.3 Da. Therefore, the autocorrelation method could probe average lipid masses with better than 1 Da error.

2.3.3. Determining the average lipid mass of nanodiscs by dual Fourier transform

A major disadvantage of the autocorrelation approach is the need for deconvolution. Although it is possible to minimize the effect, the autocorrelation method is inevitably sensitive to the deconvolution parameters. Equally important, the autocorrelation method requires a significant amount of computational cost. The resolution of the measurement is limited by the resolution of the deconvolved zero-charge mass spectra. The bin size of the

spectra was set to 0.1 Da to maintain enough resolution given that the nanodiscs tested had similar average lipid masses. Increasing the resolution however significantly extended the deconvolution period. Further, it requires exhaustive deconvolution over the range of expected lipid mass to determine the actual value of m_l (**Figure 2.4** C). Each spectrum was deconvolved over 200 times in the study. Mass spectra were acquired over a range of collision energies for three independent replicates of nanodiscs with nine different lipid compositions. Over ten thousand deconvolutions and autocorrelations were performed on the entire dataset.

Therefore, this study explored other possibilities to measure the average lipid mass (m_l) directly from the raw mass spectra (**Figure 2.5** A; Dr. Marty suggested the approach and created a tool in UniDec; the author carried the analysis). Since nESI commonly generates ions with multiple charge states (z, z + 1, z + 2, etc.), mass spectra are fundamentally the sum of m/z signals of all available charge states. Nanodiscs show a distribution of the number of lipids per complex. For a given charge state z, the distribution of the lipid counts produces an evenly spaced series of peaks that each separated by m_l/z. Because the peak spacing is periodic, it was reasoned that Fourier transforms can be of use to extract the frequency from the spectra. The Fourier transform of the charge state z alone yields a primary frequency at z/m_l in the reciprocal space v, where the units are Th^{-1} (**Figure 2.5** B). As nanodisc ions exist in multiple consecutive charge states (z, z + 1, z + 2, etc.), the Fourier transform of the mass spectrum detects frequencies of (z/m_l, (z + 1)/m_l, (z + 2)/m_l, etc.) that are 1/ m_l apart from each other.

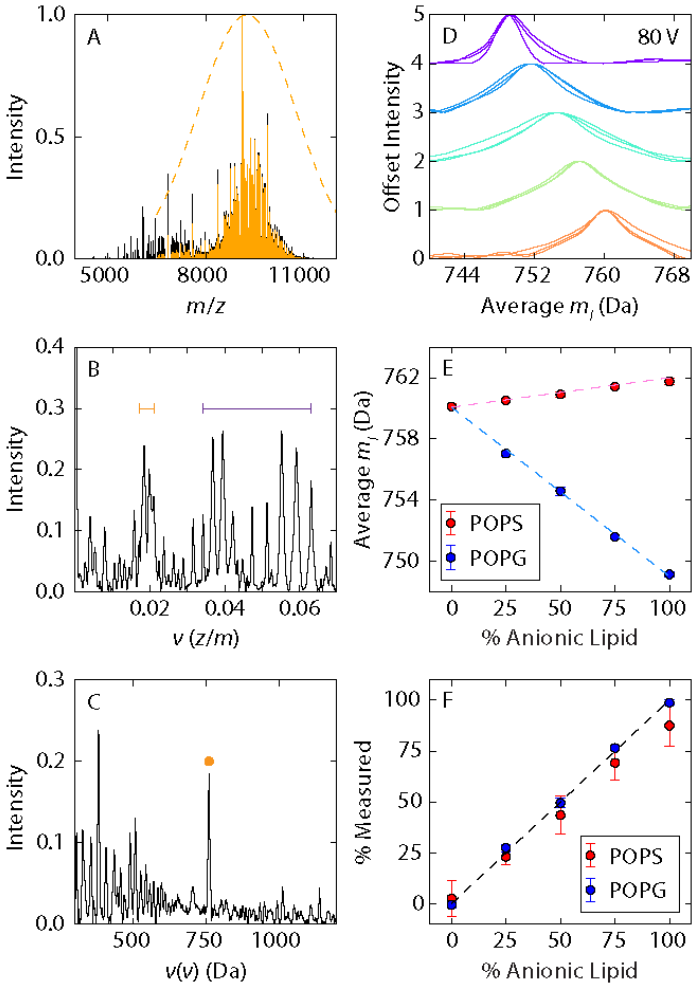

Figure 2.5 Measuring the average lipid masses (m_l) of nanodiscs by dual Fourier transform. (A) The raw spectrum of the POPC nanodisc acquired at 80 V CID (*black*) is multiplied with a Gaussian window (*orange dashed*) to generate a windowed spectrum (*orange*). (B) The Fourier transform of the windowed spectrum yields the primary frequency peaks (*orange bar*) for nanodiscs with different charge states and the harmonic peaks (*purple bar*) in the *v* space. (C) The second Fourier transform detected a frequency peak at m_l (*orange dot*) in the *v*(*v*) space. (D) As the POPG content in the nanodisc increases from 0 to 100% (*orange* to *purple*), the peak in the *v*(*v*) space shifts to a lower value. (E) The measured m_l of POPC-POPG (*blue*) and POPC-POPS (*red*) nanodiscs averaged from the mass spectra acquired at 60 to 100 V over three replicates agree within one standard deviation (*error bars*) of the theoretical m_l (*dashed*). (F) The corresponding average lipid composition is interpolated from the measured m_l in (E). Figure from Hoi *et al.*[1] and drawn by the author.

As another repeating pattern spaced with $1/m_l$ appears in the v space, a second Fourier transform can be performed. The second Fourier transform records a primary frequency at m_l in the reciprocal space $v(v)$ (**Figure 2.5** C). Because the charge in the $v(v)$ space becomes unity, the $v(v)$ space describes masses instead of m/z and has a unit of Da. Hence, the average lipid mass can be measured directly from the mass spectrum by performing two sequential Fourier transforms. In fact, another study has exploited Fourier transform to extract the lipid mass and charge state information from the spectra of homogeneous-lipid nanodiscs. This study demonstrated and supported the application of Fourier transform in native MS of nanodiscs[252].

The dual Fourier transform approach assumes a uniform m_l across the whole mass spectrum. However, the m_l of each charge state may have subtle differences, in which case the measured m_l will be the average of all charge states instead. On the other hand, this assumption is true on a local scale corresponding to the same ion species. The mass spectra contained nanodisc ions, which may be depleted of lipids by dissociation, dissociated MSP, and free lipids that could have different lipid compositions. To analyze each of these species separately, spectra were convolved with a sliding Gaussian window prior to the dual Fourier transform analysis (**Figure 2.5** A). The resulting $v(v)$ spaces were plotted as a function of the center of the Gaussian window (**Figure 2.6** B to D, E to G), allowing estimation of the m_l in the local m/z region. The overall average lipid mass in a mass spectrum reported the average lipid mass of the dominating signals in the spectrum. It was defined as the m_l measured at

the estimated centroid of the spectrum, which was the maximum position of the spectrum smoothed with a broad Gaussian (σ = 1,000 Th).

The dual Fourier transform approach enables determination of the average lipid mass directly from the raw spectra with minimal processing and assumptions. It was applied to the spectra of POPC-POPG and POPC-POPS nanodiscs acquired over a range of collision energies (**Figure 2.6** A). Because the theoretical average lipid mass of the nanodiscs is between 749 to 762 Da, the m_l was evaluated as the maximum between 740 and 770 Da in the second Fourier transform (**Figure 2.5** D). The $v(v)$ space had the same sample density as the original spectrum, which was linearized at 0.01 Th, hence it was possible to discriminate m_l with subtle differences (**Figure 2.5** D). By comparing the $v(v)$ spaces, nanodiscs with an equal fraction of POPC and POPG exhibited the largest peak width, which is likely due to a broader binomial distribution of compositional sub-states. The average lipid mass of a nanodisc was determined by the average values from spectra of 60 to 100 V CID, which had a relatively stable value of m_l (**Figure 2.6** A). Comparing the results among three independent replicates, the average lipid mass measured was very precise. The measurements had a standard deviation below 0.3 Da and generally contain the theoretical values (**Figure 2.5** E).

The results of the autocorrelation approach and the dual Fourier transform were compared using linear regression. The R-squared of both approaches were better than 0.99, indicating both set of results were linear. However, the slope and intercept of the dual Fourier transform were slightly closer to the ideal

values (**Figure 2.7**), suggesting the dual Fourier transform is a more appropriate approach to measure the m_l. In addition, the dual Fourier transform approach avoids deconvolution and may be applied directly to the raw data. While skipping the exhaustive deconvolution and requiring lower computational cost, the results obtained from the dual Fourier transform approach agreed more with the theoretical average lipid masses. Therefore, the dual Fourier transform provides a rapid, accurate and precise method to measure the average lipid mass in nanodiscs.

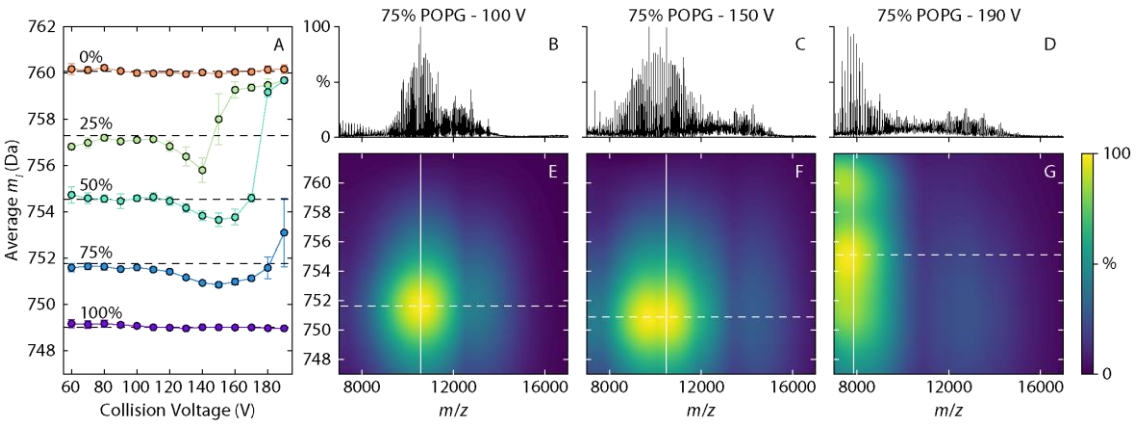

Figure 2.6 The average lipid mass (m_l) measured from nanodisc spectra at different CID states. (A) The m_l of nanodiscs with POPG:POPC compositions ranging from 0 to 100% (*orange* to *purple*) is shown as the mean and standard deviation of three replicate samples at 60 to 190 V CID. The measured average lipid masses changed as a function of collision voltages (see **Section 2.3.4** for further discussion). Dashed lines denote the theoretical m_l for nanodiscs at each composition. Representative spectra of 75% POPG nanodiscs acquired at (B) 100 V, (C) 150 V, and (D) 190 V CID are shown. (E), (F), and (G) are the corresponding plots showing the average lipid mass determined from the dual Fourier transform (ordinate) as a function of the Gaussian window position (abscissa). The intensity is normalized by the intensity of the smoothed mass spectrum used to determine the spectral centroid. The average m_l is measured as the maximum within 740 to 770 Da in the $v(v)$ space (*horizontal dashed line*) when the center

of the Gaussian window aligns with the estimated centroid (*vertical lines*). Figure from Hoi *et al.*[1] and drawn by the author.

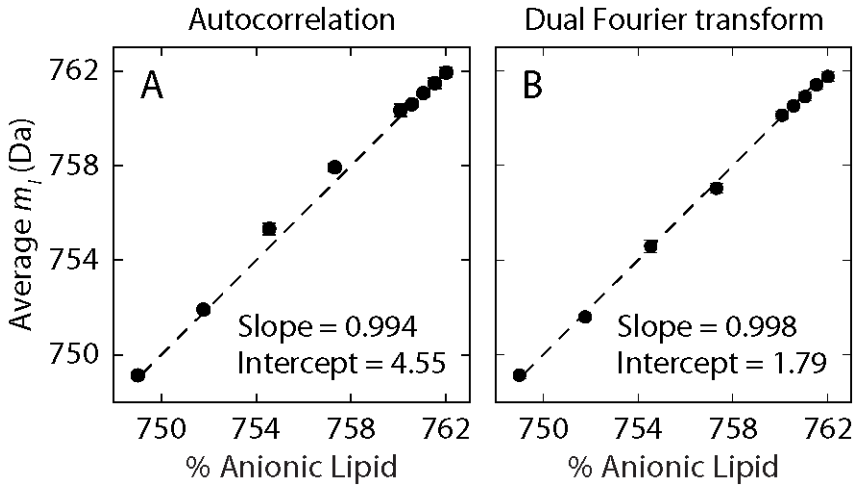

Figure 2.7 Linear regression is performed on the average lipid mass (m_l) measured for nanodiscs with different lipid compositions using (A) autocorrelation and (B) dual Fourier transform. The *dashed lines* denote the fit function. The slope and intercept are provided in the figure. Figure drawn by the author.

Proceeding with the m_l measured by the dual Fourier transform approach, it is now possible to determine the relative compositions of the two lipids in the nanodiscs. The average lipid mass correlates to the relative lipid composition and the masses of the two lipids. Because the masses of the lipids are known values, the lipid composition in the nanodisc can be interpolated from the average lipid masses (**Figure 2.5** F). The mass difference between POPC and POPG is 11 Da, which is significantly larger than the standard deviation of the measured m_l. Thus, the relative composition can be interpolated precisely with a relative standard deviation less than 3%. The calculated lipid compositions agreed closely with the lipid content input in nanodisc reconstitutions. This consistency reveals the faithful reproducibility of the lipid composition before

and after nanodisc self-assembly, illustrating the ability to tailor the lipid bilayer in nanodiscs[232].

In contrast, the compositions interpolated for the POPC-POPS nanodiscs were less promising despite having a precise measurement of m_l. Because the mass difference between POPC and POPS is small (2 Da), the standard deviation of m_l (< 0.3 Da) was converted into an error of ~15% in the relative composition. Therefore, to precisely determine the relative composition of POPC-POPS nanodiscs, it is necessary to improve the resolution in the analysis or increase the mass difference. In the latter case, isotopic labeling could be of use.

Combining the dual Fourier transform approach and UniDec, spectra of heterogeneous-lipid nanodiscs can be assigned unambiguously. The Fourier transform approach is capable of determining the average lipid masses and average lipid compositions in nanodiscs. Instead of a "guess", the measured average lipid masses can be used to optimize the deconvolution. Deconvolution of the mass spectra provides the charge states and masses of the detected ions (**Figure 2.2**), from which the stoichiometry of the MSP and total lipid can be determined. Together with the relative lipid composition, it is possible to assign the stoichiometry of each lipid component in the nanodisc complexes.

Although the framework provided here was demonstrated on heterogeneous-lipid nanodiscs containing two lipid species, it could be applied to systems with more than two lipids by combining an isotopic labeling strategy. For example, if

the measured average lipid mass of an unlabeled nanodisc is m_l, labeling a lipid component with heavy isotopes will lead to an increase in the m_l. The dual Fourier transform approach can determine the average lipid masses of the unlabeled and labeled nanodiscs as well as the increase. This increase is directly proportional to the relative fraction of the labeled lipid and thus allows its calculation. In such a way, the relative fraction of every lipid species in the system can be determined. Therefore, the framework introduced here is potentially applicable to more complex systems with multiple lipid species.

This method may facilitate analysis of lipid composition around different membrane proteins. For example, the calcium-ATPase was reconstituted in lipid bilayer containing a mixture of phosphatidylcholine (PC) and phosphatidylethanolamine (PE) at different ratios[91]. The ATPase activity was halved when the relative fraction of PC was under 20%. This was suggested due to the changes of the lipid composition in the lipid annulus, which is important to the enzymatic activity of the ATPase[80,89,92]. As demonstrated in **Chapter 3**, native MS can probe the subset of lipids making up the annulus. The dual Fourier transform approach may allow the determination of the lipid fraction of the annular lipids, which is important for understanding the modulation of protein-lipid interactions to membrane protein function.

2.3.4. The dissociation mechanisms of nanodiscs in the gas-phase

Returning to **Figure 2.6** A, the m_l measured was not uniform over the range of collision voltages applied. While surveying the m_l of POPC-POPG nanodiscs against collisional energies, the m_l remained stabilized at low collision energies

(60 and 100 V). However, the measured m_l showed a decrease followed by an increase at elevated collision energies. A closer inspection revealed that POPC was preferentially ejected at moderate energy, which led to an enrichment of POPG and a drop in the average lipid mass in the lipid-depleted nanodiscs (**Figure 2.6** E and F). In contrast, the increase of m_l at high energies was not due to an enrichment of POPC in the nanodiscs. Because nanodiscs were largely dissociated at high energies, mass spectra were dominated by dissociated MSP leading to a centroid at lower *m*/*z* (**Figure 2.6** G). Given that the released MSPs carried a small number of lipids, with a preference of POPC, the dual Fourier transform approach reported the higher m_l values of the MSP-lipid complexes as the nanodisc signals were overtaken.

The observed preference of the lipid ejection suggested that dissociation of nanodiscs in the gas-phase is a lipid-dependent process. To further investigate this relationship, the dissociation patterns of the nanodiscs were studied by monitoring their charge distributions and mass distributions as a function of collision voltage (**Figure 2.2** E and F). A previous study has shown that lipids in DMPC nanodiscs depleted gradually upon collisional activation[231]. A similar phenomenon was observed in all the homogeneous- and heterogeneous-lipid nanodiscs containing POPC, as indicated by the shift in their mass distributions (**Figure 2.8** *left*) and charge distribution (**Figure 2.9** *left*). Interestingly, POPG and POPS nanodiscs exhibited a prolonged stability against collisional activation and underwent a discontinuous dissociation to MSP monomers.

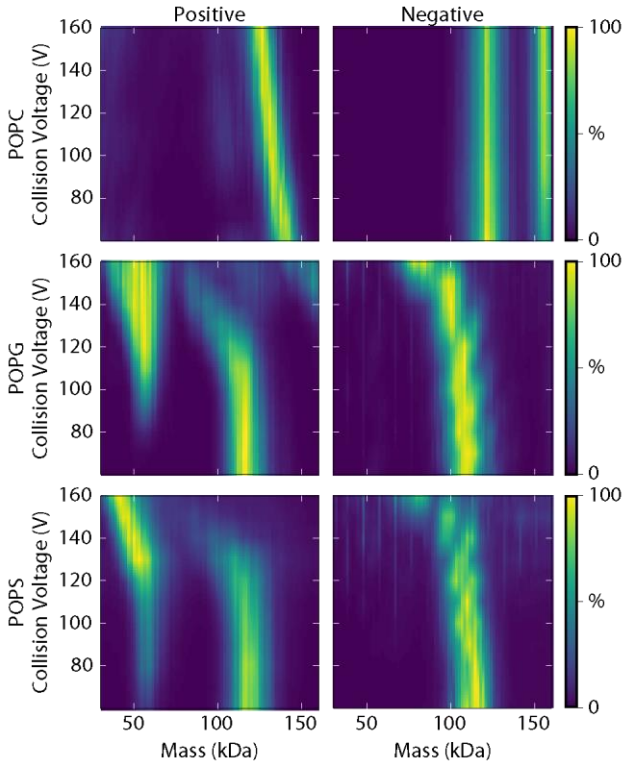

Figure 2.8 The total mass distribution of homogeneous-lipid nanodiscs. The mass distributions of the homogeneous-lipid nanodiscs are shown as a function of collision voltages from 60 to 160 V CID for POPC (*top*), POPG (*middle*), and POPS (*bottom*) nanodiscs in positive (*left*) and negative (*right*) polarity. Figure from Hoi *et al.*[1] and drawn by the author.

Since CID is often accompanied with charge reduction, the unusual stability shown by POPG and POPS nanodiscs under positive polarity may relate to the unchanged charge state of the complexes (**Figure 2.9** *left*). If nanodisc dissociation is charge-directed or involves dissociated lipids carrying charges, a dependency on instrumental polarity is expected. Switching to negative polarity, POPC nanodiscs exhibited an extraordinary stability similar to that observed for POPG and POPS nanodiscs under positive polarity (**Figure 2.8** and **Figure 2.9**). The masses and charge states of the POPC nanodiscs were constant at all energy states. However, continuous dissociation was observed for the POPG and POPS nanodiscs as well as the heterogeneous-lipid nanodiscs (**Figure 2.8**

and **Figure 2.9**). Both lipid and charge were lost gradually. Thus, nanodisc dissociations are sensitive to the lipid composition and instrumental polarity. The charge of the lipids may play an important role in the dissociation pathway.

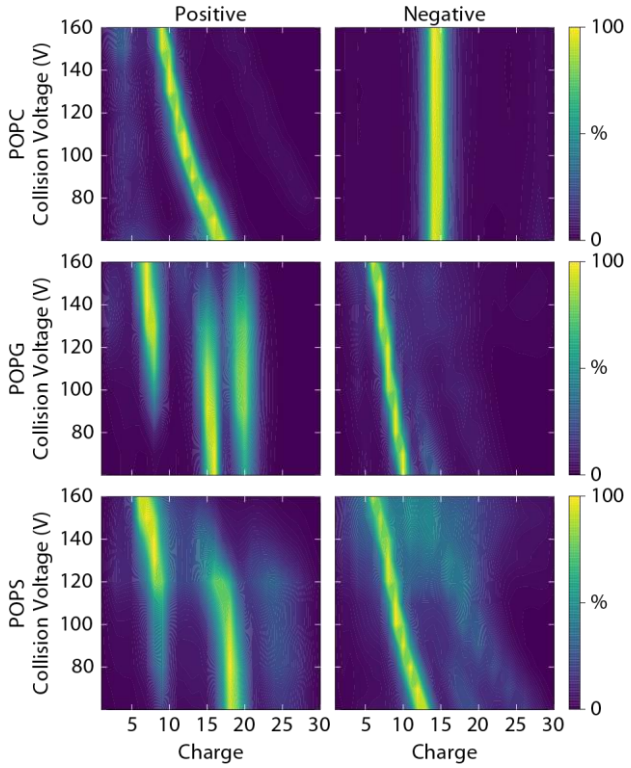

Figure 2.9 The total charge distribution of homogeneous-lipid nanodiscs. The charge distributions of homogeneous-lipid nanodiscs are shown as a function of collision voltages from 60 to 160 V CID for POPC (*top*), POPG (*middle*), and POPS (*bottom*) nanodiscs in positive (*left*) and negative (*right*) polarity. Figure from Hoi *et al.*[1] and drawn by the author.

To understand their charge properties, lipids were classified based on their pK_a values in solution lipid bilayers as well as on their ESI behavior from lipidomics experiments[253-256]. POPC, POPG ,and POPS share the same hydrocarbon tail but different head groups as shown in **Figure 2.3**. Due to the permanent positive charge on the quaternary amine, POPC can hold either a 1+ charge or a zero net charge, but not a negative charge. Thus, in the absence of adducts,

POPC can be charged under positive polarity but not negative polarity, where the POPC nanodisc displayed unusual stability.

Conversely, POPG can hold either a neutral or 1− charge. It cannot hold a positive charge and ionizes only in the negative mode. The possible charge states of POPS are between 1+ and 2− charges. Although POPS can be positively charged, this charge state is observed only at very low pH. Hence, both lipids ionize preferentially under negative polarity but stay neutral under positive polarity, where homogeneous-lipid nanodiscs of these lipids were resistant to continuous dissociation. Therefore, according to the observations in the positive and negative polarities, homogeneous-lipid nanodiscs display a unique stability under the instrumental polarity where the withholding lipids are neutral (**Figure 2.8** and **Figure 2.9**). Nanodisc dissociation is a polarity dependent process that is controlled by the chemistry of the lipid head groups.

However, nanodisc dissociation events are not exclusively controlled by charge. Continuous dissociation events were observed in the heterogeneous-lipid nanodiscs in both polarities. Although some preference was observed at high collision energies, dissociation was not initially selective for the charged lipids (**Figure 2.6** A). In fact, the ejection was unbiased toward charged and neutral lipids at low collision voltages. In addition, lipids were also ejected from homogeneous-lipid nanodiscs in both charged and neutral states under appropriate polarity (positive polarity for POPC nanodiscs; and negative polarity for POPG and POPS nanodiscs). This phenomenon was evident because the mass accounted for the charged lipids ejected met only part of the total mass

loss. Thus, charged lipids are required for continuous dissociation, but not all lipids that dissociate are charged.

2.4. Conclusion

This study demonstrates a framework for analysis of nanodiscs with a heterogeneous population of lipids by combining different strategies: use of lipids with similar masses to manipulate the resolution of mass spectra; application of Bayesian deconvolution to interpret mass spectra; and measuring the lipid composition of the nanodiscs using the average lipid mass as a reporter. Two methods, autocorrelation and dual Fourier transform, were applied to measure the average lipid mass. Although both methods provide precise and linear measurements, the dual Fourier transform method has a higher speed and accuracy but requires minimal processing and assumptions. Comparing the interpolated lipid composition with the lipid composition used in nanodisc reconstitution indicated that nanodiscs can be designed with a precise lipid composition. This framework developed here sets the stage for native MS analysis of membrane proteins in nanodiscs with a complex lipid environment. In addition, CID experiments revealed the effects of the lipid charge on nanodisc dissociation, which showed that having a chargeable lipid is required for continuous dissociation, but that not all dissociated lipids are charged. These results provide a unique perspective and highlight the importance of the chemistry of the lipid and the polarity of the instrument on the dissociation mechanisms of protein-lipid ions in the gas-phase.

Chapter 3. Gas-phase liberation of membrane proteins from nanodiscs reveals the bound lipid annulus

3.1. Introduction

Membrane proteins govern the interactions and communications between the intracellular and extracellular environments. Many of these interactions are essential cellular activities, such as energy conversion and substance transportation[257]. Abnormalities in membrane protein function correlate to pathogenesis of many diseases[111,258]. Increasing evidence has suggested that the biophysics and biochemistry of membrane proteins is influenced by the interactions with lipids in the membrane environment[81,145]. Lipids can interact with, and affect, membrane proteins' conformation, oligomeric state, dynamic, stability, functionality and ligand binding[54,61,63,73]. The details of protein-lipid interactions inform research on the function and regulation of membrane proteins. However, it is challenging to study membrane protein-lipid interactions directly in complex biological membranes found in cellular systems. Membrane proteins are intrinsically unstable outside of the native membrane environment[4,23,24]. To isolate membrane proteins from their native environment, it is necessary to maintain their hydrophobic interactions and structures.

Nanodiscs contain a lipid bilayer encircled by two molecules of membrane scaffold proteins (MSPs), which can accommodate the encapsulated membrane proteins in a native-like lipid bilayer environment, in contrast to the micelle environment provided by detergents[241,259]. The advent of nanodiscs provides a platform to characterize membrane proteins in the context of lipid bilayers. For example, the adrenergic receptor β_2AR was explored in several homogenous-

phospholipid environments using the nanodisc technology[54]. The β_2AR in different lipid environments exhibited different ligand affinity and TM6 movement, which is an indication of receptor activation. These results suggested that the lipid environment can modulate the activity of G protein-coupled receptors. In another example, the heat- and capsaicin-activated ion channel, TRPV1, reconstituted in nanodiscs were studied using cryogenic electron microscopy (cryo-EM)[90]. The resolved structures provided important molecular insights of the inhibition of phosphatidylinositides (PIPs) and the positive regulatory effect of annular lipid interactions. Many studies using nanodiscs have increased our knowledge of the effects of lipids on membrane protein function[233,234]. Nevertheless, because a limited number of techniques are capable of directly probing protein-lipid interactions, little is known about how membrane proteins interact with lipids.

This chapter couples mass spectrometry (MS) with nanodisc technology to interrogate the protein-lipid interactions of two model membrane proteins: the ammonium channel, AmtB, and aquaporin Z, AqpZ. Mass spectra showed that membrane proteins released from nanodiscs in the gas-phase can retain a large number of bound lipids. By integrating a previously established Bayesian deconvolution algorithm for mass spectra with overlapping charge states, a high-resolution Orbitrap mass spectrometer, and a novel mass defect analysis, it is possible to count the number of lipids bound to a membrane protein ejected from nanodiscs. Molecular dynamics (MD) simulations indicate that the distributions of bound lipids observed in MS correspond to the distinct populations of annular lipids featured by different types of protein-lipid

interactions. These results demonstrate the potential of MS in combination with nanodisc technology to define the cohort of lipids that are in direct interactions with membrane proteins.

3.2. Methods

3.2.1. Sample preparation and characterization

Methods for protein concentration, concentration determination, filtration and centrifugal buffer exchange are in the **Section 2.2.1**. All results were obtained from three biological replicates to establish reproducibility.

Membrane protein expression and purification: Expression plasmids of AmtB and AqpZ, which were generated by cloning AmtB or AqpZ gene into modified pET15b vectors, were provided by Dr. Timothy Allison (see reference[63] for detail). Briefly, the plasmids were transformed into *E. coli* BL2-Gold (DE3) (Bio Labs New England) for protein expression. AmtB expressed from the construct contains a N-terminal fusion of 10× His-tag followed by a maltose binding protein (MBP) and a TEV cleavage site. AqpZ was expressed with a C-terminal fusion (green fluorescent protein (GFP) and 6× His-tag) jointed by a TEV recognition site.

AmtB and AqpZ were prepared following previously described methods[226,241]. AmtB and AqpZ were overexpressed in *E. coli* grown in Luria Broth (LB) medium supplemented with 100 µg/mL ampicillin. At OD_{600} of 0.6, protein expression was induced by adding 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG; 1 M IPTG stocks). After 3 h of protein expression, cells were harvested by centrifugation (5,000 ×*g*, 10 min), re-suspended in buffer (20 mM Tris pH 7.4, 150 mM NaCl; 25 mL per liter of cell culture) and lysed using a microfluidizer M-110P (Microfluidics Corp.). The cell debris in the lysate was removed by centrifugation (20,000 ×*g* for 20 min). Supernatant containing the membranes

was collected. Cell membranes were pelleted by ultracentrifugation (27,000 ×*g* for 2 hours). Membrane proteins were extracted from the membrane by homogenization (Potter-Elvehjem Teflon pestle and glass tube) using 1% (w/v) OG dissolved in loading buffer (50 mM Tris, 200 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, 0.025% (w/v) DDM, and 5 mM 2-mercaptoethanol; 10 mL per gram of membrane) and rotated overnight. Sample solution was loaded on 5 mL His-Trap HP column equilibrated with loading buffer, washed (250 mL), and eluted using loading buffer with 0.5 M imidazole. Eluent was pooled, concentrated, filtered, and injected on a Superdex 200 16/600 PG column equilibrated with loading buffer to isolate the protein oligomers (**Figure 3.1**). Peak fractions were collected and combined. Following concentration to ~50 µM, purified membrane proteins were snap frozen and stored at −80 °C.

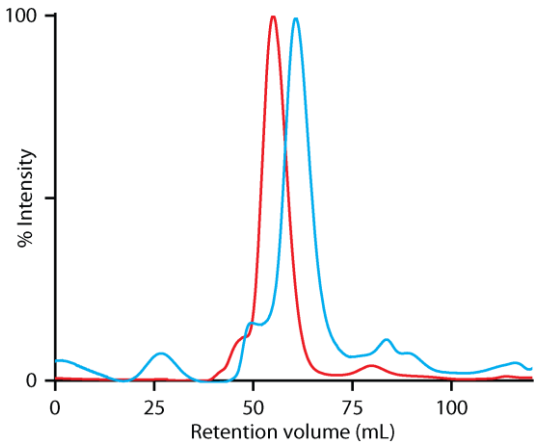

Figure 3.1epresentative SEC profiles (Superdex 200 16/600 PG column) of AmtB (*red*) and AqpZ (*blue*). Figure drawn by the author.

Nanodisc reconstitution and purification: Membrane proteins were reconstituted in POPC-MSP1D1(−), DMPC-MSP1D1(−), POPC-MSP1E3D1(−), and POPG-MSP1D1(−) nanodiscs based on previously described

procedures[2,241]. Purified membrane proteins were mixed with a specific combination of a MSP and a lipid (see below). Excess MSP and lipid were added at 10-fold molar equivalent of the membrane protein to enhance membrane protein incorporation into nanodiscs. In addition, membrane proteins displace lipids from empty nanodiscs and reduce the optimal lipid/MSP ratio of the complex. To keep the lipid/MSP ratio same as that of empty nanodiscs, an excess of empty nanodiscs was added to reduce the deviation of lipid/MSP ratio over the optimal value.

MSPs (400 μM stocks) and cholate-solubilized lipids (0.05 M stocks) were prepared as described in **Chapter 2**. For each preparation, 1 mL MSP was used. Lipids were added at a molar ratio of 1:55 for POPC-MSP1D1(−) and POPG-MSP1D1(−), 1:70 for DMPC-MSP1D1(−), and 1:120 for POPC-MSP1E3D1(−) nanodiscs. Membrane proteins (50 μM stocks) were added at 0.05-fold molar equivalent of the MSP. The final concentration of cholate was adjusted to 23 mM by adding cholate buffer (0.1 M sodium cholate, 20 mM Tris pH 7.4, 0.15 M NaCl) or buffer (20 mM Tris pH 7.4, 0.15 M NaCl). The reconstitution mixtures were incubated at 4 °C (for POPC and POPG) or room temperature (for DMPC) for 2 hours, followed by overnight incubation with 1:1 (v/v) Amberlite XAD-2 hydrophobic beads to remove the detergents and drive nanodisc self-assembly.

Following incubation, the supernatant was recovered by filtration and loaded onto a 1 mL His-Trap HP column in loading buffer (40 mM Tris pH 7.4, 0.3 M NaCl, and 20 mM imidazole) to remove nanodiscs without embedded

membrane proteins. Following washing (loading buffer, 10 mL), the membrane protein nanodiscs were eluted using loading buffer with 0.5 M imidazole. Eluent was pooled, concentrated, filtered, and injected on a Superose 6 Increase 10/300 SEC column in loading buffer (**Figure 3.2** A).

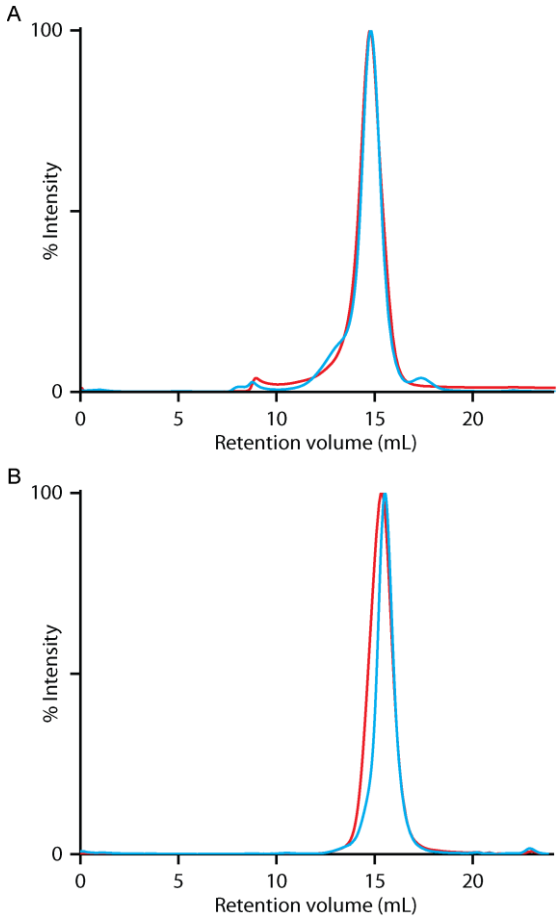

Figure 3.2epresentative SEC profiles (Superose 6 Increase 10/300 column) of protein nanodiscs. (A) The SEC chromatograms of AmtB with the fusion tag in POPC-MSP1D1(−) (*red*) and POPC-MSP1E3D1(−) (*blue*) nanodiscs. (B) The corresponding SEC chromatograms after removal of the tag. Removal of the tag is indicated by the slightly later retention volume. Figure drawn by the author.

Peak fractions were combined and incubated overnight with 1:100 (w/w) TEV protease with a histidine tag (Sigma) and 5 mM 2-mercaptoethanol to cleave

the MBP and GFP fusions on AmtB and AqpZ, respectively. Samples were purified using r-IMAC (1 mL His-Trap HP column in loading buffer) to remove the fusions. Flow through containing the nanodiscs was concentrated and injected on a Superose 6 increase 10/300 SEC column equilibrated with 0.2 M ammonium acetate pH 7.4 (**Figure 3.2** B). Peak fractions were pooled and concentrated to an absorbance ranging from 1 to 5 AU at 280 nm (A280). Purified nanodiscs can be directly used or snap frozen and stored at −80 °C before MS measurements. In some measurements where residual sodium adducts were observed, nanodiscs were buffer exchanged to 0.2 M ammonium acetate pH 7.4 using a Micro Bio-Spin^{TM} 6 column.

Characterization: Nanodiscs were analyzed by MS in positive polarity using the method described in **Chapter 2**. In addition to MS, purified nanodiscs were also characterized by electrophoresis (SDS-PAGE) to analyze their protein constituents and SEC (Superdex 200 Increase 3.2/200 column) to estimate their Stokes radius. The measured Stokes radii were calibrated using five different gel filtration standards: carbonic anhydrase (29 kDa), albumin (66 kDa), alcohol dehydrogenase (147 kDa), apoferritin (443 kDa), and thyroglobulin (669 kDa). Three independent replicates were prepared for each type of nanodiscs, and each was characterized and analyzed independently.

3.2.2. Data analysis

Deconvolution: Mass spectra of nanodiscs were deconvolved using UniDec as described in **Chapter 2**. Deconvolution considered possible charge states between 1 and 25 and possible masses below 220 kDa. No significant peaks

were observed outside the defined range. The peak shape used a split Gaussian/Lorentzian peak shape with a FWHM of 5 Th. The oligomer mass used was the theoretical mass of POPC or DMPC. Deconvolution used the charge state and *m*/*z* values to generate mass distributions, which were linearized at 10 Da intervals. Because the typical peak width of the high-mass region was around 100 Da, 10 Da intervals were adequate to accurately define the peak while saving unnecessary computational cost. Deconvolution was also performed using 1 Da intervals to obtain exact masses for mass defect analysis. Mass accuracy was judged by comparing the deconvolved mass to the theoretical mass of the MSP, where the mass difference was generally smaller than one bin size of 10 or 1 Da depending on the resolution defined in the mass distribution.

Mass defect analysis: Dr. Marty suggested the mass defect analysis and created a tool in UniDec; the author carried out the analysis. To assign the composition of the masses obtained from deconvolution, a mass defect analysis was developed. A mass defect is the remainder of a mass divided by a reference mass. Normalized mass defects are defined as:

$$\frac{M}{M_l} - Floor\left(\frac{M}{M_l}\right) \qquad \textbf{Equation 3.1}$$

where M is the measured mass and M_l is the mass of the reference lipid. *Floor* function returns the greatest integer that is smaller than, or equal to, the variable. Reference lipid masses used were 760.076 Da for POPC and 677.933 Da for DMPC, which are the average masses reported by Avanti Polar Lipids. For MSP1D1(−) (22,044 Da), the mass defect referred to POPC is 0.02 (22,044 / 760.076 = 29.002).

Mass defect analysis was performed by the built-in module of Kendrick mass tools in UniDec. The analysis used the average mass distribution (see **Figure 3.6** D and H), which was the sum of the deconvolved mass distributions of spectra acquired at HCD energies from 50 to 190 V. The normalized mass defects of every signal in the mass distribution were calculated. In the output, measured masses were binned to the nearest integer multiple of the lipid mass that is not larger than itself, which is defined as:

$$M_l \times Floor\left(\frac{M}{M_l}\right) \qquad \textbf{Equation 3.2}$$

The mass defect analysis returned the results as a heat map of mass defects versus measured masses, where the intensity values corresponded to the intensity in the mass distribution for each M (see **Figure 3.7**).**Average mass defect distributions:** Dr. Marty did this work (see Marty *et al.*[2] for detail). Briefly, to determine the average mass defect of the high-mass peaks, mass defect distributions for mass values greater than 90 kDa were averaged (see **Figure 3.9** *left*). The average mass defect distributions were fit to a Gaussian to determine the mean and standard deviation of the mass defect of the high-mass peaks (see **Table 3.1**).

Distribution of the number of bound lipids: Dr. Marty did this work (see Marty *et al.*[2] for detail). Briefly, average mass distributions were converted to average lipid distributions (**Figure 3.9** *right*). Average lipid distributions were fit with a three-component Gaussian model (determined by a T-test). The mean and standard deviation of the fit Gaussians are summarized in **Table 3.2**.

Characterization of protein-lipid interactions by MD: Dr. Marty did this work (see Marty *et al.*[2] for detail). Briefly, to rationalize the lipid distributions observed in MS, molecular dynamics of AmtB and AqpZ in POPC bilayer were performed. Three lipid populations are counted based on their interactions with protein: annular lipids, sub-population of annular lipids that interact with the protein through the head group, and sub-population of annular lipids interacting through ionic contacts. Mean and standard deviation of these populations were calculated (see **Table 3.2**).

3.3. Results and discussion

3.3.1. Preparation of membrane protein nanodiscs

Nanodiscs of membrane proteins were produced for MS analysis (the author acknowledges Dr. M. Marty for the training and supervision)[2,241]. The sample preparation here took multiple incubation and purification steps to optimize the sample purity and the polydispersity of the lipid stoichiometry per nanodisc. However, these lengthy steps could affect protein stability and sample recovery yield. Due to their high stability and expression level[260,261], the trimeric ammonium transporter, AmtB, and tetrameric aquaporin, AqpZ, were selected to prepare nanodiscs. Another advantage of using AmtB and AqpZ is that both proteins have been well characterized previously by MS and have known crystallographic structures[63,262,263]. This information is useful to comprehend the molecular detail of the membrane protein-lipid interactions. The two targeted proteins were expressed, purified and reconstituted into nanodiscs via established protocols[63,241].

AmtB and AqpZ were reconstituted into nanodiscs with either POPC or DMPC, MSP1D1(−) or MSP1E3D1(−). POPC (16:0-18:1) and DMPC (14:0) were selected as lipid candidates because they are frequently applied in nanodisc formation as these lipids form a stable lamellar lipid phase[233]. Additionally, composition assignment for mass spectral data was deduced by comparing the results obtained from the POPC and DMPC nanodiscs using the mass defect analysis (see detail **Section 3.3.4**). This analysis observes the mass defect of a detected complex, which is sensitive to the masses of its constituents (i.e., the membrane protein, lipid, and MSP). By changing either or both of the masses of

the MSP and lipid, the composition of the complex can be inferred by the change in the mass defect. Ideally, isotopic labeling may be used to change the mass of the constituent without altering the complex identity. This study however used two chemically similar lipids, POPC and DMPC, instead of isotopic labeling. Further, MSP1D1(−) and MSP1E3D1(−) form nanodiscs of different sizes[227]. They were used to investigate the effect of nanodisc size on the MS analysis, such as their dissociation behavior and their ability to preserve protein-lipid interactions. Moreover, switching to different MSPs also aids composition assignments since the mass defect also depends on the scaffold protein.

To ensure nanodiscs homogeneity and proper incorporation of the membrane proteins, reconstituted membrane protein nanodiscs were characterized by electrophoresis (sodium dodecyl sulfate-polyacrylamide gel electrophoresis, SDS-PAGE) and SEC. Electrophoresis revealed successful reconstitution of the targeted proteins into nanodiscs (**Figure 3.3)**. Additionally, MSP1D1(−) and MSP1E3D1(−) nanodisc preparations showed single peaks in their respective SEC chromatograms (**Figure 3.2** B), suggesting a monodisperse nanodisc population. The Stokes diameters of the nanodiscs were estimated by the retention times calibrated by five different gel filtration standards. The average Stokes diameter of MSP1E3D1(−) nanodiscs was 12.0 ± 0.1 nm, which agreed well with the 12.1 nm Stokes diameter of POPC-MSP1E3D1 nanodiscs previously reported[227]. The nanodiscs with MSP1D1(−), however, had a larger Stokes diameter of 12.2 ± 0.1 nm compared with the 9.7 nm of POPC-MSP1D1 nanodiscs in other literature[227].

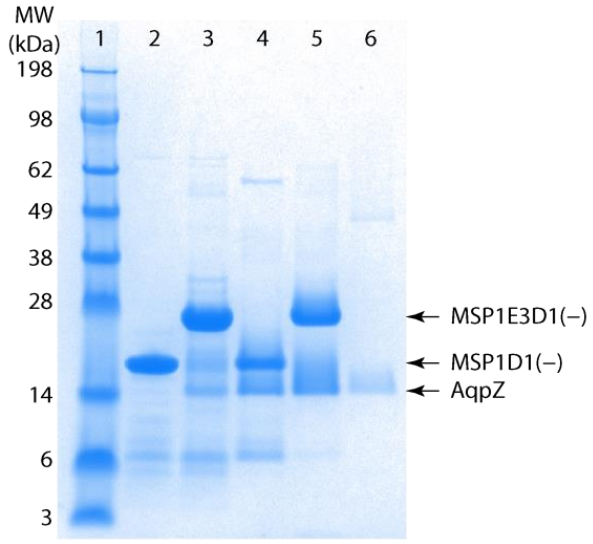

Figure 3.3 Representative SDS-PAGEs of membrane protein nanodiscs. Lanes correspond to (1) molecular weight marker, (2) MSP1D1(−), (3) MSP1E3D1(−), (4) AqpZ-POPC-MSP1D1(−) nanodisc, (5) AqpZ-POPC-MSP1E3D1(−) nanodisc, and (6) AqpZ. Figure drawn by the author.

The increased Stokes diameter is possibly due to formation of larger nanodiscs with more than two MSPs. As AmtB and AqpZ do not have large soluble domains, the embedded AmtB and AqpZ would have minor effect on the Stokes diameter of the nanodisc. Therefore, MSP1E3D1(−) nanodiscs had similar Stokes diameters reported in literature. In contrast, MSP1D1(−) forms smaller nanodiscs and may not be able to accommodate the AmtB and AqpZ oligomers. Larger nanodiscs with more than two MSP1D1(−) may be required to encapsulate the large protein oligomers, leading to larger Stokes diameters. Furthermore, nanodiscs of different sizes can be formed with one or two MSP2 depending on the MSP/lipid ratio[264]. Nanodiscs can also be formed with multiple 18 amino acid peptides[265]. Both of these examples implies that the stoichiometry of MSP1D1(−) may be different from two. Nevertheless, no intact nanodisc was observed in the spectra (see **Section 3.3.4**). The stoichiometry of

the MSP in intact nanodiscs remains unknown. Because all MSP molecules were dissociated from the nanodiscs as indicated by the mass defect analysis, the MSP stoichiometry did not affect the assignment of the membrane protein-lipid complexes observed in the spectra. More importantly, the unexpected stroke diameters of MSP1D1(−) nanodiscs did not affect the membrane protein-lipid interactions observed, as indicated by the similarity between the mass spectra of MSP1E3D1(−) and MSP1D1(−) nanodiscs (more detail below).

3.3.2. MS measurement of membrane protein nanodiscs

Reconstituted membrane protein nanodiscs were analyzed on an Orbitrap mass spectrometer optimized for the transmission and detection of high *m*/*z* ions (**Figure 1.15**)[209]. The nanodiscs were introduced into the gas-phase through nESI. Heat was applied at the ion transfer capillary to improve desolvation. To dissociate the nanodiscs and release the encapsulated membrane protein complexes, ions were activated by applying voltages in the HCD cell (50 to 190 V) prior to the detection in the Orbitrap mass analyzer.

The mass spectra of membrane protein nanodiscs were highly complex, containing a “forest” of peaks that span over a broad range of *m*/*z* (typically from 2,500 to 15,000 *m*/*z*). These peaks represented a heterogeneous collection of CID products, which included membrane protein oligomers, dissociated membrane protein, MSPs, and lipid clusters (no intact nanodiscs were observed; peak assignment is discussed in **Section 3.3.4**). All these products carried a number of lipid adducts, especially the membrane protein oligomers observed in the high *m*/*z* region (**Figure 3.4** A). These peaks

displayed a characteristic pattern of overlapping peaks resembling that observed for “empty” nanodiscs (nanodiscs devoid of membrane proteins; **Figure 3.4** B)[231], implying a dissociation product with a broad lipid distribution (i.e., the number of lipids per complex may range from a few to dozens). Thus, spectra of nanodiscs contained a large variety of complexes of different masses. Because these complexes existed in multiple charge states (**Figure 3.4** C), they presented hundreds of peaks in the spectra.

In spite of the high density of peaks, nanodisc spectra acquired in this study were at near-baseline resolution. This is attributable to the high resolution afforded by the Orbitrap mass spectrometer. If the spectra were acquired at inadequate resolution, many peaks would have overlapped with each other (**Figure 3.5**). Peak overlapping not only leads to unresolved spectra, lower signal-to-noise ratio, and loss of information, but also hinders charge state and mass assignment. Hence, the advent of the Orbitrap technology provides mass spectra with unprecedented quality and facilitates explicit assignments and interpretations, which was not achievable until recently[209].

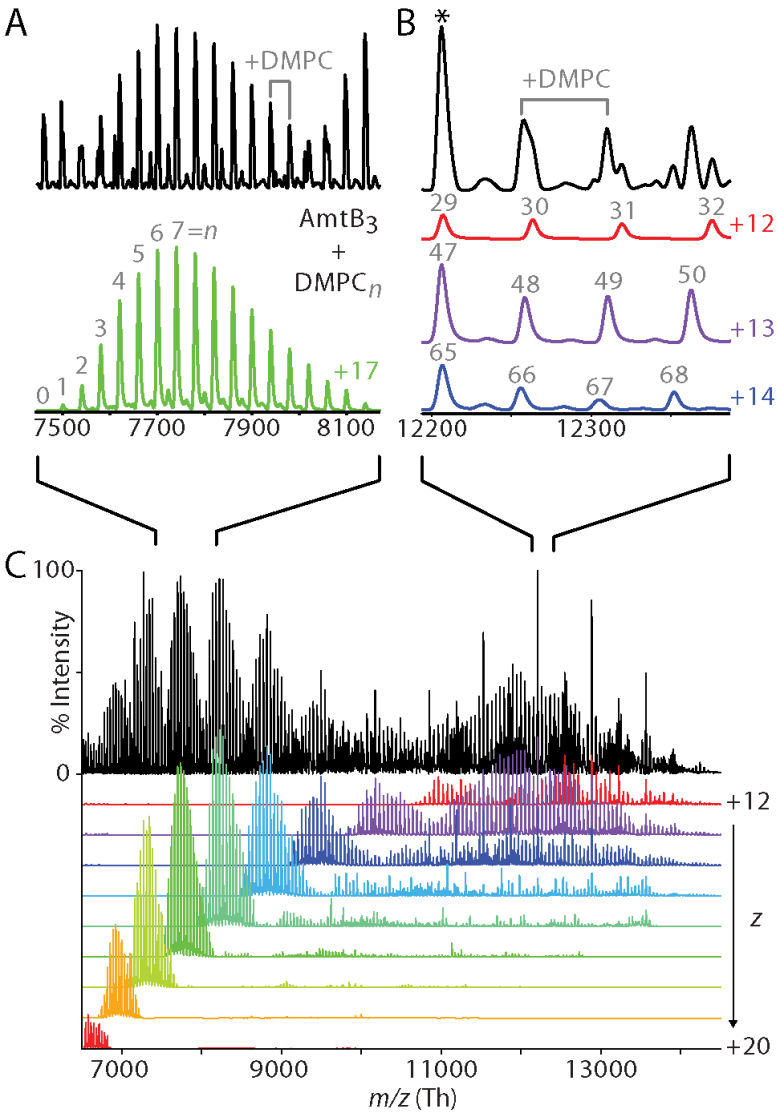

Figure 3.4 The representative spectrum (*black*) of an AmtB nanodisc with DMPC and MSP1D1(−) acquired with a modified Orbitrap mass spectrometer using 130 V in the HCD cell. (A) shows the enlarged *m*/*z* region (*black*) and its predominant charge state (+17, *green*) containing a series of distinct peaks differing by the number of bound DMPC (*grey*). (B) shows another *m*/*z* region (*black*) and the three major charge states (+12 to + 14). The three major charge states overlap more significantly at the integer multiple of the lipid mass (*asterisk*), leading to resonant peaks and a characteristic pattern in the spectrum. Regions (A) and (B) are liberated AmtB oligomer-lipid complexes (corresponding to the *yellow* and *red* species in **Figure 3.6** A; assignment is discussed in **Section 3.3.4**). (C) Deconvolved charge states from +12 to +20 are shown in various colors. Figure from Marty *et al.*[2].

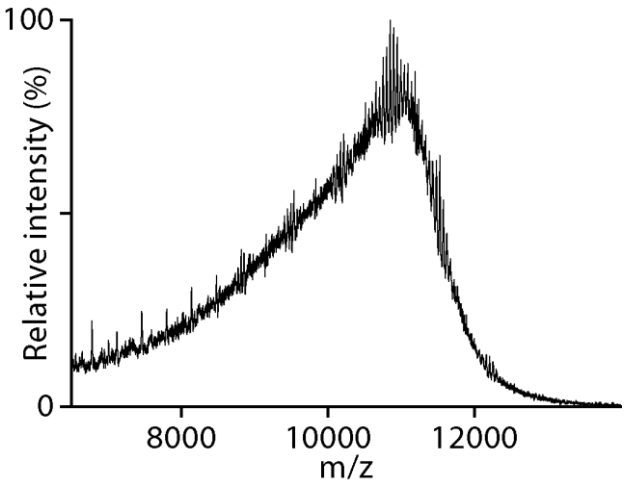

Figure 3.5 A representative mass spectrum of AmtB-POPC-MSP1D1(−) nanodiscs acquired with a modified Synapt G2 mass spectrometer using 130 V collisional voltage in the trap. In contrast to **Figure 3.4**, peaks were poorly resolved due to a lower resolution. Figure drawn by the author.

3.3.3. Charge state disambiguation and mass assignment

Even though high-resolution MS provided clearer spectra, its interpretation remains challenging. First, manual assignment is impractical as the spectra contained hundreds of peaks. Second, the mass assignment requires disambiguation of the overlapping charge states[2]. The peaks in the high *m*/*z* region displayed a constructive overlapping pattern similar to those observed for "empty" nanodiscs (**Figure 3.4** B). This pattern involves overlapping of peaks from neighboring charge states, which cannot be unambiguously determined by the conventional approach for charge state assignment[239,252].

To address these difficulties that are similar to those encountered by "empty" nanodiscs (see **Section 1.4.2**), UniDec was used to assign the mass spectra. This deconvolution algorithm provides an automated and robust approach to disambiguate the charge states of mass spectra, which has been demonstrated

for multiple complex systems, such as hetero-oligomeric heat shock proteins and nanodiscs devoid of membrane proteins[239]. UniDec deconvolves raw spectra into separate charge states (**Figure 3.4** C), which are converted into masses by multiplying by their respective charge. In this manner, masses were assigned from the mass spectra and plotted as a mass distribution called zero-charge mass spectrum (**Figure 3.6** A and B, E and F).

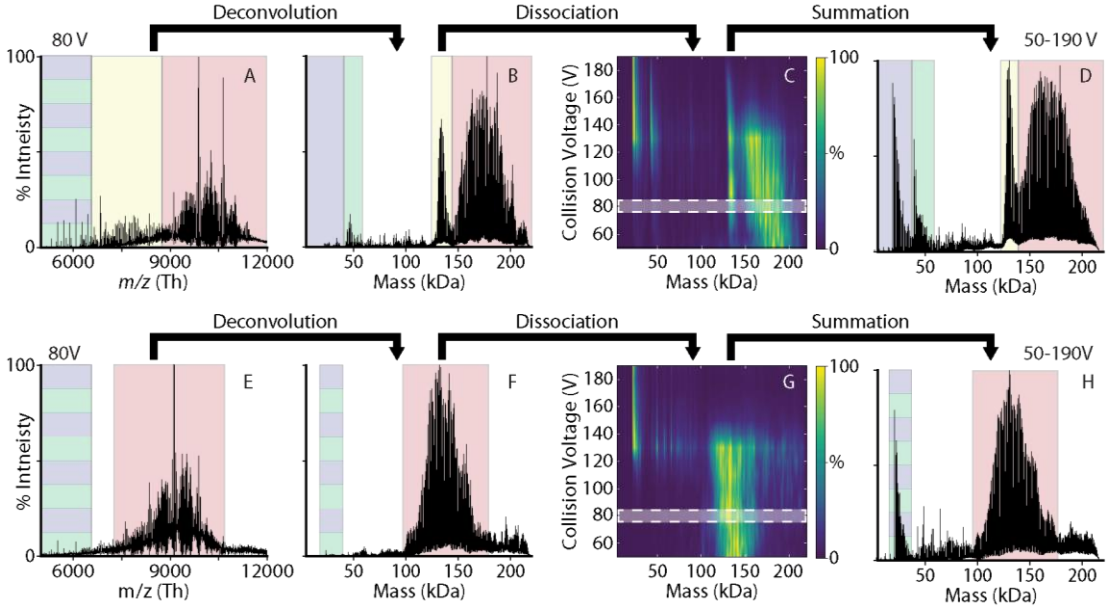

Figure 3.6 The representative mass spectra of membrane protein nanodiscs and their deconvolution. UniDec was used to deconvolve the mass spectra of the AmtB (A) and AqpZ (E) nanodiscs with POPC and MSP1D1(−) acquired at 80 V CID. The deconvolution yielded the zero-charge mass spectra/mass distributions (B) and (F). The zero-charge mass spectra are plotted as a function of collision voltages (C and G). The summed mass distributions over all CID states (D and H) show four distinct dissociation products: membrane protein oligomers with broad (*red*) and narrow (*yellow*, AmtB only) distributions of bound lipids, membrane protein monomers (*green*) and MSPs (*blue*), which may appear at the same *m*/*z* and mass ranges. These four CID products are similarly shaded in all spectra/distributions. Figure from Marty *et al.*[2] and drawn by the author.

Deconvolution revealed several distinct features in the spectra. These features represented diverse CID products of nanodiscs, which varied depending on the collisional activation energy. For instance, features of less than 60 kDa were barely observable at low CID energies, whereas they dominated the spectra at CID above 130 V (**Figure 3.6** C and G). Thus, the deconvoluted mass distributions were summed from 50 V to 190 V CID to identify the major CID products during the dissociation process (**Figure 3.6** D and H). Prominent features were distinguished based on their mass range and the number of lipid adducts. Strikingly, the major features were remarkably similar among all nanodiscs regardless of the encapsulated membrane proteins, lipids, MSPs, and independent replicates. Essentially, four distinct features were found, one of which was unique to the AmtB nanodiscs (**Figure 3.6** D, *yellow*) and the other three were in common to both AmtB and AqpZ nanodiscs.

All the four features identified from the mass distributions consisted of an ensemble of peaks that were separated by the mass of the reconstituted lipid (i.e., 760 Da for POPC and 675 Da for DMPC), indicating a range in the number of lipids per complex. One of these features was located in the high-mass region and showed a wide distribution in the number of bound lipids (**Figure 3.6** D and H, *red*). Due to the broad lipid distribution, neighboring charge states of the same feature overlapped with each other and manifested a characteristic pattern in the mass spectra (**Figure 3.4** B). The overlapping was more effective near the integer multiple of the lipid mass, resulting in resonated peaks in the spectra (**Figure 3.4** B asterisk)[231]. Another feature found in the high-mass region, which was only observable for the AmtB nanodiscs, showed a narrow

distribution of bound lipids that was roughly Gaussian (**Figure 3.6** D and H, *yellow*). In this case, individual charge states were well separated and directly observed in the mass spectra (**Figure 3.4** A). The other two features appeared in the low-mass region. The corresponding CID products carried only a cadre of lipids and displayed a clear charge state series in the mass spectra (**Figure 3.6** D and H, *blue* and *green*).

3.3.4. Composition assignments reveal membrane protein complexes with many bound lipids

The compositions of these features were assigned primarily by fitting their masses to different combinations of the membrane protein monomer, MSP, and lipid. Based on their masses, the two low-mass features were assigned to MSPs (22,044 and 29,980 Da for the MSP1D1(−) and MSP1E3D1(−), respectively) and membrane protein monomers (42,261 and 24,726 Da for the AmtB and AqpZ monomer, respectively) with a cadre of bound lipids (**Figure 3.6** D and H, *blue* and *green*). More importantly, the lowest mass in the series corresponded to unbound MSP and membrane protein monomers. This explains why no further dissociation occurred. Additional evidence is provided by comparing the spectra between nanodiscs formed by different MSPs and membrane proteins. The mass distributions of these two features shifted according to the masses of the MSP and the membrane protein monomer, implying a correlation. Although the identity of the MSP and membrane protein monomer is clear and comprehensive, it is impossible to assign the other two features of higher masses from a single spectrum. For example, the masses of MSP1D1(−) (22,044 Da) and $POPC_{29}$ (22,042 Da) differ by 2 Da, which is not

resolvable in these measurements. Therefore, additional restraints are required to unambiguously assign the composition of these features.

To help assign these high-mass features, a novel mass defect analysis was developed based on Kendrick mass defect analysis (Dr. M. Marty proposed the method and created a tool in UniDec; the author carried out the analysis). A mass defect is the difference between an exact mass (e.g., 1.0078 Da for ^{1}H) and its nominal mass (1 Da for ^{1}H). Instead of a mass scale in Dalton, Kendrick mass defect analysis redefines the mass scale based on a reference unit so that its mass defect is zero[266]. For the analysis of a hydrocarbon mixture, the reference mass is typically methylene. In the Kendrick mass scale, methylene groups have a zero mass defect. The number of the methylene groups in a molecule becomes irrelevant to the mass defect of that molecule. Compounds of the same family, which are atomically identical except for the number of methylene units (e.g., CH_3Cl and C_2H_5Cl), share the same mass defect value. Consequently, Kendrick mass defect analysis is particularly useful to identify families of halogenated compounds and/or hydrocarbon families with the same level of unsaturation. This principle was also found useful in the MS analysis of biomolecules[267]. Here, the masses of the lipids serve as the reference masse, 760.076 Da for POPC and 677.933 for DMPC. For a given mass, the remainder of this mass divided by the reference mass is defined as the normalized mass defect, which has no units. For example, the mass defect of MSP1D1(−) (22,044 Da) is 0.002 (22,044 / 760.076 = 29.002) and 0.516 (22,044 / 677.933 = 32.516) with respect to POPC and DMPC, while it is 0 for the lipid when referred to itself. For a given complex, its mass defect is the sum of the mass

defects of all the constituents. Since lipids have a zero mass defect, the analysis returns the information regarding only the protein components in the complex. Complexes with the same protein composition share the same mass defect.

Mass distributions averaged from 50 to 190 V CID (**Figure 3.6** D and H) were plotted against their mass defects (**Figure 3.7**). The mass defects of the low-mass peaks are in excellent agreement with the theoretical values of the MSP and membrane protein monomers[63,241]. The accuracy of the mass defects as determined from MSP monomers was better than 0.01. Because the typical FWHM for high-mass peaks was around 100 Da, differences below 0.1 can be sufficiently resolved in the plot. The high-mass peaks had the same mass defect (**Figure 3.7**), indicating the same protein composition for the two high-mass features. Mass defect distributions were determined for the high-mass peaks by sampling the mass defect of all mass species greater than 90 kDa from 45 spectra (three replicates of spectra acquired from 50–190 V CID; **Figure 3.9** *left*; Dr. Marty made the mass defect distributions). These distributions were fit to Gaussian distributions to determine the mass defect values of the high-mass peaks (**Table 3.1**; Dr. Marty did the fitting and determined the values).

Given that the mass defect of MSP1D1(−) to POPC is effectively zero, the mass defects reported from the POPC-MSP1D1(−) nanodiscs thus depend exclusively on the oligomeric state of the membrane protein in the complex (**Figure 3.7** *left*). For both AmtB and AqpZ nanodiscs, the mass defects of the

high-mass features agreed with the values of the expected oligomeric states, which is trimeric for AmtB and tetrameric for AqpZ (**Table 3.1**). Although it may be possible to rationalize the mass defects with higher oligomeric structures, these possibilities exceed the limitation of the detected mass. These results confirm that AmtB and AqpZ remained oligomeric after reconstitution and CID experiment.

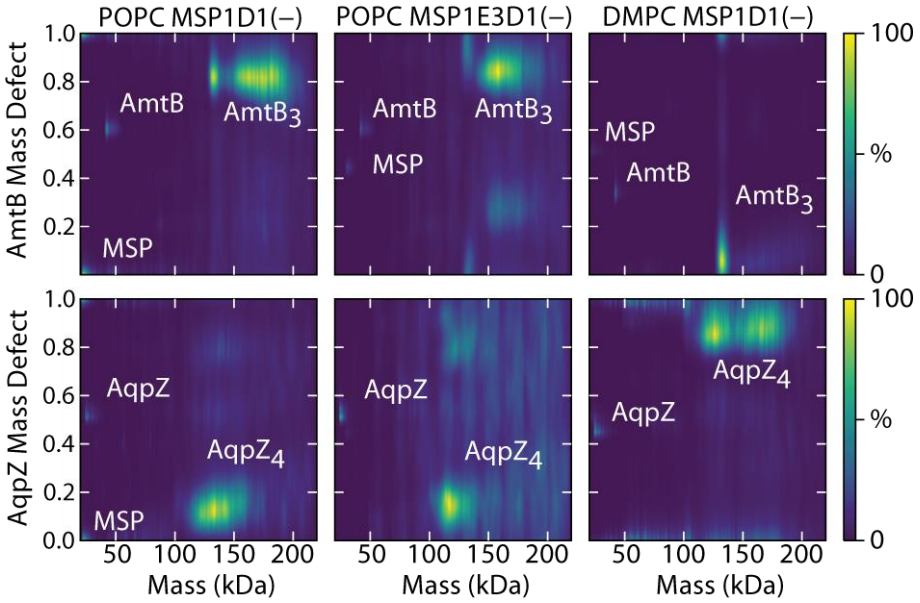

Figure 3.7 The mass defect analysis of the dissociated products of membrane protein nanodiscs. Representative plots are shown for the mass defect analysis for the AmtB (*top*) and AqpZ (*bottom*) in the POPC-MSP1D1(−) (*left*), POPC-MSP1E3D1(−) (*middle*), and DMPC-MSP1D1(−) (*right*) nanodiscs, which are calculated from the average mass distributions of all CID states. The regions for MSP, membrane protein monomers, and membrane protein oligomers are annotated. Figure from Marty *et al.*[2] and drawn by the author.

The identity of the high-mass peaks was deduced by comparing the mass defects of nanodiscs with different MSPs and lipids based on two assumptions. The first assumption stated that the only components in the systems were the

membrane proteins, scaffold proteins, and lipids. This is reasonable since no additional components were observed in the mass spectra even when the complex was completely dissociated (**Figure 3.8**). The second assumption proposed that the same CID products were observed for different types of nanodiscs, which is supported by the remarkably similar spectral features and deconvolution outcomes.

Table 3.1 The normalized mass defect of the trimeric AmtB and tetrameric AqpZ in three different types of nanodiscs. For each nanodisc, the mass defects of all masses greater than 90 kDa and all collision energies from three replicates are averaged (see also **Figure 3.9** *left*) and fit to a Gaussian distribution to determine the mean and standard deviation. The systematic and slightly higher value of the measured values than predicted is likely caused by solvent or ion adducts[268]. Table from Marty *et al.*[2].

Nanodisc Type	**Mean**	**Std. Dev.**	**Predicted**
AmtB Trimer			
POPC MSP1D1(−)	0.843	0.090	0.803
POPC MSP1E3D1(−)	0.861	0.065	0.803
DMPC MSP1D1(−)	0.068	0.089	0.014
AqpZ Tetramer			
POPC MSP1D1(−)	0.146	0.070	0.124
POPC MSP1E3D1(−)	0.131	0.065	0.124
DMPC MSP1D1(−)	0.938	0.090	0.891

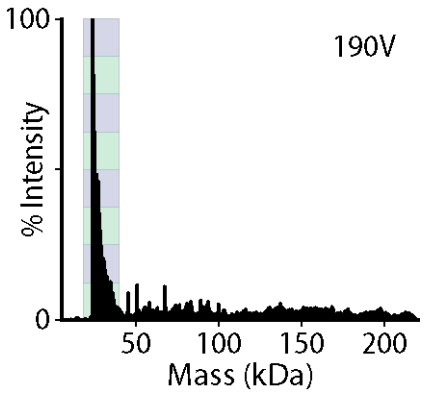

Figure 3.8 A representative mass distribution of nanodisc spectra acquired at high CID voltages. The mass distribution was deconvolved from the AqpZ-POPC-MSP1D1(−) nanodisc spectrum acquired at 190 V CID voltage. Nanodiscs are largely dissociated. Only signals corresponding to MSPs and AqpZ monomers (*blue* and *green* region) were observed.

The MSP was found to be absent from the complexes giving the mass defects of these complexes were independent of the mass defect of the MSP. The mass defects of MSP1E3D1(−) (0.443) and MSP1D1(−) (0.002) in POPC differ by 0.441, which will cause a shift in the mass defect by more than a standard deviation of the peak (smaller than 0.1 in mass defect units). Despite this difference, similar mass defects were observed for the high-mass species when the scaffold protein was changed to MSP1E3D1(−) (**Figure 3.7** *middle* and **Table 3.1**). The insensitivity of the mass defect to the mass defect of the scaffold protein denotes that the high-mass species do not have any MSP. The same conclusion can be drawn by switching the lipid to DMPC (**Figure 3.7** *right*), which will change the mass defect of all the components. In this case, the mass defect remained consistent to the value predicted for the membrane protein oligomers devoid of the two MSP belts (**Table 3.1**). Overall, the mass defect analysis from the three different types of nanodisc converges on the same

assignment for the high-mass peaks — complexes containing only lipids and membrane protein oligomers (**Figure 3.9** *left*).

The high-mass peaks were confidently assigned to the AmtB trimer and AqpZ tetramer with a range of 0 to 120 bound lipids (**Figure 3.9** *right*; Dr. M. Marty made the lipid distributions). Since no intact nanodisc was observed, both MSP scaffolds may be lost during ionization and/or desolvation, leaving membrane protein oligomers with many bound lipids. The low-energy dissociation of nanodiscs allowed preservation of a large extent of bound lipid on the membrane proteins. Because the number of bound lipids observed here was much greater than previously observed for any other membrane mimetics, nanodiscs present a unique opportunity to probe a wider range of membrane protein-lipid interactions.

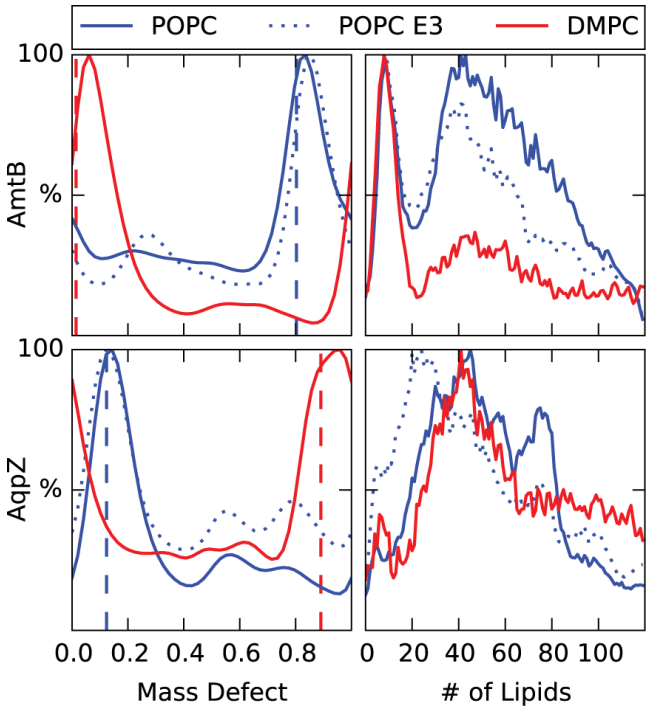

Figure 3.9 Average mass defect values and distributions of the number of lipids in AmtB and AqpZ nanodiscs. The figure shows the mass defect distributions for AmtB (*top left*) and AqpZ (*bottom left*) in POPC-MSP1D1(−) (*blue*), POPC-MSP1E3D1(−) (*blue dashed*), and DMPC-MSP1D1(−) (*red*) nanodiscs. The mass defect distribution shown is an average sampling all the mass species above 90 kDa from three replicates for spectra acquired from 50-190V CID. Vertical dashed lines are aligned to the predicted mass defects of the theoretical oligomers in POPC (*blue*) and DMPC (*red*). *Top right* and *bottom right* show the extracted lipid distributions corresponding to the nanodiscs on the *left*. Figure from Marty *et al.*[2].

3.3.5. MS of membrane protein nanodiscs informs different chemical shells of lipid interactions

The population of bound lipids surviving in the complex after CID was similar between nanodiscs with different MSPs and lipids, indicating that protein-lipid interactions were conserved in these nanodisc systems (**Figure 3.9** *right*). Therefore, the distribution of bound lipids observed here represents a particular set of interactions being captured in the MS. Although the lipid distributions were remarkably similar between different nanodisc systems, the lipid distributions observed for AmtB and AqpZ are clearly different. Two well-separated distributions were observed for AmtB nanodiscs, one of which was broad and has tens of lipids bound and the other was narrow with a small number of lipids. In contrast, AqpZ nanodiscs presented only a broad distribution.

To rationalize these lipid distributions from the molecular perspective, MD was used to simulate AmtB and AqpZ oligomers in a POPC bilayer (Dr. M. Marty did all the works for the simulations). The simulations surveyed the lipids based on the nature of their interactions with the membrane protein (i.e., the distance

between two designated subsets of atoms from the lipid and the membrane protein, respectively). Three populations of lipids were investigated, including the annular lipids; the sub-population of annular lipids that interact via the head groups; and the sub-population of annular lipids that interact via ionic contact. The lipids in each populations were counted for each extracted simulation frame and their mean and standard deviation were determined over the course of the simulation (**Table 3.2**).

Table 3.2 The numbers of lipids of different sub-populations of the annular belt determined by MD and MS. MD was used to determine the number of lipids in the annular belt, and the sub-population of annular lipids that interact with the protein through the head groups and ionic contacts. Also listed in the table are the mean and standard deviation of Peak 1, 2, and 3 corresponding to the three overlapping Gaussian distributions fit to the MS lipid distributions (**Figure 3.10** A and B). Table from Marty *et al*.[2].

Protein	Method	Annular Belt	Head Groups	Ionic Contacts
AmtB	MD	78.8±2.5	50.8±2.7	7.5±2.5
Trimer	MS	73.1±30.3	39.6±16.8	8.7±4.2
AqpZ	MD	72.0±2.9	47.5±2.9	12.9±2.6
Tetramer	MS	74.4±16.5	36.6±15.7	5.8±1.8

By comparing the MD results, the lipid distributions observed in MS were found to be associated with the stoichiometry of annular lipids or sub-populations of annular lipids. The number of lipids counted by MD was aligned with three Gaussian distributions fit to the average lipid distributions from MS (**Figure 3.10**

A to D; Dr. M. Marty did the fitting and determined the mean and stand deviation). These distributions generally included the number of lipids predicted for the annular belt, head group interacting lipids, and ionic contacts within one standard deviation (**Table 3.2**). Nanodisc ions had different charge states and experienced different numbers of collisions during CID. These factors lead to slightly different dissociation among nanodisc ions and thus broad lipid distributions in the spectra. Therefore, the numbers of lipids as determined from these distributions have a large standard deviation comparing to the results of simulations. Comparison between individual CID states revealed a transition of the number of bound annular lipids. (**Figure 3.10** E and F). While the center of the lipid distribution at low energies corresponded to the population of lipids predicted for the annular belt, it shifted to the number of annular lipids in the head group sub-population at intermediate energies. The transition of the lipid distributions indicates a progressive dissociation of weakly bound annular lipids that eventually exposes the sub-population of annular lipids binding through their head group.

The narrow distribution observed for AmtB nanodiscs (**Figure 3.6** D, *yellow*) agreed with the prediction of the number of ionic contacts. This suggests a second dissociation pathway that preserves only the lipids interacting via ionic interactions. Although a similar distribution was observed in some spectra of AqpZ nanodiscs, they deviated from the number of ionic contacts predicted and were much weaker. The absence of an equivalent distribution for AqpZ implied that the AqpZ-lipid complexes undergo oligomer dissociation rather than lipid

stripping upon CID. This is likely due to the larger number of ionic contacts (**Table 3.2**) and the generally stronger interactions with the membrane bilayer.

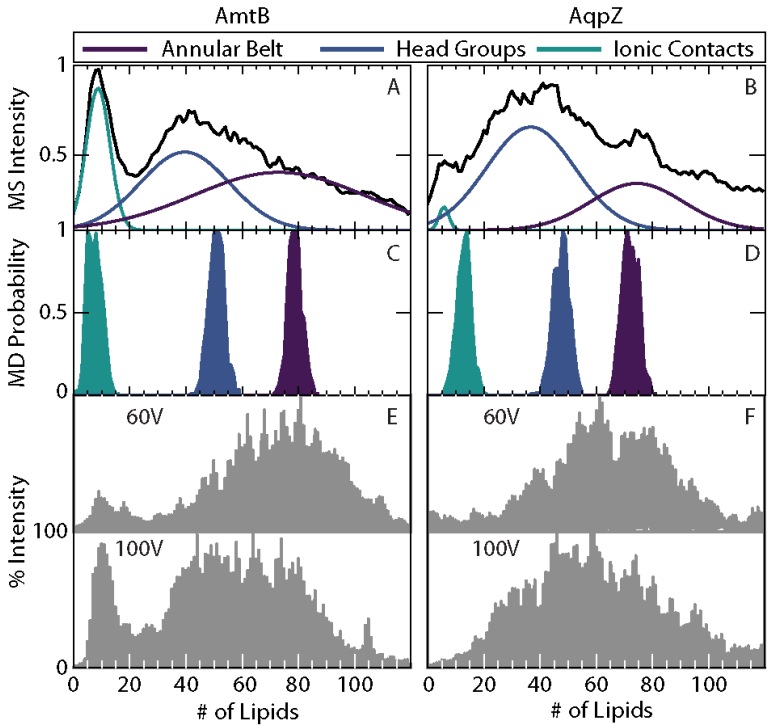

Figure 3.10 Comparison of the protein-lipid interactions observed by MS and predicted by MD. The MS determined average lipid distributions (averaged over a range of collision energies) of the AmtB (A) and AqpZ (B) is fit to three overlapping Gaussian distributions (the number of Guassian was determined by F-test; see **Section 3.2.2**). These Gaussians are compared with the number of lipids predicted by MD (C and D) for the annular belt (*purple*), the sub-population interacting with the head group (*blue*), and the sub-population binding through ionic contacts (*cyan*). (E) and (F) are the representative lipid distributions at 60 V and 100 V CID, which demonstrate shifting of the lipid distributions. Figure adapted from Marty *et al*.[2].

Since nanodisc dissociation has two competing pathways, delipidation and dissociation of the protein oligomer, the pathway that takes place should be related to the chemistry of the lipids. An AmtB-MSP1D1(−) nanodisc prepared with POPG (16:0-18:1) was analyzed by MS. The zero-charge mass spectrum

of the POPG nanodisc showed a broad lipid distribution similar to that of POPC nanodiscs (**Figure 3.11** A and B); however, peaks corresponding to the ionic contacts were not observed for the POPG nanodisc. More importantly, the spectra of POPG nanodisc showed a significant increase in the dissociated protein monomer. This is possibly because that lipid dissociation in nanodiscs is charge-driven. Because POPG cannot carry positive charges, it is more difficult to dissociate in positive polarity comparing to POPC. Therefore, POPG nanodiscs favor the dissociation pathway toward oligomer dissociation. In this case, dissociation of annular lipids is expected to be polarity sensitive. Experiments using negative polarity are required to confirm this. However, the broad lipid distributions observed for POPC and POPG nanodiscs are similar, demonstrating that charge-driven dissociation may play a less important role in the early stage of nanodisc dissociation. Furthermore, POPG cannot carry positive charges and its dissociation does not lead to charge reduction[255,256]. An elevated positive charge is loaded on the AmtB oligomer (**Figure 3.11** C and D), leading to destabilizing charge repulsion[269]. Thus, the AmtB-POPG complex undergoes oligomer dissociation before exposure of the ionic contact sub-population. This result confirms that nanodisc dissociation is dependent on the relative strength between the protein-protein interactions and protein-lipid interactions, which determines the lipid distribution observed.

MD simulations provide insight into the lipid distributions observed in MS and allow the mechanism of membrane protein nanodisc dissociation to be hypothesized (**Figure 3.12**). Ions transmitted into the gas-phase are collisionally activated to desolvate and reveal resolved ion signals. Meanwhile, the MSP

scaffolds and bulk lipids are removed, leaving the membrane protein with an annular belt (**Figure 3.12** *red*). Based on the known properties of intermolecular bond strengths in the gas-phase, the strengths of non-covalent interactions follow: ionic interactions > hydrogen bonds > dipole interactions > non-polar interactions[270]. Therefore, further activation begins to strip the weakly bound lipids on the complexes, leaving lipids that bind via their polar head groups. In the case where the protein oligomer is relatively stable (e.g., AmtB-PC lipid complexes), lipid dissociation can proceed further until the sub-population of ionic contact-annular lipids is remained (**Figure 3.12** *yellow*). High activation energy eventually dissociates the complexes into protein monomers with a few bound lipids (**Figure 3.12** *green*).

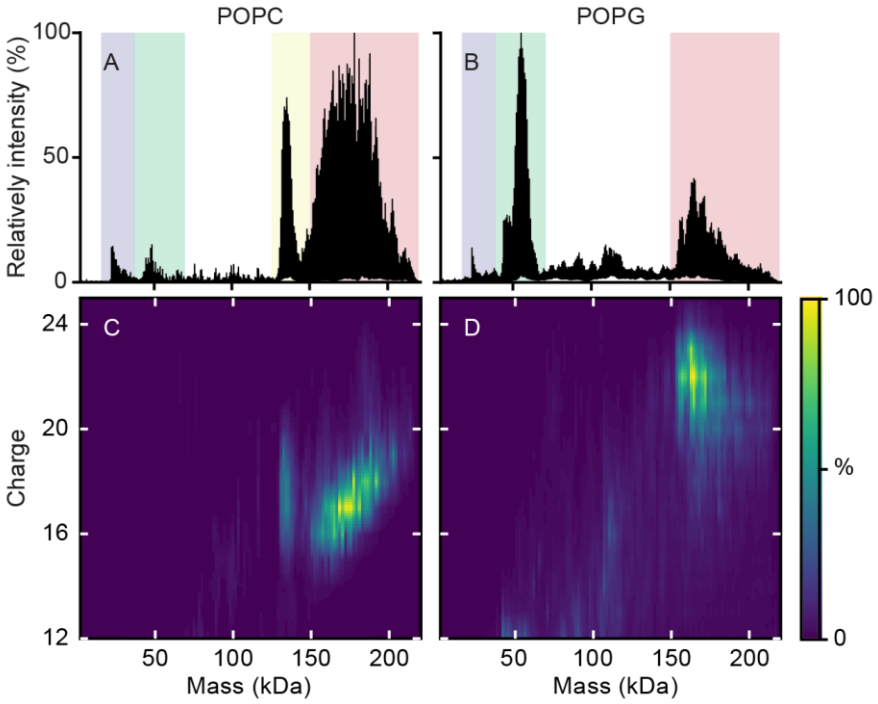

Figure 3.11 The mass distributions of AmtB nanodiscs containing POPC or POPG. The zero-charge mass distributions of (A) AmtB-POPC and (B) AmtB-POPG nanodiscs are deconvolved from the mass spectra acquired at 80V. The shaded regions correspond to the AmtB trimers

with many bound lipids (*red*), AmtB trimers bound to approximately nine lipids (*yellow*), dissociated AmtB monomers (*green*), and MSP monomers (*blue*). (C) The AmtB-POPC complexes have a lower charge state than that of (D) the AmtB-POPG complexes. Figure drawn by the author.

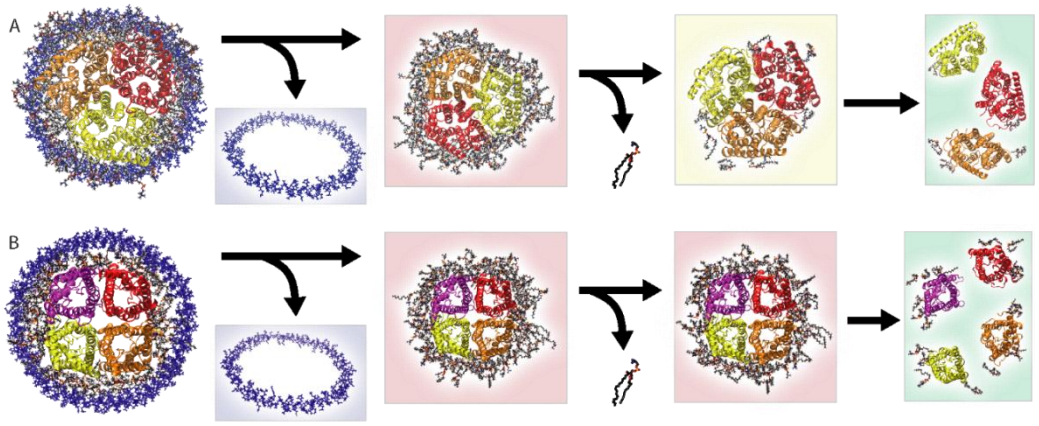

Figure 3.12 Proposed dissociation pathway of membrane protein nanodiscs. In the gas-phase, MSPs and bulk lipids of AmtB (A) and AqpZ (B) nanodisc are removed prior to the resolution of the spectra (*blue box*). Membrane protein oligomers with annular lipids (*red box*) are released at low activation energies. Upon further activation, lipids are gradually stripped until reaching the sub-population of annular lipids interacting via the head groups (for AqpZ) or ionic-contacts (for AmtB, *yellow box*), followed by oligomer dissociation into membrane protein monomers (*green box*). MSPs are in *blue*. Lipids are in *dark grey*. Individual protein monomers are in *purple*, *red*, *yellow*, and *orange*. Figure from Marty *et al*.[2].

The dissociation pathway that membrane protein nanodiscs undergo in the gas-phase enables the capture of a wide range of protein-lipid interactions around the membrane protein using MS. As notable differences were observed between the AmtB and AqpZ nanodiscs, POPC and POPG nanodiscs, the exact mechanism is expected to be dependent on the membrane protein, lipid, polarity, and nature of the gas-phase activation. Nevertheless, the mass spectra and the overall dissociation mechanism of these nanodisc systems were

broadly similar. Membrane proteins were initially ejected from nanodiscs with a similar number of annular lipids. Therefore, these results demonstrate the use of MS with nanodiscs to interrogate the lipid annular belt and protein-lipid interactions in the context of lipid bilayers.

3.4. Conclusion

This chapter used MS with nanodiscs to interrogate membrane protein-lipid interactions in the context of lipid layers. The mass spectra of membrane protein nanodiscs were highly complex and difficult to analyze. This can be attributed to the following: the impracticality of manual inspection; the ambiguity of charge states assignment; and the ambiguity of composition assignment. Assimilating novel strategies with the previously established methodology on “empty” nanodiscs, this work provides a framework to extract the information of lipid interactions encoded in the mass spectra of the membrane protein nanodiscs. The combination of MS and MD revealed the dissociation pathway of membrane protein nanodiscs in the gas-phase. Because the distinct chemical shells of different lipid interactions can be retained on the membrane proteins in the gas-phase, it is possible to determine the lipid stoichiometries in the individual annular subshells/sub-populations. These stoichiometries are parameters that were previously considered difficult to measure, but are useful in characterizing the protein-lipid interactions in the annular belt. Given that annular lipids are in direct contact with membrane proteins, they are likely to mediate properties of membrane proteins such as stability and conformation. Our methodology may facilitate understanding the membrane protein structures and functions in the membrane environment.

Chapter 4. An investigation of GPCR-lipid interactions and their role in G protein signaling pathway by native MS

4.1. Introduction

G protein-coupled receptors (GPCRs) are the largest family of integral membrane proteins in vertebrates; they mediate a vast variety of signaling pathways and physiological processes. Playing an important role in many pathogenesis processes, GPCRs account for 12% of human protein drug targets and are the targets of roughly one third of the currently marketed drugs (**Figure 1.2**)[9]. Despite the high pharmaceutical value of GPCRs, their signaling mechanisms are elusive and challenging to study. Their complexity is a result of the conformational plasticity of GPCRs to convey signals through interactions with distinct transducers (**Figure 1.7**)[117,271] and the various ligands that regulate the degree of the signaling response and the selectivity of different transducer-mediated pathways[126,272,273]. The signal transduction of GPCRs thus involves a complex interaction network between the ligand, receptor, and transducer that are difficult to monitor and elucidate.

GPCR activation involves a relatively large-scale rearrangement of the transmembrane (TM) bundle that leads to different conformations and coupling to different transducers[218]. Previous studies have shown that binding of allosteric ligands at the protein-membrane interface, distinct to the long-recognized ligand binding at the orthosteric binding pocket, can module the conformation of the GPCRs[54,105,274,275]. For example, BPTU, an antagonist of purinergic receptor ($P2Y_1R$), binds to the receptor at the external interface of the transmembrane domain, preventing receptor activation by restricting the

movement of TM2 and TM3[105]. The serotonin$_{1A}$ receptor exhibits increased ligand binding in cholesterol-depleted membranes[276]. Phospholipids were seen to modulate the TM6 outward movement in the β_2-adrenergic receptor (β_2AR), indicating their functional significance on GPCR activation[54]. This evidence, therefore, highlighted the functional significance of the protein-lipid interfaces in regulating GPCR functions.

Given the fact that lipid-mediated effects on GPCR function are well-documented[54,138,142,275,277-279], GPCR-lipid interactions must play an important role in GPCR signaling. However, it is difficult to directly capture the GPCR-lipid interactions, as well as the interactions with other small-molecule modulators and binding partners. The molecular insight of how lipids influence GPCR function remains elusive. A recent study established a native MS method to interrogate P2Y$_1$R-ligand interactions and achieved, for the first time, preservation of non-covalent complexes of a class A GPCR and its ligand in the gas-phase[240]. This breakthrough demonstrated the new opportunity of native MS in GPCR studies and shed light on characterization of GPCR-lipid interactions.

Here, this chapter characterized GPCR-lipid interactions by employing native MS in combination with coarse-grained (CG) simulation, lipidomics, mutagenesis, and multiplexing strategies. Three well-characterized class A GPCRs, the β_1-adrenergic receptor (β_1AR), adenosine A_{2A} receptor (A_{2A}R), and neurotensin receptor 1 (NTSR1) were studied and compared to generalize the lipid interactions of class A GPCRs. We discovered that phosphatidylinositol

bisphosphate (PIP_2) binds preferentially at the intracellular face of these receptors, stabilizes the receptors in active states, and selectively enhances the coupling efficiency of G proteins by tethering the receptor with specific G proteins.

4.2. Methods

4.2.1. Sample preparation

Methods for protein concentration, concentration determination, filtration and centrifugal buffer exchange are in the **Section 2.2.1**. All results were obtained from three biological replicates to establish reproducibility, except that lipidomics results were obtained from a single experiment.

Proteins: Isoprenaline-bound (β118) and ligand-free (β114-E130W) thermostabilized turkey (*M. gallopavo*) β_1AR[280-282], human wild-type $A_{2A}R$[134], Mini-Gα (Mini-$G\alpha_s$ and Mini-$G\alpha_{i(s)}$)[283,284], and Nb6b9[132,285,286] were obtained from Professor Chris Tate's laboratory. Heterotrimeric G protein ($G\alpha_{i1}\beta_1\gamma_1$) was obtained from Professor Andreas Plückthun's laboratory. See Yen *et al*.[3] for detail.

Constructs of NTSR1: Expression plasmids of two rat NTSR1 stabilized variants, NTSR1 HTGH4-ΔIC3B and NTSR1 HTGH4 43–421, were obtained from the Plückthun laboratory[287,288] (See Yen *et al*.[3]). Briefly, the expression plasmids of NTSR1 were optimized for expression in *E. coli*. The protein sequence of NTSR1 HTGH4-ΔIC3B spans from amino acid residues 50 to 390 with an ICL3 deletion between the residues 273 and 290. The protein was thermostabilized with 26 point mutations. NTSR1 HTGH4 43–421 is identical to NTSR1 HTGH4-ΔIC3B except that the protein sequence covers residues from 43 to 421.

Deletion of the residues between 394 and 398 (RRKKR) in NTSR1 HTGH4 43–421 and the subsequent TM1 (R43G, K44G, and K45G), TM4 (R135I, R137T, K139L, and K140L), and TM7 (R311N) mutants were generated using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, USA) (the author did the mutagenesis; Dr. Hsin-Yung Yen designed the primers). All mutagenesis was confirmed by Sanger sequencing service of Source Biosciences.

Proteins as expressed from these constructs have a N-terminal fusion of maltose binding protein, 8× His-tag, and 3C protease recognition site, and a C-terminal fusion of 3C protease recognition site, thioredoxin, and 10× His-tag.

Production of NTSR1: The author did this work. Expression plasmids of NTSR1 were transformed into *E. coli* BL21-Gold (DE3) via heat shock. Cells were grown in 2YT medium (16 g tryptone, 10 g yeast extract, and 5 g NaCl per liter of water) supplemented with 1% (w/v) glucose and 100 µg/mL ampicillin at 37 °C, 200 rpm. At OD_{600} of 0.5, 1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG; 1 M IPTG stocks) was added to initiate overexpression. Cells were cultivated at 28 °C for 16 h and then harvested by centrifugation (5,000 ×*g*, 10 min). The cell pellet was resuspended using a solubilization buffer (50 mM HEPES pH 8, 0.2 M NaCl, 20% (v/v) glycerol, 5 mM $MgCl_2$, 0.15 % (w/v) lysozyme, 0.0025 % (w/v) DNase I, 2 % (w/v) n-Dodecyl-β-D-maltoside (DDM), and 0.1% (w/v) cholestery hemisuccinate (CHS) Tris salt; 25 mL per liter of cell culture) and stirred for 1 h at 4 °C. After that, cells were lysed using a

microfluidizer M-110P (Microfluidics Corp.). The lysate was centrifuged for 30 min at 28,000 ×*g* to remove insoluble materials.

The supernatant containing the detergent-solubilized NTSR1 was supplemented with 15 mM imidazole (1 M stocks) and loaded on a 5 mL HiTrap Talon column equilibrated with loading buffer (25 mM HEPES pH 8, 10% (v/v) glycerol, 0.6 M NaCl, 0.1% (w/v) DDM, and 20 mM imidazole). Following washing with loading buffer (250 mL), NTSR1 was eluted with elution buffer (25 mM HEPES pH 8, 10% (v/v) glycerol, 0.15 M NaCl, 0.1% (w/v) DDM, and 0.25 M imidazole). The eluent was concentrated and injected on a Superdex 200 Increase 10/300 size exclusion chromatography (SEC) column equilibrated with elution buffer without imidazole (**Figure 4.1** A). Desalted protein fractions were pooled and the protein concentration was determined from the A280. HRV 3C protease (0.25-fold molar ratio to the protein; Takara Bio) and 0.1% LMNG (5% stocks) was added to the protein and incubated for 1h to cleave the fusion proteins. To remove the fusion proteins and HRV 3C protease, the sample was injected on a Superdex 200 Increase 10/300 SEC column equilibrated with SEC buffer (10 mM HEPES pH 8, 10% (v/v) glycerol, 0.15 M NaCl, and 0.01% (w/v) LMNG) (**Figure 4.1** B). Peak fractions containing the NTSR1 were combined and concentrated to about 100 µM. The NTSR1 was mixed with 1:1 (v:v) SEC buffer containing 50% glycerol to make a 50 µM NTSR1 stock. Protein purity was estimated using SDS-PAGE (**Figure 4.1** C). The stock was snap frozen and stored at −80 °C before use.

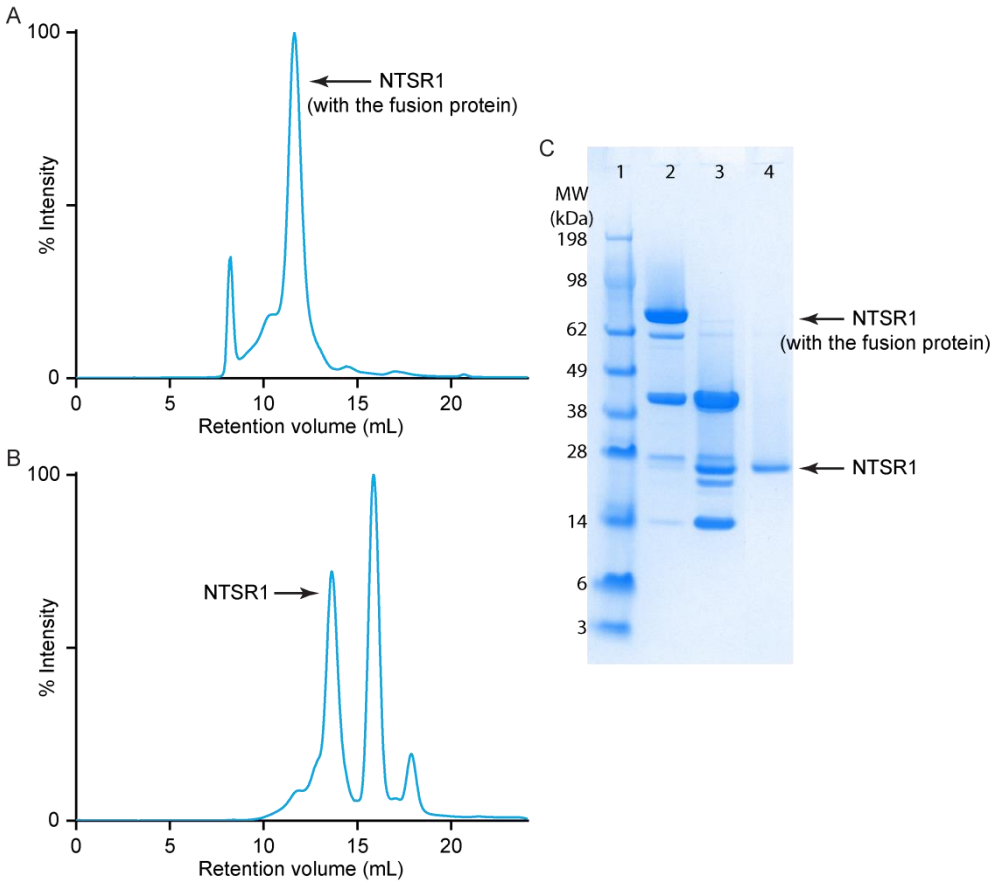

Figure 4.1epresentative SEC profiles (Superdex 200 Increase 10/300 column) and SDS-PAGE of NTSR1. (A) and (B) are SEC chromatograms of NTSR1 before and after removal of the fusion proteins. (C) Protein purity was estimated using SDS-PAGE. Lanes correspond to (1) molecular weight marker, (2) NTSR1 with fusion proteins obtained from (A), (3) after incubation with HRV 3C protease, and (4) NTSR1 obtained from (B). Figure drawn by the author.

Lipids: The author did this work. All lipids were purchased from Avanti Polar Lipids (Alabama, AL). Powder lipids of phosphatidylinositol (PI, 18:1/18:1), phosphatidylinositol monophosphate (PI(4)P, 18:1/18:1), phosphatidylinositol bisphosphate (PI(4,5)P_2, 18:1/18:1), and phosphatidylinositol trisphosphate (PI(3,4,5)P_3, 18:1/18:1) were dissolved in spray buffer (see **Section 4.2.2**). Lipid concentration was determined by phosphate analysis and prepared as 3 mM stocks[243]. The stocks were stored at −20 °C before use.

PI(4)P, PI(4,5)P_2, and PI(3,4,5)P_3 are referred as PIP, PIP_2, and PIP_3 respectively, in the rest of this chapter.

4.2.2. MS and simulation methods

Native mass spectrometry analysis: MS measurements were performed using a modified Orbitrap mass spectrometer as described in **Chapter 2** and **Chapter 3**. Ions were collisionally activated in the HCD cell by applying voltages between 100 and 130 V. The C-trap entrance lens was set to 4 V. The temperature of the ion transfer capillary was set to 60 °C. The S-lens RF potential was set to 100 V. For all MS measurements or assay preparations, purified GPCRs were buffer exchanged (Zeba^{TM} Spin) into spray buffer (for NTSR1: 200 mM ammonium acetate pH 7.5, 0.001% LMNG, 0.001% foscholine, and 0.001% CHS; for A_2AR and β_1AR: 200 mM ammonium acetate pH 7.5, 0.01% DDM, 0.001% foscholine, and 0.001% CHS)[240].

Data analysis: Mass spectral results were analyzed by Universal Deconvolution (UniDec) software as introduced in **Section 1.4.2** to determine the relative intensity of each lipid-bound state[239]. Deconvolution typically used a mass range of 10–100 kDa, a charge range of 5–25, and a Gaussian peak shape with a 1.2 Th FWHM. The deconvoluted mass distribution was linearized at 1 Da interval.

Endogenous lipid identification: Dr. Yen did this work. Detail can be found in Yen *et al*.[3]. Briefly, purified protein-lipid complexes were analyzed using a

tandem MS method to dissociate bound lipid adducts (**Figure 4.2**). Dissociated lipids were detected at low *m*/*z* region and the *m*/*z* can be determined accurately. Accurate mass of endogenous lipids were used for identification.

Lipidomics analysis: Dr. Yen prepared the lipid extract. Dr. Di Wu ran the LC-MS measurements. The author analyzed the data and assigned the lipids. Detail can be found in Yen *et al*.[3]. Briefly, co-purified lipids were extracted from GPCRs for lipidomics analysis. Lipids observed in lipidomics were compared to the lipid adducts observed in native MS spectra. Lipids were assigned using Lipidomics Gateway by matching the masses of precursor and product ions. Mass tolerance used for database search was 0.01 and 0.1 *m*/*z* for precursor and product ions, respectively.

Lipid titration experiments: Dr. Yen did this work. Detail can be found in Yen *et al*.[3]. Briefly, receptors were incubated with lipids of interest at different concentrations to compare the interactions of the receptors with different lipids.

Lipid titration of mix PIs: The author did this work. PI, PI(4)P, PI(4,5)P_2, and PI(3,4,5)P_3 were mixed at equal molar ratio. Protein stocks were buffer exchanged (Zeba^{TM} Spin) to spray buffer. Receptors (10 µL, 5 µM) were incubated with 0—80 µM (total lipid concentration) PIs for 5 min at 4 °C prior to MS analysis. The binding level of each lipid adducts was evaluated as an intensity ratio to the apo receptor. Only one lipid-bound species were

considered in the analysis. One binding site model (Prism software) was used to generate binding curves.

Identification of preferential PIP_2-binding sites on NTSR1: The author did this work. Unmodified NTSR1 HTGH4 43–421 Δ394–398 and its mutants (buffer exchanged in spray buffer) were mixed at 1:1 molar ratio with a total protein concentration of 12 µM, 10 µL. The protein mixture was incubated with 1.25-fold molar equivalents of PIP_2 for 5 min on ice prior to MS analysis. The spectral results were analyzed using UniDec. The intensity of the PIP_2-bound states was normalized to the apo-states. The effects of the mutations were evaluated as the intensity ratios of the PIP_2-bound states between the mutants and unmodified receptor.

Comparison of PIP_2 binding between inactive and active β_1AR: Dr. Yen prepared the assay and acquired spectra. The author analyzed the data. See Yen *et al*.[3] for detail. Briefly, equal molar mixtures of isoprenaline-bound (β118) and ligand-free (β114) β_1AR were incubated with PIP_2 for MS analysis. The PIP_2 binding level was compared between the two conformations. Analysis was same as above.

GTPase assay: Dr. Yen did this work. Detail can be found in Yen *et al*.[3]. Briefly, the GTPase activity of trimeric $G\alpha_i\beta\gamma$ was studied in the absence or presence of NTSR1 and PIP_2.

MS analysis of β_1AR-Gα_s and β_1AR-Nb6b9 complexes: Dr. Yen did this work except that the author acquired the β1AR-Nb6b9 spectra. Detail can be found in Yen *et al*.[3]. Briefly, β_1AR coupling with Gα_s and Nb6b9 was studied without lipids.

Effects of lipids on β_1AR-effector coupling: Dr. Yen did all the works for β_1AR-Gα_s coupling in the presence of PIP_2 or PS. For the analogue experiments of β_1AR-Nb6b9 and β_1AR-mini-G$\alpha_{i(s)}$, Dr. H. Yen and Dr. Idlir Liko prepared the samples; the author acquired the spectra and analyzed the data. Detail can be found in Yen *et al*.[3]. Briefly, β_1AR was coupled with different effectors in the presence of PIP_2 or PS to investigate the effect of PIP_2 and PS on β_1AR-effector coupling. The spectral results were analyzed using UniDec. For each spectrum, the intensity ratios of the lipid-bound states to the apo form of the receptor in the uncoupled/coupled state were determined and plotted as a bar chart.

Coarse-grained molecular dynamic simulations: Dr. George Hedger did this work. Detail can be found in Yen *et al*.[3]. Briefly, GPCRs were simulated in lipids bilayers containing PIP_2 to investigate the potential PIP_2-binding hotspots in GPCRs. Analogue simulations were performed using PS for comparison.

Calculation of potential of mean forces: Dr. Wanling Song did this work. Detail can be found in Yen *et al*.[3]. Briefly, Potential of mean forces (PMFs) of the interactions between mini-Gα_s and A_{2A}R in lipid bilayers with and without

PIP_2 were calculated. The results were used to evaluate the stabilization effect of PIP_2 on GPCR-effector coupling.

Results and discussion

4.2.3. Characterization of class A GPCR interactions with different lipids

To investigate how GPCRs interact with lipids, this study started with identification of endogenous lipids that directly interact with GPCRs using native MS (The author received supervision from Dr. Hsin-Yung Yen). Typically, native MS is used to probe lipids that directly bind to membrane proteins. The mass of the lipid adduct is determined according to the mass difference between the apo-protein and the lipid adduct species. This mass is used to match the lipid species detected in lipidomics for the identification of the lipid[87,209,214]. However, it was shown that even high-resolution mass spectrometry struggles to resolve lipid adducts which differ by two methylene groups in the acyl chains[209]. A lipid adduct peak may not correspond to a single lipid species but a heterogeneous mixture of lipids of similar masses. The mass of an adduct peak measured thus represents an average mass. Additionally, it might be ambiguous whether an adduct peak results from a single adduct or multiple smaller adducts in some measurements.

To address these problems, an approach was established to determine the stoichiometry and accurate mass of the lipid adducts (**Figure 4.2**). Briefly, voltage is applied at the sample cone to activate membrane protein complexes and liberate them from the protective detergent micelles[289]. Protein-lipid complexes are subsequently isolated by the quadrupole and further activated in the collision cell (e.g., the trap in the Synapt and HCD cell in the Q-Exactive) to dissociate the bound lipids. The dissociated lipid adducts are detected at the low *m*/*z* region, allowing accurate mass determination for these lipids.

Compared to the conventional MS^1 workflow that activates micellated proteins in the collision cell, activating ions in the sample cone spares (200 V) the quadrupole and collision cell for MS^2 experiments of the protein-lipid complex. This enables direct detection of the lipid adducts with high mass accuracy.

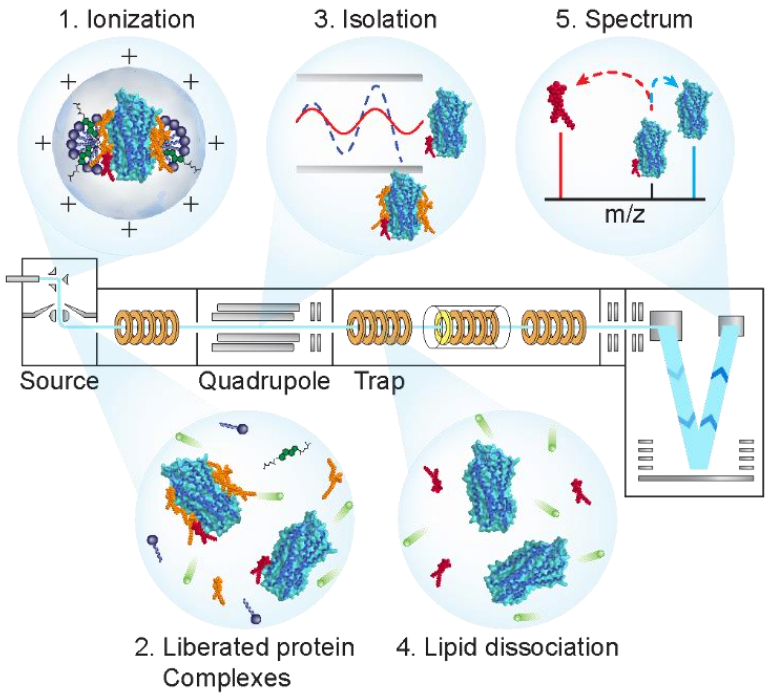

Figure 4.2 Schematic of tandem MS based lipid identification using a modified Synapt G2 mass spectrometer. (1) Protein-lipid complexes in detergent micelles are introduced into the mass spectrometer by nano-electrospray ionization (nESI). (2) A voltage (200 V) is applied to the sample cone to liberate protein-lipid complexes from detergent micelles, allowing the subsequent (3) quadrupole isolation of the targeted protein-lipid complexes. (4) Lipid adducts are dissociated by activating the protein-lipid complexes in the trap and (5) detected by the time-of-flight mass analyzer. This method enables interrogation of the binding stoichiometry and accurate mass measurement of the lipid adducts. Figure from Yen *et al*.[3] and drawn by the author.

This method was applied to identify the co-purified endogenous lipids of the NTSR1 and β_1AR (**Figure 4.3** A to D; Dr. Yen did this work), which were expressed and purified from *E. coli* and insect cells, respectively (β_1AR (β114)[281,282] was from the Tate lab). The mass spectra of NTSR1 showed a major lipid adduct of 701 Da. This mass was assigned as a phosphatidic acid (PA, 36:2) using the lipidomics gateway database[290]. The presence of the lipid species was further confirmed by lipidomics (**Figure 4.3** E; for lipid lipidomics: Dr. Yen prepared the samples; Dr. Di Wu did the LC-MS measurements; the author assigned the spectra). Similarly, the major lipids of β_1AR at 760, 788, and 1,022 Da were identified as two phosphatidylserines (PS, 34:2 and 36:2) and a phosphatidylinositol phosphate (PIP, 42:5) (**Figure 4.3** C, D, and F). Despite the two receptors showed binding with different lipids, PA, PS, and PIP are anionic lipids, implying a preference of class A GPCRs to bind anionic lipids.

Rat NTSR1 was heterologously expressed in *E. coli*. The *E. coli* membrane is enriched with PE, PG, and CDL[291,292], whereas the major phospholipids in eukaryotic (including mammalian and insect) cell membranes are PC, PE, PS, and PI[31,293]. Therefore, NTSR1 was studied for its interactions with several lipids commonly found in eukaryotes, including a zwitterionic lipid (PC), a neutral lipid (DAG), and two anionic lipids (PS, PI) as well as PA (Dr. H. Yen did this work). NTSR1 was incubated with different concentration of lipids (0 to 160 μM) and analyzed using native MS. The fraction of receptor binding to the lipid was plotted against the lipid concentration. Lipid binding was observed for the anionic lipids (PS, PI, and PA) but not PC and DAG (**Figure 4.4**), supporting the preference of class A GPCRs in binding anionic lipids.

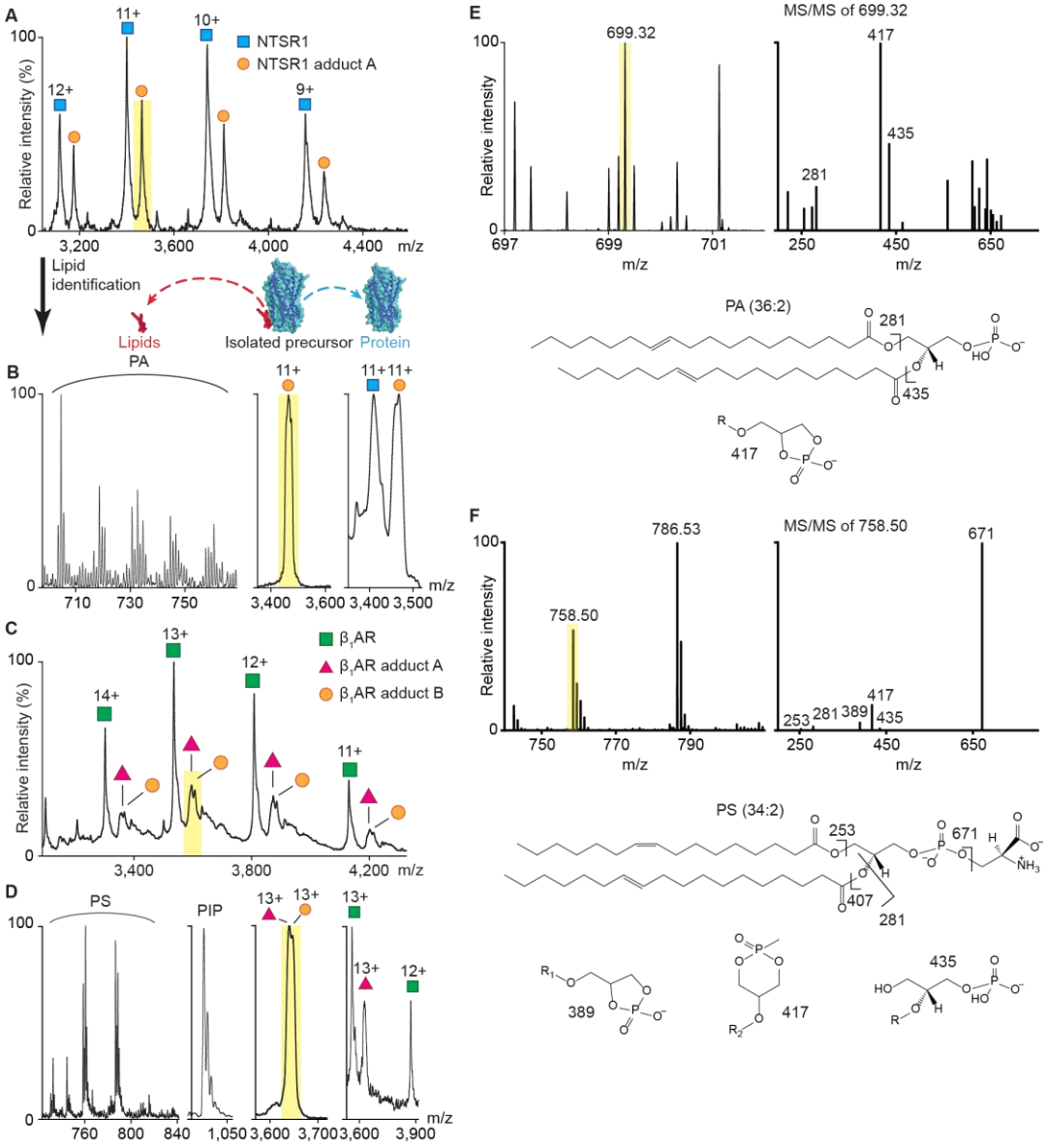

Figure 4.3 Identification of endogenous lipids. The mass spectra were acquired from (A) NTSR1 (HTGH4-ΔIC3B) and (C) β_1AR (β114), which were expressed and purified from *E*. coli and insect cells, respectively. Adduct peaks (*highlighted*) were selected by the quadrupole as shown in (B) and (D) and subjected to tandem MS analysis. In addition to the apo-receptor observed at the high *m*/*z* range (*right*), dissociated lipids (PA for NTSR1; PS and PIP for β_1AR) are observed at the low *m*/*z* range. (E) and (F) show the representative lipidomics results supporting the presence of the identified lipids in (B) NTSR1 and (D) β_1AR. *Left* and *right* spectra correspond to the survey full-scan of the highlighted lipid species and their MS/MS spectra. Insets denote the fragmented product ions identified in the MS/MS spectra. Figure from Yen *et al.*[3] and drawn by the author.

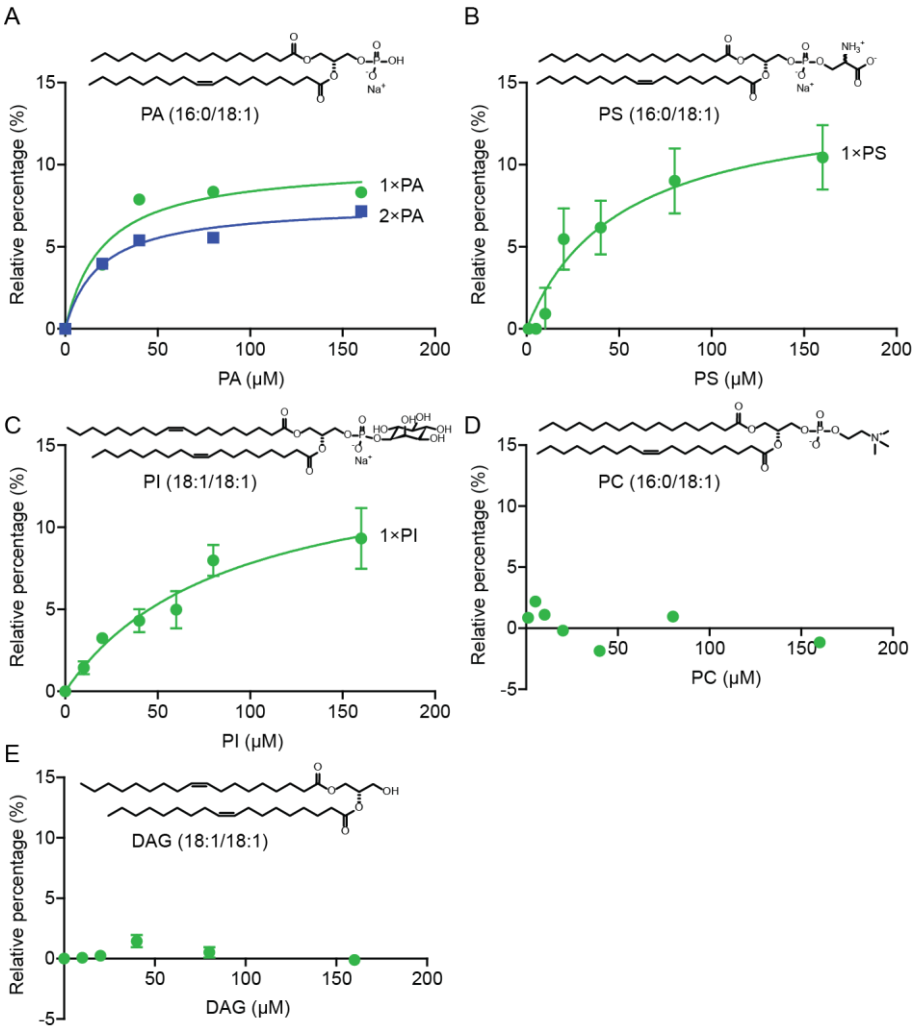

Figure 4.4 The lipid-binding preference of NTSR1. NTSR1 (HTGH4-ΔIC3B) was incubated with 0–160 μM (A) PA, (B) PS, (C) PI, (D) PC, and (E) DAG. Percentage of lipid binding of all lipid-bound states (intensity ratios of lipid-bound states to the apo-state) is plotted against the lipid concentrations and fit by one-site total binding model (GraphPad Prism software). Figure from Yen *et al.*[3] and drawn by the author.

Similarly, β_1AR was incubated and showed interactions with PS (16:0-18:1) and PI(4)P (18:1) (Dr. H. Yen did this work), confirming the interactions as observed for endogenous lipids (**Figure 4.5**). Interestingly, PI(4)P displayed a higher binding affinity compared to PS. This high affinity was consistent with the observation for the endogenous lipid interactions discussed above, given that PIP is a minor lipid (~1%) relative to PS (~20% of total phospholipids)[293,294]. As

PIP and its relatives (PI and phosphoinositides) are involved in various signaling cascades, including GPCR signaling pathways[67,68,295], the high affinity of PIP interaction might have significance in GPCR function.

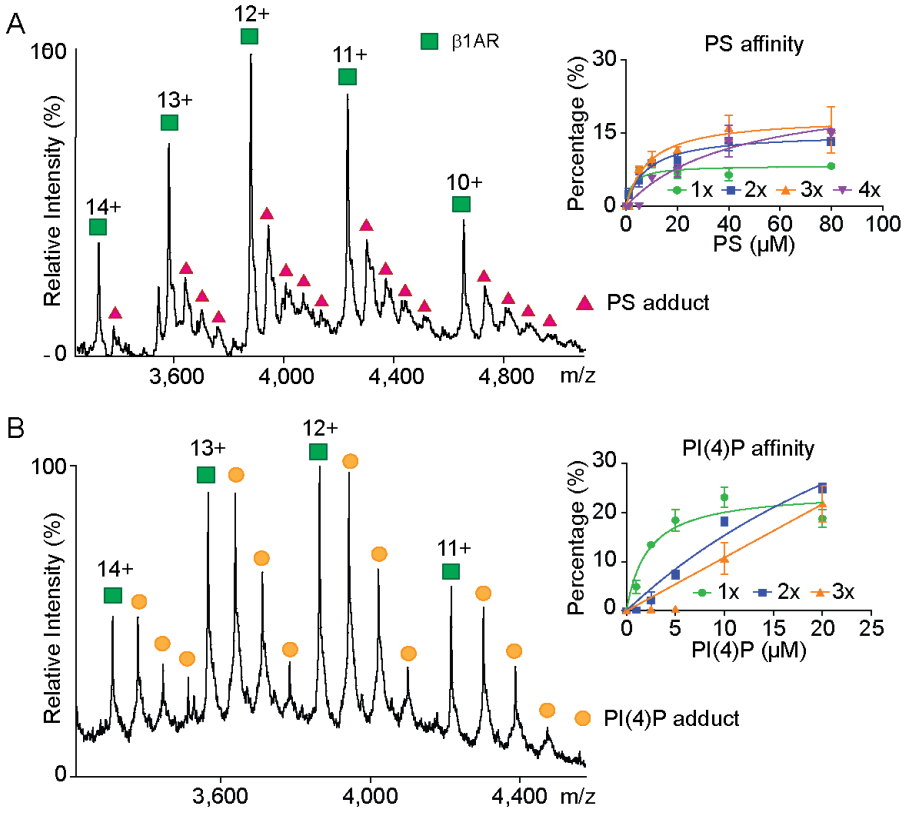

Figure 4.5 The Interactions of β_1AR with PS and PIP. β_1AR (β114) was incubated with 0–20 μM (A) PS (0–80 μM) and (B) PI(4)P. Mass spectra correspond to the lipid concentration at 10 μM. Percentage of one (1×) to four (4×) adduct bound species were plotted against the lipid concentration. Figure from Yen *et al.*[3].

A multiplexing strategy was used to investigate the lipid-binding preference of β_1AR to PI and its derivative phosphoinositides. Lipid-binding experiments were performed by incubating the receptor with an equal molar mixture of PI, PI(4)P, PI(4,5)P_2, and PI(3,4,5)P_3. This competitive lipid-binding assay exploits the ability of MS to separate heterogeneous ions and offers two distinct advantages.

First, different lipid adducts that resolved in the spectra can be studied in a single MS measurement, reducing the analysis time and number. Second, lipid adducts experience the same ionization process and collisional activation energy in the same measurement. The lipid adducts share the same spectral baseline. These minimize the bias between measurements, allowing a straightforward and accurate comparison between different lipid adducts. Accounting for the fact that the mass difference between the progressively phosphorylated PIs is 80 Da, these measurements employed a modified Orbitrap mass spectrometer (**Figure 1.15**) optimized for high-resolution measurements of high-mass ions [203,209,244,245].

The binding experiments revealed that β_1AR preferentially binds PIP_2 > PIP > PIP_3 ≈ PI at all concentration points (**Figure 4.6** A). For PI, PIP, and PIP_2, each lipid differs by an additional phosphate group on the inositol moiety, which leads to an additional negative charge on the lipid head group. The order of the lipid affinity is correlated with the negative charges on the head group, suggesting the importance of the electrostatic interactions for the binding. Considering that electrostatic interactions become stronger in the gas-phase, this gas-phase effect may contribute to the difference of binding observed in the spectra. However, the influence of this effect should be minor, because PIP_3 showed lower affinity than PIP_2 and PIP. If the gas-phase effect were the main factor, PIP_3 would be expected to show the highest binding giving it has the highest negative charge. Therefore, the difference of the binding level may be elevated in the gas-phase but the order of affinity observed should be the same as in solution. In addition, the reduced affinity of PIP_3 is speculated to be due to the

steric hindrance of the bulky head group, suggesting that the interaction of PIP_2 is charge selective and structure selective. Further, analogous experiments were performed using NTSR1 and $A_{2A}R$ ($A_{2A}R^{134}$ were obtained from the Tate lab). Both receptors showed similar binding preference to the PIs as observed for β_1AR. Interestingly, PI-bound NTSR1 (**Figure 4.6** B) was not detected although PI was observed bound to the receptor in the absence of its phosphorylated derivatives (**Figure 4.4** C). The absence of the PI-bound NTSR1 suggested that PI was outcompeted by the phosphoinositides, implying that the PIs may share common binding sites. Overall, the three class A GPCRs displayed high affinity and selectivity to PIP_2. This implies that preferential binding sites for PIP_2 may exist in these receptors.

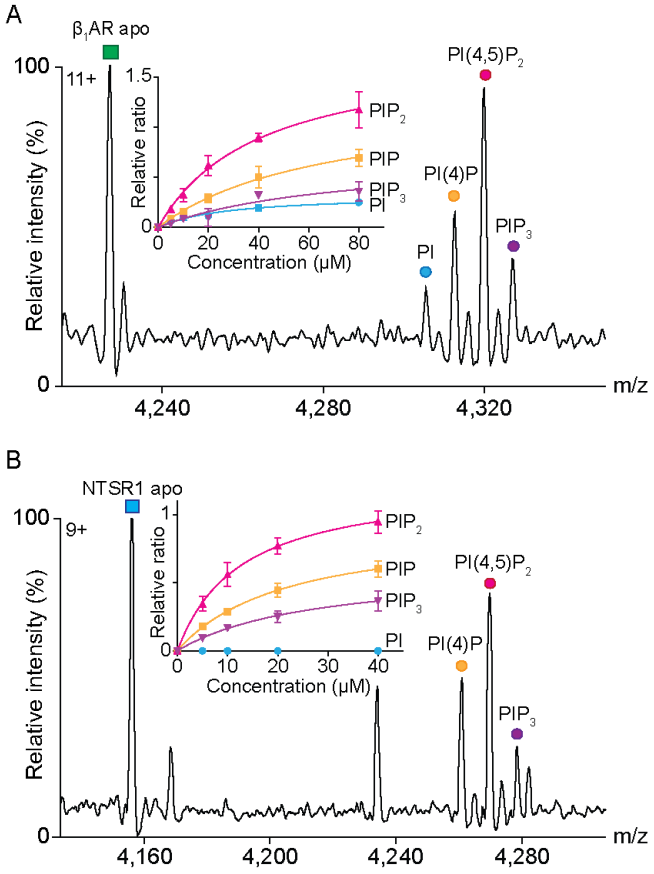

Figure 4.6 Selectivity of GPCRs to PIs. Representative spectra are shown for (A) β_1AR (β114) and (B) NTSR1 (HTGH4-ΔIC3B) incubated with an equal molar concentration of PI, PI(4)P, PI(4,5)P_2, and PI(3,4,5)P_3 at 20 μM total lipid concentration. Receptors were incubated with the lipid mixture at different concentrations (total lipid concentration was between 0 and 80 μM). Lipid binding was evaluated as an intensity ratio to the apo-receptor and plotted against the lipid concentrations (insets). For all the receptors, β_1AR, NTSR1, and A_{2A}R, the PIP_2 adduct had the highest intensity among the PIs. Figure from Yen *et al*.[3] and drawn by the author.

4.2.4. Investigating the PIP_2-binding motifs of class A GPCRs

CG simulation was applied to characterize the nature of GPCR-PIP_2 interactions (Dr. George Hedger did the simulations)[26]. The simulations were performed using NTSR1 (PDB: 4BUO) in a mixed PC/PIP_2 membrane. Ten independent simulations of 5 μs were performed. Protein-lipid contacts were defined using a bond distance cut-off of 0.6 nm. The numbers of PIP_2 contacts were counted for each protein residue. The simulations revealed multiple binding interfaces of PIP_2 formed by the intercellular loops (ICL) linking TM1, TM2, TM4, and TM7 (**Figure 4.7** A). PIP_2 was seen to interact with the basic side chains at the interfaces via the biphosphorylated inositol head group. In contrast, the acyl chains of PIP_2 participated in less specific and thus more dynamic interactions, which yielded more diffuse probability densities. This PIP_2 interaction pattern remained similar in an extended simulation of 100 μs. Analogous simulations using a PC/PS or PC/PIP_2/PS membrane revealed that PS molecules had fewer contacts with the receptor than PIP_2 and these contacts were more diffuse (**Figure 4.7** B). Additionally, PS did not compete or interfere with the PIP_2 interactions, therefore, suggesting distinct interactions between PIP_2 and other anionic lipids. A more comprehensive comparison of simulations for nine class A GPCRs revealed conservation of this PIP_2

interacting pattern at the TM1, TM4, and TM7 interfaces (**Figure 4.8**), implying a functional significance of the PIP_2 binding.

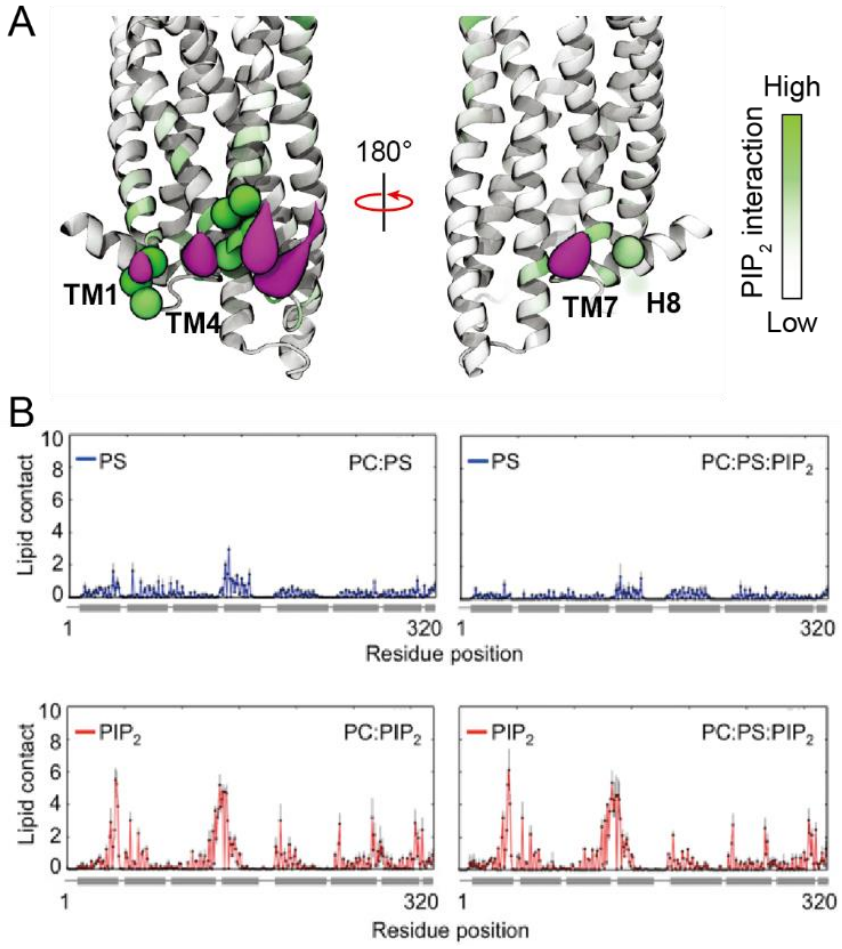

Figure 4.7 Coarse-grained molecular dynamics (CGMD) simulations of NTSR1 in a PC membrane containing PIP_2. (A) NTSR1 (TM86C-ΔIC3B) (PDB: PDB: 4BUO) was simulated in PC/PIP_2 bilayers by CGMD. Residues in the protein structure are colored by a white-green scale (low to high) to indicate the number of PIP_2 contacts, which was defined by a 6 Å cut-off in the radial distribution function of CG particles. Purple spaces represent the most occupied regions of PIP_2 CG particles in the simulations. Since the lipid tails have more diffused interactions with the receptor, the purple spaces are mainly contributed by the CG particles representing the PIP_2 head group. (B) PS (*top*) and PIP_2 (*bottom*) contacts with NTSR1 are plotted as a function of residue position for simulations performed in a PC/PS (*top left*), PC/PIP_2 (*bottom left*), PC/PIP_2/PS (*top and bottom right*) membrane. Figure from Yen *et al*.[3].

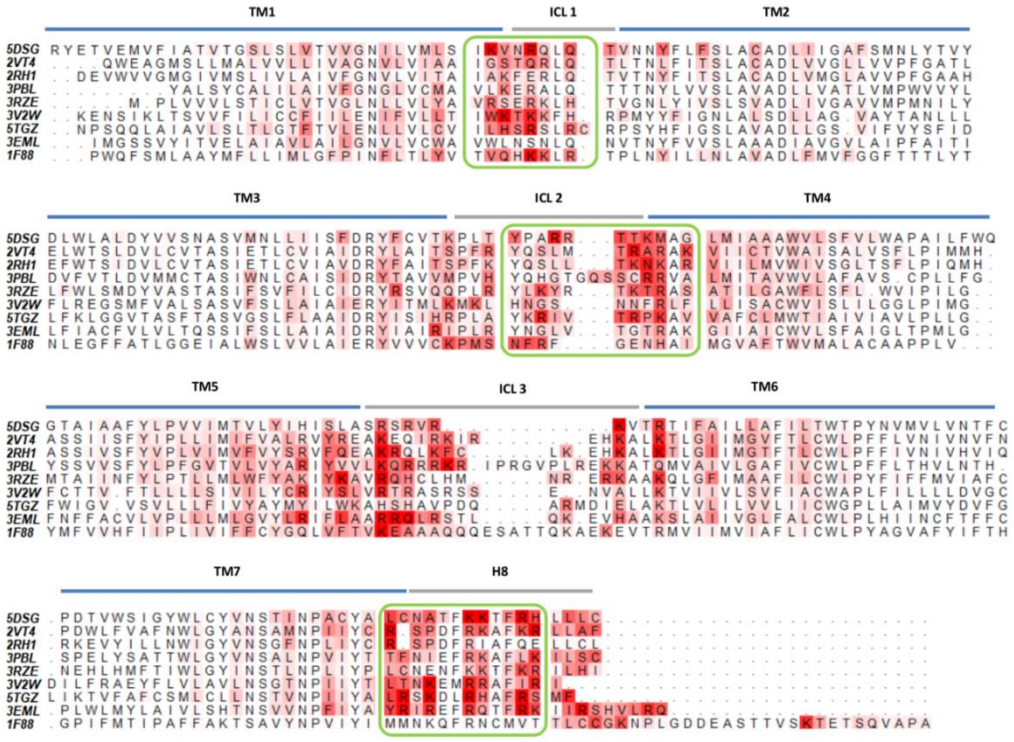

Figure 4.8 PIP_2-residue contacts in class A GPCRs observed in simulations. CGMD simulations were performed using nine class A GPCRs (5DSG = M4 muscarinic acetylcholine receptor; 2VT4 = β_1 adrenergic receptor; 2RH1 = β_2 adrenergic receptor; 3PBL = dopamine D3 receptor; 3RZE = histamine H1 receptor; 3V2W = sphingosine 1-phosphate receptor; 5TGZ = CB1 cannabinoid receptor; 3EML = adenosine A_{2A} receptor; 1F88 = rhodopsin). Sequences of the receptors are aligned with their transmembrane helices, intercellular loops (ICL), and H8 helices indicated by horizontal lines. Residues are color-coded (*red*) according to their numbers of contacts with PIP_2. The three green boxes correspond to the motifs with frequent PIP_2 contact as discussed in the text for NTSR1. Figure from Yen *et al.*[3].

To further characterize the PIP_2 interactions, a NTSR1 construct HTGH4 43–421[287,288] was used in this study (expression plasmid was obtained by the Plückthun lab). This construct provides an advantage by preserving the intracellular loop 3 (ICL3) and a more intact sequence of C-terminal sequence (**Figure 4.9** A), which are flexible motifs that were removed in NTSR1 HTGH4 ΔIC3B to facilitate crystallization but unnecessary for MS analysis[288]. Prior

investigation of the PIP_2 interaction, NTSR1 HTGH4 43–421 was expressed and characterized by native MS (**Figure 4.9** B). Interestingly, the receptor exhibited a higher PA binding level compared to NTSR1 HTGH4 ΔIC3B (**Figure 4.9** D). Inspection of the two sequences revealed a conserved polybasic motif between residues 394 and 398 (RRKKR) that was considered responsible for the increased PA binding. Deleting the RRKKR motif yielded the NTSR1 HTGH4 43–421 Δ394–398 with a reduced PA-binding level (**Figure 4.9** C), implying that this motif is a potential lipid-binding site. To avoid the possible interference of the RRKKR motif to the PIP_2-binding analysis, the NTSR1 HTGH4 43–421 Δ394–398 was used for later experiments.

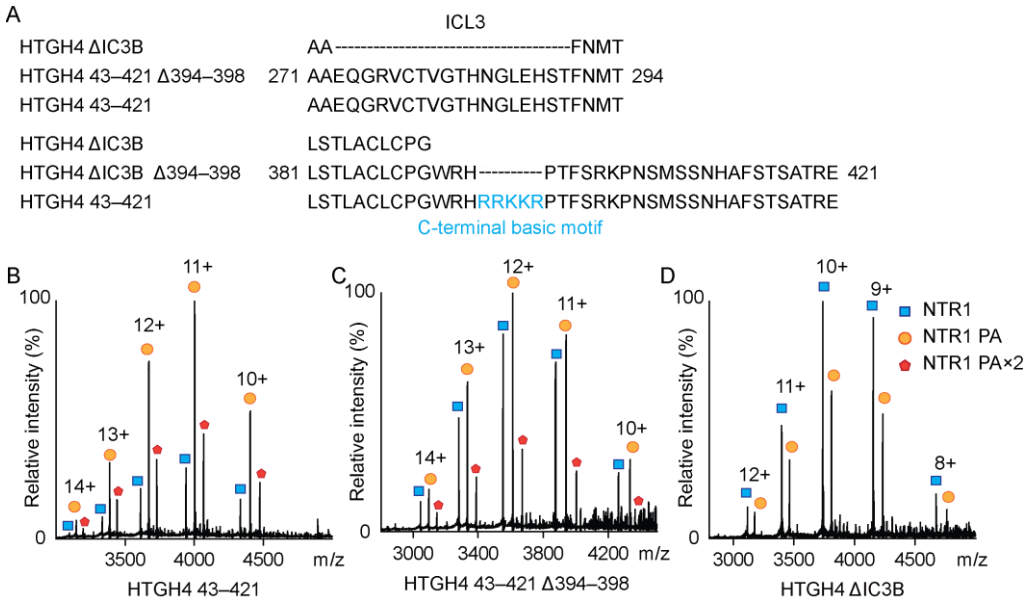

Figure 4.9 Endogenous co-purified PA of different NTSR1 constructs. (A) NTSR1 HTGH4 43–421 preserves the intercellular loop 3 (ICL3) and a longer C-terminal sequence. Polybasic residues between position 394 and 398 are colored in *blue*. Residues use the position number of the corresponding residue in the wild-type receptor. Mass spectra of purified (B) NTSR1 HTGH4 43–421, (C) NTSR1 HTGH4 43–421 Δ394–398, and (D) NTSR1 HTGH4 ΔIC3B showed a different level of PA adducts. Figure drawn by the author.

To investigate the binding preference of PIP_2 toward the binding interfaces discovered from the CGMD simulations, mutations were introduced to the NTSR1 HTGH4 43–421 Δ394–398 to disrupt the PIP_2 interactions at these interfaces (Dr. H. Yen designed the primers; the author did the mutagenesis, protein expression, MS measurement and data analysis). The changes in the PIP_2-binding level will indicate the contribution of these interfaces to the PIP_2 binding of the receptor. Three mutants were designed to contain mutations at the TM1 motif (R43, K44, and K45), the TM4 motif (R135, R137, K139, and K140), and the TM7 motif (R311), which were the amino acid residues that showed the most frequent contact with PIP_2 from the simulations. Because most GPCRs are intrinsically unstable, mutations might perturb the structure, stability, and expression level of the receptor. Previous study has exhaustively evaluated the expression level and activity of a library of NTSR1 HTGH4 43–421 mutants (i.e., an ensemble of "all-versus-all mutations"), which included mutants that comprise all 64 codons for each amino acid position (this information was obtained from the Plückthun lab)[296]. Guided by this information, the Lys/Arg residues were mutated to the non-basic amino acids that give the highest expression level and activity (TM1: R43G, K44G and K45G; TM4: R135I, R137T, K139L and K140L; and TM7-H8: R311N). Because Lys/Arg are heavier amino acids than the targeted residues (i.e., Gly, Ile, Thr, Leu, and Asn), the mutations created "light" receptor mutants in contrast to the "heavy" unmodified receptor.

The effect of the mutations was evaluated by comparing the PIP_2-binding level of the "light" and "heavy" receptors at the same PIP_2 concentration. The

receptors were incubated with a low concentration of PIP_2 to ensure that the receptors formed adduct with maximum of one PIP_2. Therefore, the receptors existed in either the apo form or single PIP_2-bound adduct. The intensity of the PIP_2 adduct (relative to the apo) related the sum of the PIP_2 binding at all binding sites. The intensity difference between the PIP_2 adduct of the "light" and "heavy" receptors reported the contribution of the mutation sites to the total PIP_2 binding.

Although this method allowed a straightforward comparison for the "light" and "heavy" receptors, it struggled to obtain reproducible measurements. Because this method used a low PIP_2 concentration, the PIP_2 adduct had a low spectral intensity and low signal-to-noise ratio. The low-intensity peak was more sensitive to the spectral baseline, ionization efficiency, and activation energy, which influence the measurement of the peak intensity. Variation of these factors between different MS analysis hindered precise measurements of the PIP_2 adduct peak intensity.

To address this challenge, a multiplexing strategy was applied[297]. The "light" mutants were mixed with an equal molar ratio of "heavy" receptor prior to the PIP_2 incubation and MS analysis (**Figure 4.10** A). Because the mass difference between the "light" and "heavy" receptor was > 42 Da (roughly 4 *m*/*z* unit in the mass spectra), the high-resolution Orbitrap mass spectrometer afforded sufficient resolution to resolve the "light" and "heavy" receptors. For each "light" mutant, its PIP_2-binding level was estimated as its PIP_2 adduct/apo ratio relative to that of the "heavy" receptor (**Figure 4.10** B). The multiplexing strategy took

advantage of the resolution of MS and the mass difference between the “light” and “heavy” receptors. By analyzing the “light” and “heavy” receptors in one measurement, both receptors experienced the same measurement condition and thus reduced the bias in the measurements.

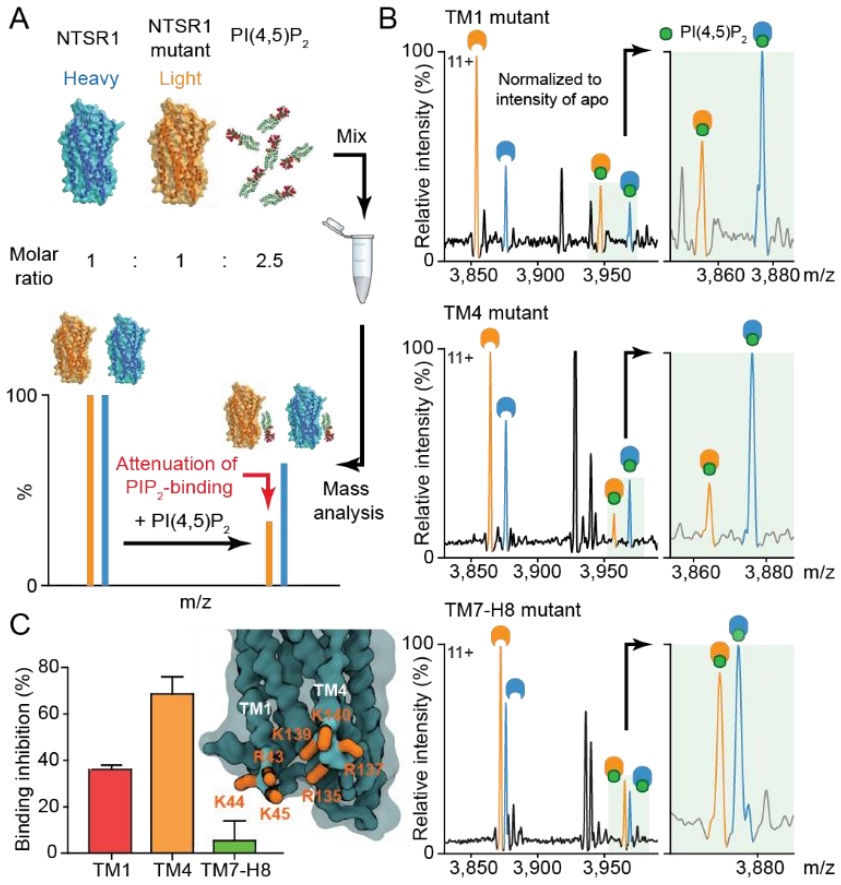

Figure 4.10 The preferential PIP_2-binding sites in the NTSR1. (A) Site-directed mutagenesis generated mutated “light” mutants of NTSR1 (TM1: R43G, K44G and K45G; TM4: R135I, R137T, K139L and K140L; and TM7-H8: R311N; residue position refers to the corresponding position number in the wild-type receptor) that are resolvable from the unmodified “heavy” receptor in the mass spectra. The unmodified “heavy” and mutated “light” receptors were mixed at equal molar ratio and incubated with a low concentration of PIP_2 for MS analysis, allowing a direct readout of the effect of the mutations to the corresponding binding interfaces. (B) The PIP_2-binding level of each mutant was determined as the PIP_2 adduct:apo ratio to that of the

unmodified “heavy” receptor. (C) The effects of the mutations were evaluated as the attenuation of the PIP_2-binding level in the mutants, from which the TM4 interface is shown to be a preferential binding site over the TM1 and TM7-H8 interfaces. The inset shows the structure of NTSR1, with the mutated basic residues in TM1 and TM4 highlighted in *orange*. Figure from Yen *et al*.[3] and drawn by the author.

The three NTSR1 mutants showed different levels of attenuation in PIP_2 binding compared to the unmodified receptor. The most significant attenuation was observed for the mutant with mutations on the TM4, which showed a 70 ± 0.13% decrease (**Figure 4.10** C). For the TM1 mutant, PIP_2 binding diminished by 35 ± 0.03%. In contrast, the mutation in TM7 mutant had marginal effect on PIP_2 binding. Despite multiple mutations being introduced in the mutants, PIP_2 binding could not be completely inhibited. This implied that more residues may be involved in the binding or multiple PIP_2-binding sites may exist in the receptor, which was consistent with the results as predicted by CGMD simulations. Overall, the mutagenesis results indicated that the major PIP_2-binding hotspots of NTSR1 are on the cytoplasmic faces of the TM1 and TM4.

Interestingly, the unmodified and mutated NTST1 have different ionization efficiency. This phenomenon was also observed for the Parkinson’s-associated protein, DJ-1[297]. Both D149A and A104T mutations caused different ionization efficiency comparing to the unmodified protein. While D149 is a surface residue, A104 is a residue in the hydrophobic core of the protein. Therefore, the difference of ionization efficiency may not due to changes of the protein surface charge. It is possible that ionization efficiency is sensitive to the subtle

difference of the protein conformation caused by the mutation. Further studies are necessary to understand the exact mechanism.

4.2.5. PIP_2 interactions enhance GPCR-G protein coupling

Considering the selectivity of GPCRs to PIP_2 and the conservation and location of the PIP_2 hotspots in class A GPCRs, PIP_2 interactions are speculated with modulatory effect on transducer coupling of GPCRs. To examine the functional effect of PIP_2 on G protein-coupling[117], the GTPase activity of a heterotrimeric G protein ($G\alpha_i\beta\gamma$) was measured in the presence of active NTSR1, stimulated by neurotensin$_{8–13}$, and PIP_2 (Dr. H. Yen did this work; $G\alpha_i\beta\gamma$ was from the Plückthun lab). The GTP hydrolysis by the $G\alpha_i\beta\gamma$ was enhanced by 1.3-fold in the presence of PIP_2 (**Figure 4.11**), implying the positive modulatory effect of PIP_2 on G protein coupling.

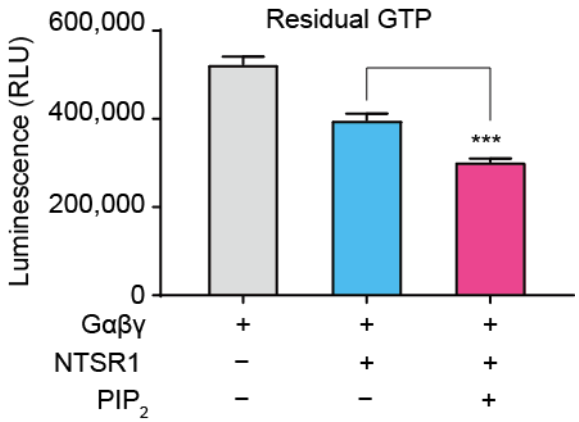

Figure 4.11 The effect of PIP_2 to the GTPase activity of the G protein. The GTPase activity of the $G\alpha_i\beta\gamma$ was measured using three different systems: $G\alpha_i\beta\gamma$ (*grey*), NTSR1 (HTGH4 ΔIC3B)-$G\alpha_i\beta\gamma$ (*blue*), and NTSR1-$G\alpha_i\beta\gamma$ in the presence of PIP_2 (*magenta*). A T-test revealed a statically significant increase ($p < 0.001$ ***, $n = 3$) of the GTP hydrolysis in the presence of PIP_2. Figure from Yen *et al*.[3].

PIP_2 interactions may facilitate receptor activation and thus G protein coupling. To test this hypothesis, we compared the PIP_2-binding level between the active agonist-bound and inactive ligand-free β_1AR. PIP_2 was incubated with the β_1AR, co-purified with the agonist isoprenaline, and the β_1AR E130W variant[298], which is engineered to increase the functional expression of the receptors in the ligand-free state[298] (β_1AR (β118)[281,282] was obtained from the Tate lab; Dr. H. Yen prepared the sample and acquired the spectra; the author analyzed the data). The spectral results revealed a 30.9 ± 1.4% increase in PIP_2 binding to the β_1AR in the agonist-bound state over the ligand-free state (**Figure 4.12**), suggesting that PIP_2 may stabilize GPCRs in the active conformation. This finding coincided with the previously observed effect of anionic phospholipids for β_2AR[54]. Anionic lipids (PG, PS, PI), in contrast to zwitterionic lipids (PC and PE), were more effective in increasing the agonist affinity of the β_2AR. Anionic lipids were also shown to induce the outward movement of the cytoplasmic end of TM6, which is the most profound conformational change upon receptor activation and its engagement with G proteins. Equally important, a recently available structure of the kappa opioid receptor revealed that receptor activation involves not only a substantial movement of the TM5 and TM6, but also the ICL2[299], which is the PIP_2-binding hotspot identified in the present study. This implies that PIP_2 may stabilize GPCRs in active conformation through interactions with the ICL2.

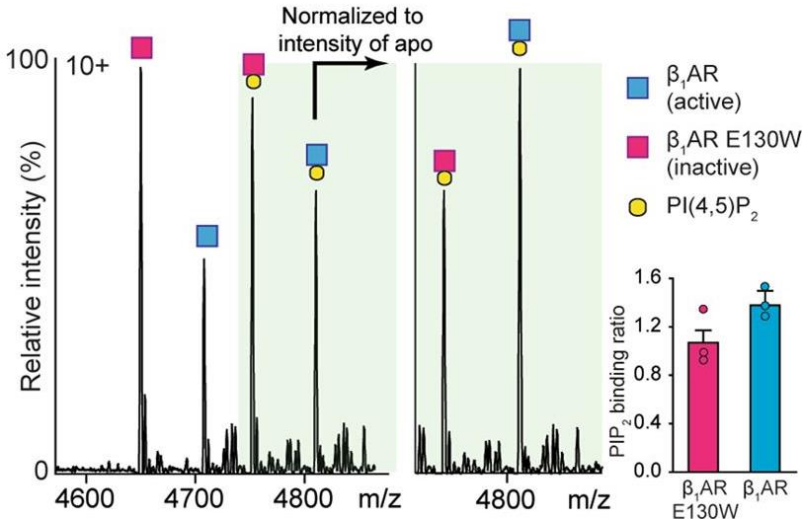

Figure 4.12 Comparison of the PIP_2 affinity between the active agonist-bound and inactive ligand-free β_1AR. The mass spectrum was recorded from an equal molar mixture of the inactive ligand-free β_1AR (β114-E130W) and active β_1AR (β118, co-purified with agonist isoprenaline) incubated with PIP_2. PIP_2 binding was observed for both receptors. After normalization to account for difference in ionization efficiency, the result revealed a higher PIP_2 binding to the active state. The bar chart shows the intensity ratio of the PIP_2 adducts to the corresponding apo-receptor. Although the increase was not statistically significant, more binding was observed for the active state in all three replicates. Figure from Yen *et al*.[3].

To investigate the effect of PIP_2 on G protein-coupling, a method was developed to preserve the β_1AR-mini-Gα_s complex in the gas-phase (mini-Gα_s was obtained from the Tate lab; Dr. H. Yen did this work). Mini-G proteins are engineered Gα subunits that mimic the interactions at the receptor-G protein interface. Mini-G coupling can recapitulate the same effects of the heterotrimeric G protein coupling on increasing the agonist affinity of the receptor and stabilizing the receptor in the fully active conformation[134]. Mini-G proteins thus render an appropriate system to investigate the effect of PIP_2 on GPCR-G protein coupling. Using this method, the agonist-bound β_1AR and mini-Gα_s were incubated and studied by MS. The spectra presented two distinct

charge envelopes for the β_1AR and β_1AR-mini-Gα_s complexes (**Figure 4.13** A). The large population of β_1AR-mini-Gα_s complexes observed in the mass spectra highlights the viability of native MS in characterizing GPCR-G protein coupling.

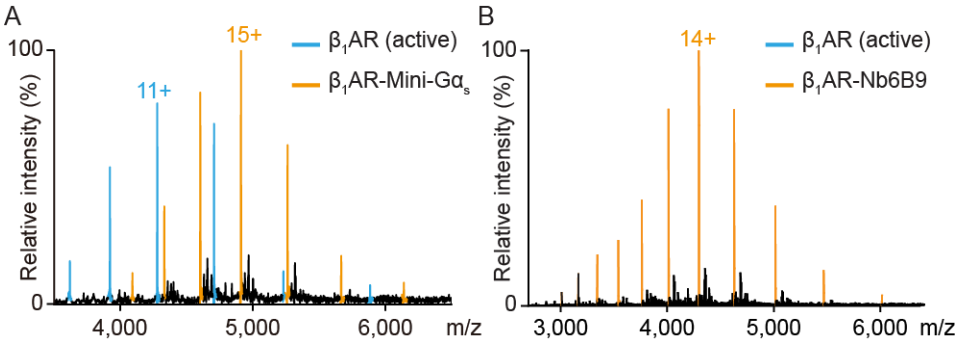

Figure 4.13 The representative mass spectra of the β_1AR-mini-G and β_1AR-Nb complexes in the gas-phase. Native MS is able to capture the β_1AR (β118), stimulated by isoprenaline, complexed with (A) mini-Gα_s and (B) Nb6B9. Signals of the receptor monomer and receptor-transducer complexes are highlighted in *blue* and *orange*, respectively. Receptor monomer is absent in (B). All β1AR form complexes with Nb6B9, indicating a high affinity of the receptor to the nanobody. As mini-Gα_s and Nb6B9 are soluble protein, they did not survived at high CID energies and were not observed in the spectra. The most abundant charge states of each species are labelled. The spectra are from Yen *et al*.[3]. Figure drawn by the author.

The agonist-bound β_1AR was incubated with mini-Gα_s in the presence of PIP_2 or PS and analyzed by native MS (Dr. H. Yen did this work). The spectra showed two charge envelopes corresponding to the β_1AR monomer (**Figure 4.14** A *upper* spectrum) and the β_1AR-mini-Gα_s complex (*lower* spectrum). Both receptor and receptor-mini-G complex showed binding to 0 to 3 PIP_2 adducts (color code *blue*, *cyan*, *orange*, and *red*). For each PIP_2 adduct, the effect of PIP_2 on receptor coupling was evaluated as the intensity ratio of the PIP_2 adduct to the corresponding apo state of the uncoupled/coupled receptor

(**Figure 4.15** D). The extent of complex formation was similar when in the absence and presence of one PIP_2. Complexes were formed in a higher extent in the presence of two (2.7-fold) and three PIP_2 (4.5-fold). This phenomenon was consistent with the results of a time course experiment (**Figure 4.14** C). The coupling efficiency is defined as the intensity ratio of the coupled receptor to the total receptor at the appropriate PIP_2-bound state. Mini-$G\alpha_s$ coupling was increased by 21 ± 6 % and 33 ± 5 % in the presence of two or three bound PIP_2, respectively. In contrast, analogous experiments performed with PS showed a minor increase of mini-$G\alpha_s$ coupling, highlighting the specific effect of PIP_2 (**Figure 4.14** B).

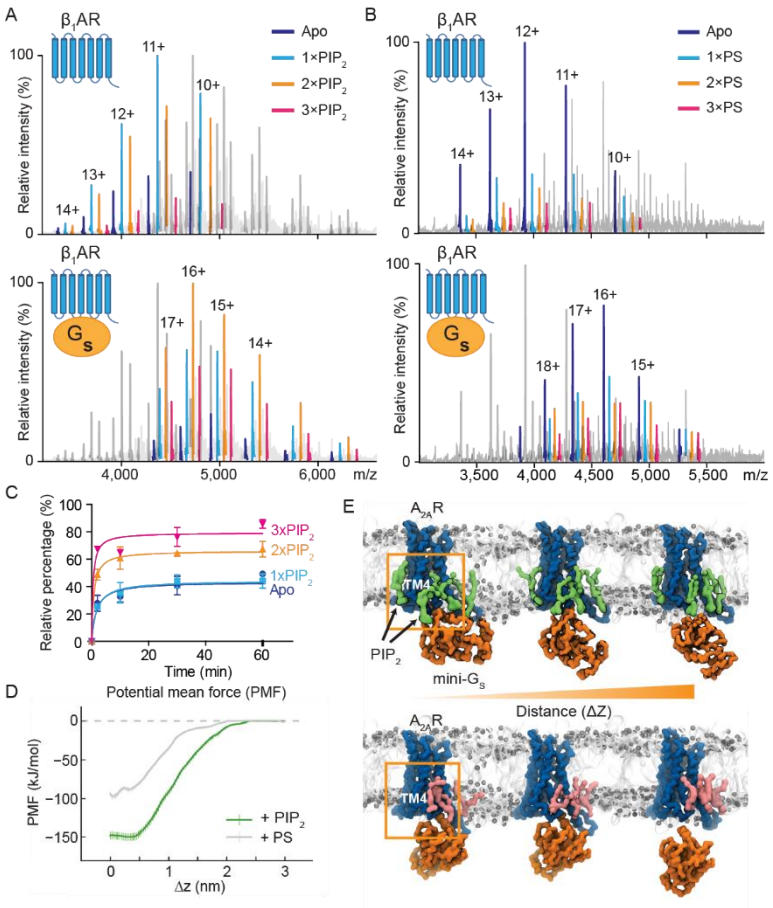

Figure 4.14 The effect of PIP_2 and PS on β_1AR-mini-Gαs coupling. β_1AR was incubated with

mini-$G\alpha_s$ in the presence of (A) PIP_2 or (B) PS for 10 min. The spectrum shows two charge envelopes corresponding to the β_1AR monomer (*upper*) and β_1AR-mini-$G\alpha_s$ complex (*lower*) with 0 to 3 bound lipids (color coded with *blue*, *cyan*, *orange*, and *red*). Complex formation was enhanced in the presence of two or three PIP_2 (**Figure 4.15** D). Same observation was also seen for (C) different incubation time. PS interactions had minor effect on the mini-G coupling. The experiments used 3-fold higher concentration of PS than PIP_2 so thus to compensate the lower affinity of β_1AR to PS. The relative percentage was defined as the spectral intensity ratio of the coupled receptor to the total receptor (receptors in both the coupled and uncoupled states) at the appropriate PIP_2 bound state. (D) Potential of mean force (PMF) of A_{2A}R-mini-$G\alpha_s$ interactions in the presence of PIP_2 (*green*) or PS (*grey*) was determined in silico and plotted as a function of the center-center separation (Δz) between the A_{2A}R and mini-$G\alpha_s$. A_{2A}R-mini-$G\alpha_s$ complex was ~50 kJ/mol more stable when interacting with PIP_2 than PS. (E) Snapshots in the PMF calculations reveals different binding behavior (*orange boxes*) of PIP_2 (*green*) and PS (*pink*). Figure from Yen *et al*.[3].

The different effects on β_1AR-mini-$G\alpha_s$ coupling between PIP_2 and PS implied that β_1AR-mini-$G\alpha_s$ complexes are stabilized by PIP_2, but not PS. To investigate this lipid-mediated effect on the free energy landscape of G protein-coupling, CGMD simulations were performed using the A_{2A}R-mini-$G\alpha_s$ complex, the only available structure of a receptor-mini-$G\alpha_s$ complex[136], in the presence of PIP_2 or PS (Wanling Song did the simulations). PMF was calculated for the A_{2A}R-mini-$G\alpha_s$ complex as a function of the distance between the A_{2A}R and mini-$G\alpha_s$ (**Figure 4.14** D). The PMF calculations revealed that more energy (~50 kJ/mol) was required to dissociate the mini-$G\alpha_s$ from the A_{2A}R interacting with PIP_2 than PS, indicating that PIP_2 stabilizes the A_{2A}R-mini-$G\alpha_s$ complex.

Notably, when mini-$G\alpha_s$ was pulled away from A_{2A}R in simulations, PIP_2 molecules occupied the interface between the receptor and mini-$G\alpha_s$ (**Figure**

4.14 E). PIP_2 appeared to interact with both the receptor and mini-$G\alpha_s$ through its polyanionic head group to stabilize the complex. In contrast, singly-charged PS didn't exhibit the same behavior (**Figure 4.14** E). These observations are similar to those of the Kir2.2 channel (**Figure 1.5**)[73]. PIP_2 was found tethering the transmembrane domain (TMD) and cytoplasmic domain (CTD) of the channel via interactions between its polyanionic head group and multiple basic sidechains from both the TMD and CTD[73]. Such interactions and effects were not replicable by singly-charged PA, which engages only the TMD region. The similar observations on the Kir channel and receptor-G protein complex imply that PIP_2 may stabilize the complexes through bridging interactions.

According to the results of the CGMD simulations, PIP_2 bound to $A_{2A}R$-mini-$G\alpha_s$ via interactions with both the $A_{2A}R$ and mini-$G\alpha_s$. PIP_2 contacts on the $A_{2A}R$ were the same as those discussed in **Section 4.2.4**. For the mini-$G\alpha_s$, PIP_2 interacted with the basic residues proximal to the PIP_2 contacts on the TM3, TM4, and TM5 interfaces of the $A_{2A}R$ (**Figure 4.15** A). To interrogate the functional significance of these interactions, a nanobody (Nb6B9[132,285,300]) was recruited (Nb6B9 from the Tate lab; Dr. Idlir Liko and Dr. H. Yen prepared the samples; the author acquired the spectra and analyzed the data). The Nb6B9 is a G protein surrogate. The conformations of receptors coupled to the G protein and nanobody are virtually identical[301]. However, the PIP_2 interaction sites on the mini-$G\alpha_s$ do not exist in the nanobody (**Figure 4.15** A)[132]. At an equal molar ratio of Nb6B9, agonist-bound β_1AR coupled to Nb6B9 at 100% efficiency and no β_1AR monomer was observed (**Figure 4.13** B). To compare the PIP_2-binding level between β_1AR in the monomer and complex form, β_1AR was incubated

with 0.3 molar equivalents of Nb6B9 in the presence of PIP_2. The β_1AR and β_1AR-Nb6B9 complex showed a similar degree of PIP_2 binding (**Figure 4.15** B and D), implying that the interactions between the β_1AR and Nb6B9 are insensitive to the presence of PIP_2. The different responses of β_1AR-mini-Gα_s and -Nb6B9 complexes to PIP_2 indicated that PIP_2 stabilizes β_1AR-mini-Gα_s complexes via bridging interactions specific to the receptor and mini-Gα_s.

To further explore the effect of PIP_2 on the coupling of other G proteins, analogous experiments were performed using mini-G$\alpha_{i(s)}$, which is a chimera mini-Gα_s with residues transferred from the Gα_i (the mini-G$\alpha_{i(s)}$ was obtained from the Tate lab; Dr. Idlir Liko and Dr. H. Yen prepared the samples; the author acquired the spectra and analyzed the data). Coupling of mini-G$\alpha_{i(s)}$ to β_1AR was enhanced in the presence of PIP_2 but at a lower extent compared to that of mini-Gα_s (**Figure 4.15** C and D). Given that the PIP_2 interacting residues (T40, H41, R42, K216, R380) of mini-Gα_s identified in the simulations were substituted in the mini-G$\alpha_{i(s)}$ (E40, V41, K42, D216, T380), the reduced effect is attributed to the alternation of the PIP_2-binding sites in the mini-G$\alpha_{i(s)}$. Importantly, the different results observed for the mini-Gα_s and mini-G$\alpha_{i(s)}$ implied that PIP_2 discriminates different G proteins and stabilizes the corresponding complexes to a different extent. Since the effect of PIP_2 on GPCR-G coupling is G protein-dependent, our discovery sheds light on the development of lipid-like molecules that bias the selectivity of different G protein pathways.

Ligands of different efficacy have been shown to regulate GPCRs by changing their conformations, the rate of GPCR-G protein coupling, and the efficiency of nucleotide exchange that leads to G protein activation[128,135,282]. Our results demonstrated the effect of PIP_2 on the conformation of the receptor and GPCR-mini-G coupling. Since PIP_2 directly interacts with the mini-G when bound to receptor-mini-G complexes, it is interesting to see if PIP_2 affects the efficiency of nucleotide exchange and G protein activation as observed for ligands. While mini-$G\alpha_s$ does not have a GTP/GDP binding site, future experiments that use functional G proteins may provide answers to this question.

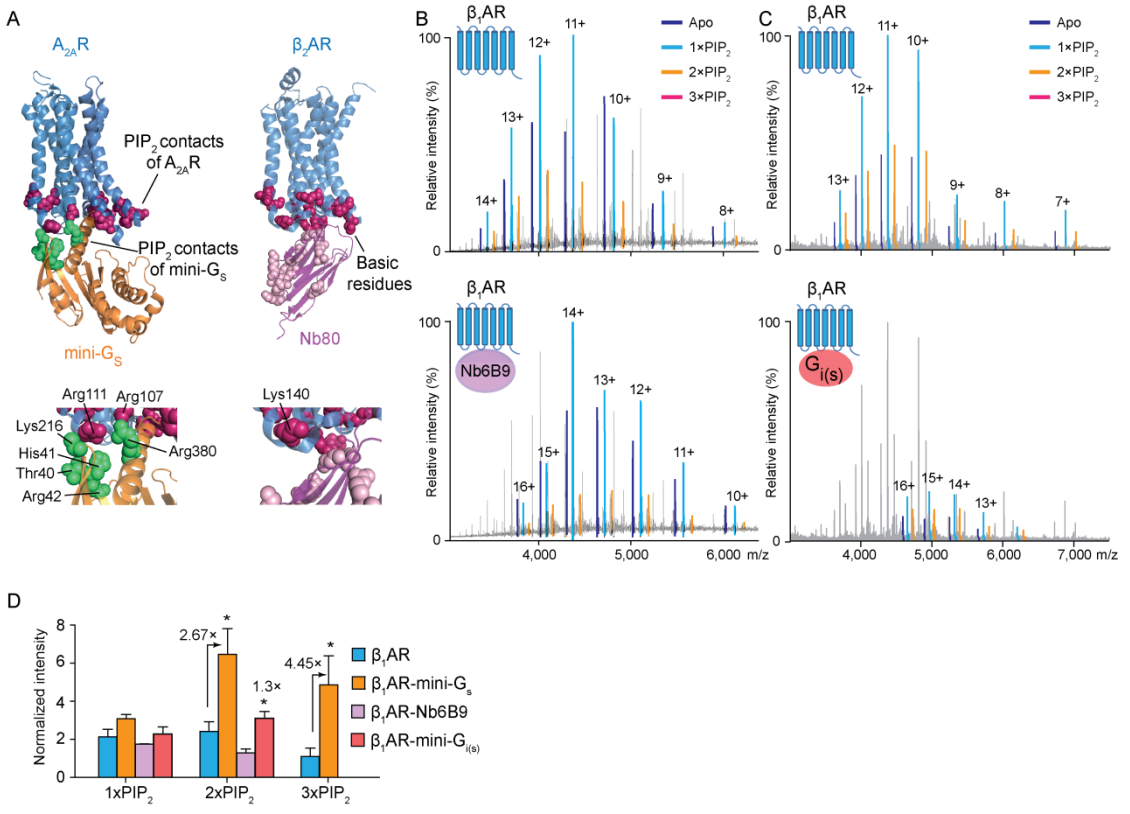

Figure 4.15 The effect of PIP_2 on the complex formation of β_1AR with mini-G and nanobody. (A) The CGMD simulations of the A_{2A}R-mini-$G\alpha_s$ in a PIP_2 containing membrane revealed the PIP_2 contacts on the A_{2A}R (*space filled magenta*) and mini-$G\alpha_s$ (Thr40, His41, Arg42, Lys216, and Arg 380; *green*). The structure of β_2AR-Nb80 is shown for comparison and the basic residues of

the Nb80 are shown as *space filled purple*. (B) and (C) are the mass spectra of β_1AR incubated with Nb6B9 (0.3 molar equivalents) and mini-G$\alpha_{i(s)}$ in the presence of PIP_2, respectively. (D) The bar chart shows the intensity ratio of the lipid-bound states to the apo form of the receptor in the uncoupled/coupled state. One-way ANOVA with Dunnett's multiple comparison test was performed, and statistically significant difference is denoted by asterisks ($p < 0.05$ *, $n = 3$). Figure from Yen *et al*.[3].

4.3. Conclusion

This study used state-of-the-art native mass spectrometry in combination with mutagenesis and CGMD simulations to explore the selectivity, binding sites, and binding patterns of lipids for class A GPCRs, to give the molecular insights into the functional significance of lipids on GPCR-G protein coupling. The MS results demonstrated that PIP_2 binding to β_1AR is conformation-dependent, implying the allosteric effect of PIP_2 in stabilizing GPCRs in an active conformation. Further investigation revealed that PIP_2 is able to enhance the coupling efficiency between GPCR and Gα subunits (e.g., β_1AR and $G\alpha_s$, NTSR1 and $G\alpha_i$) via bridging interactions to both the receptor and G protein. PIP_2 is, therefore, proposed here as a positive allosteric modulator that binds to the cytoplasmic face of class A GPCRs and enhances G protein coupling. This novel PIP_2-binding site revealed by the mutagenesis study illuminates a new lipid-mediated signal transduction pathway for GPCRs and may act as an intracellular drug-binding motif in the design of GPCR pharmaceuticals.

Chapter 5. Conclusion

To obtain a complete picture of the structure and function of membrane proteins, it is necessary to understand the coordination between the membrane proteins and their membrane environments. With a limited number of detection methods, membrane protein-lipid interactions have proved challenging to capture due to their dynamic and versatile interaction modes. Following the successful preservation of an intact membrane protein complex using native mass spectrometry (MS), it was demonstrated that protein-lipid complexes may be preserved in the gas-phase as well. Since then, native MS provides a high-speed, high-sensitivity and high-resolution analytical tool to study protein-lipid complexes, revealing many important insights about membrane protein-lipid interactions.

While detergents are certainly useful and compatible membrane mimetics for native MS, alternative mimetics, such as nanodiscs, offer many attractive advantages. One of these advantages is a consistent lipid bilayer structure, which faithfully recapitulates the structure of biological membranes while also providing the ability to tailor the bilayer composition. Nanodiscs may be formed by a mixture of lipids with a precisely controlled composition, although this causes the heterogeneity and polydispersity of the system. To address these challenges, two novel methods, autocorrelation and dual Fourier transform methods, were developed, allowing determination of the average lipid mass of heterogeneous-lipid nanodiscs. Although both methods are highly accurate and precise, the dual Fourier transform is preferred due to a minimal requirement of data processing and assumptions. Based on the dual Fourier method, a

framework was established to analyze heterogeneous-lipid nanodiscs and their lipid composition.

Transitioning from “empty” nanodiscs to membrane protein nanodiscs, the mass spectra became more complex and difficult to interpret. By adapting the previously established deconvolution algorithm (UniDec) and designing a novel mass defect analysis, it was possible to assign the stoichiometries of the protein subunits and lipid components for each peak. Surprisingly, a large number of bound lipids were observed and multiple distributions of bound lipids were identified. Comparison with molecular dynamics (MD) simulations revealed that these distributions correspond to the different populations of the annular lipids that involving in different interactions. As annular lipid interactions are relatively short-lived compared to non-annular lipid interactions, they are generally harder to capture and less understood. These results demonstrate the invaluable potential of using this nanodisc-based MS method to study annular lipid interactions. In combination with the dual Fourier transform, monitoring the change in lipid composition as a function of different sub-populations of annular lipids may be feasible.

Additionally, notable differences in the dissociation pathways of “empty” and membrane containing nanodiscs were observed. While “empty” nanodiscs dissociate by ejecting their lipids, membrane protein nanodisc dissociation starts from the exterior by losing MSPs, followed by delipidation and oligomer dissociation upon increasing collisional activation. Although the exact mechanism remains unclear, the results presented in this thesis indicate a

dependency on the instrumental polarity, chemistry of the lipids, embedded membrane protein of the nanodisc, and the relative strength between protein-protein and protein-lipid interactions. A better understanding of the dissociation mechanism may allow for a better control of the liberation of the membrane protein from nanodiscs and the preservation of the annulus.

The framework developed here allows analysis of nanodiscs by native MS and sets ground for more complex nanodisc systems. UniDec allows interpretation of the spectra. Mass defect analysis and dual Fourier transform method provide information of the protein and lipid composition. However, additional tools may be required to analyze systems that are more complex. For example, if the membrane protein is hetero-oligomeric, combinations of different subunits may share the same mass defect. Mass defect analysis cannot disambiguate the protein composition of the nanodisc complex. Mutagenesis and tandem MS are suitable tools in this case. Mutagenesis can be used to change the mass of the protein subunit and therefore the mass defect of the nanodisc complex. Using tandem MS, complexes are isolated and dissociated into different sub-complexes, from which the composition and stoichiometry of the subunits may be determined. Furthermore, the dual Fourier transform method was tested on nanodiscs containing lipids with similar masses. Further developments may be required to adapt this method to nanodiscs containing lipids with significantly different masses.

As well as nanodisc-based methods, this thesis develops detergent-based methods to study G protein-coupled receptors (GPCRs), which are important

but challenging targets in health and disease. Recently developed detergent-based native MS has enabled the analysis of Class A GPCRs in the gas-phase. Here, this method was combined with MD simulations, mutagenesis, multiplexing strategy, and *in vitro* assays to obtain a comprehensive view of the role of lipids in class A GPCR-G protein signal transduction. This study demonstrated that class A GPCRs preferentially interact with phosphatidylinositol bisphosphate (PIP_2), which predominately binds at the cytoplasmic interface of TM4. This interaction was proposed to stabilize the receptor in its active conformation and bridge the receptor with specific G proteins, thus enhancing the efficiency of G protein coupling. In light of these findings, the novel PIP_2-binding site and the regulatory mechanism may illuminate new directions of rational drug design. Furthermore, successfully preserving the ternary complex of lipid-GPCR-transducer demonstrates the remarkable ability of native MS to study non-covalent interactions in a label-free manner.

Beside lipid interactions, native MS may provide a promising technique to interrogate GPCR-transducer coupling. GPCR signaling involves interactions of many different proteins/ligands/lipids. While it is a challenge for many techniques to study complex mixtures/systems, MS provides an alternative solution. The high resolution of MS allows us to monitor different complexes in one spectrum as shown in **Chapter 4**. The correlation between the G protein- and arrestin-coupling remains poor understood. Using MS to investigate GPCR coupling in the presence of lipids, G proteins and arrestins may be feasible. Studying the relative population of the G protein-coupled receptor and arrestin-

coupled receptor may help us understanding the selectivity of these pathways as well as the effects of lipids. In addition, G proteins anchor on the cell membrane through lipidation, which is important to their coupling with GPCRs. Future native MS studies using nanodiscs may provide additional information to how membrane properties affect GPCR signaling.

In addition to detergents and nanodiscs, SMALP also presents a promising mimetic for native MS and protein-lipid studies. One of the major challenges in studying protein-lipid interactions is to identify important functional lipids. This requires both preservation and identification of endogenous lipids. MS is useful to identify lipids directly bind to membrane proteins, such as non-annular lipids. However, currently available detergent-based and nanodisc-based methods are not completely detergent-free. Important lipids could have been removed during detergent extraction. SMALP, a completely detergent-free method, sheds light on preserving the native lipid environment of membrane proteins. SMA directly extracts and solubilizes membrane proteins and their lipid environments from the membrane. Therefore, combination of MS and SMALP may become a powerful technique to identify lipid ligands, which is urgently needed in the field.

Although SMALP-based MS method is attractive, two fundamental questions need to be addressed. First, we need to know if SMALP is a compatible membrane mimetic for native MS. SMALPs might have different gas-phase behavior comparing to detergent micelles and nanodiscs. It is unknown whether SMALPs can be effectively ionized with ESI and whether protein-lipid complexes can be liberated from SMALPs. Second, SMALPs contain a

heterogeneous mixture of lipids. This may lead to spectra more complex than those of nanodiscs. In this case, spectral resolution may become demanding. Further, as spectra become more complex, new deconvolution methods may be required to interpret the results.

A particularly critical instrumental advance in the studies here is the use of a high-resolution Orbitrap mass spectrometer. The increased resolution improves the capacity of MS in handling more complex and heterogeneous samples. Without this high-resolution, nanodisc ions of different lipid adduct states may overlap extensively in the mass spectra. It would be impossible to have a precise and accurate measurement of the lipid composition of the nanodiscs and less information could be extracted from the mass spectra. In addition, a multiplexing strategy was used to quantitatively compare different non-covalent interactions. This strategy relies heavily on the resolution of the instrument to separate analogue complexes with small mass differences.

In summary, native MS has enabled the study of protein-lipid interactions, including both annular and non-annular lipids. In combination with other techniques and methodologies, a diverse array of information may be derived from the acquired mass spectra. This has enabled us to characterize the nature of protein-lipid interactions and their contribution to the function of membrane proteins. With the consistent improvement in instrumentation and methodology, native MS is expected to be applied to increasingly challenging protein systems. By providing new insight into the modulation of membrane protein structure and

function, native MS will therefore remain a leading technique at the frontier of protein-lipid interactions.

Reference

1 Hoi, K. K., Robinson, C. V. & Marty, M. T. Unraveling the Composition and Behavior of Heterogeneous Lipid Nanodiscs by Mass Spectrometry. *Anal. Chem.* **88**, 6199-6204, (2016).

2 Marty, M. T., Hoi, K. K., Gault, J. & Robinson, C. V. Probing the Lipid Annular Belt by Gas-Phase Dissociation of Membrane Proteins in Nanodiscs. *Angew. Chem. Int. Ed.* **55**, 550-554, (2016).

3 Yen, H.-Y. *et al.* PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, 423-427, (2018).

4 Lodish, H. F. *et al.* in *Molecular cell biology* Ch. 7, (W.H. Freeman-Macmillan Learning, 2016).

5 Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567-580, (2001).

6 Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, (2015).

7 Wallin, E. & Heijne, G. V. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**, 1029-1038, (1998).

8 Reece, J. B. & Campbell, N. A. in *Campbell Biology* Ch. 7, (Pearson Education, 2014).

9 Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19-34, (2017).

10 de Armas-Rillo, L., Valera, M.-S., Marrero-Hernández, S. & Valenzuela-Fernández, A. Membrane dynamics associated with viral infection. *Rev. Med. Virol.* **26**, 146-160, (2016).

11 Hopkins, A. L. & Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727-730, (2002).

12 Watson, H. Biological membranes. *Essays Biochem.* **59**, 43-69, (2015).

13 von Heijne, G. Membrane-protein topology. *Nat. Rev. Mol. Cell Biol.* **7**, 909-918, (2006).

14 Reichelt, R. *et al.* Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *The Journal of Cell Biology* **110**, 883, (1990).

15 Spillane, J., Kullmann, D. M. & Hanna, M. G. Genetic neurological channelopathies: molecular genetics and clinical phenotypes. *Journal of Neurology, Neurosurgery & Psychiatry* **87**, 37, (2016).

16 Sangwan, V. & Park, M. Receptor tyrosine kinases: role in cancer progression. *Current Oncology* **13**, 191-193, (2006).

17 Paul, M. K. & Mukhopadhyay, A. K. Tyrosine kinase – Role and significance in Cancer. *Int. J. Med. Sci.* **1**, 101-115, (2004).

18 Miller, S. I. Antibiotic Resistance and Regulation of the Gram-Negative Bacterial Outer Membrane Barrier by Host Innate Immune Molecules. *mBio* **7**, (2016).

19 Pielak, R. M. & Chou, J. J. Influenza M2 proton channels. *BBA - Biomembranes* **1808**, 522-529, (2011).

20 Blumenthal, R., Durell, S. & Viard, M. HIV Entry and Envelope Glycoprotein-mediated Fusion. *J. Biol. Chem.* **287**, 40841-40849, (2012).

21 Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nature Reviews Molecular Cell Biology* **18**, 361, (2017).

22 Lacapère, J.-J., Pebay-Peyroula, E., Neumann, J.-M. & Etchebest, C. Determining membrane protein structures: still a challenge! *Trends Biochem. Sci.* **32**, 259-270, (2007).

23 Zorman, S., Botte, M., Jiang, Q., Collinson, I. & Schaffitzel, C. Advances and challenges of membrane–protein complex production. *Curr. Opin. Struct. Biol.* **32**, 123-130, (2015).
24 Ghosh, E., Kumari, P., Jaiman, D. & Shukla, A. K. Methodological advances: the unsung heroes of the GPCR structural revolution. *Nat. Rev. Mol. Cell Biol.* **16**, 69-81, (2015).
25 Rawson, S., Davies, S., Lippiat, J. D. & Muench, S. P. The changing landscape of membrane protein structural biology through developments in electron microscopy. *Mol. Membr. Biol.* **33**, 12-22, (2016).
26 Hedger, G. & Sansom, M. S. P. Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations. *BBA - Biomembranes* **1858**, 2390-2400, (2016).
27 Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *BBA - Biomembranes* **1838**, 1548-1559, (2014).
28 Koshy, C. & Ziegler, C. Structural insights into functional lipid–protein interactions in secondary transporters. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1850**, 476-487, (2015).
29 Marsh, D. Protein modulation of lipids, and vice-versa, in membranes. *BBA - Biomembranes* **1778**, 1545-1575, (2008).
30 Barrera, N. P., Zhou, M. & Robinson, C. V. The role of lipids in defining membrane protein interactions: insights from mass spectrometry. *Trends Cell Biol.* **23**, 1-8, (2013).
31 van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112-124, (2008).
32 Berg, J. M., Tymoczko, J. L., Stryer, L. & Gatto, G. J. in *Biochemistry* Ch. 12, (W.H. Freeman, 2012).
33 Mbamala, E. C., Ben-Shaul, A. & May, S. Domain Formation Induced by the Adsorption of Charged Proteins on Mixed Lipid Membranes. *Biophys. J.* **88**, 1702-1714, (2005).
34 Lewis;, R. N. A. H. & McElhaney, R. N. in *The structure of biological membranes* (ed Philip Yeagle) Ch. 2, (CRC Press, 2012).
35 Garcia-Manyes, S., Oncins, G. & Sanz, F. Effect of Ion-Binding and Chemical Phospholipid Structure on the Nanomechanics of Lipid Bilayers Studied by Force Spectroscopy. *Biophys. J.* **89**, 1812-1826, (2005).
36 Pomorski, T. G., Nylander, T. & Cárdenas, M. Model cell membranes: Discerning lipid and protein contributions in shaping the cell. *Adv. Colloid Interface Sci.* **205**, 207-220, (2014).
37 Lippincott-Schwartz, J. & Phair, R. D. Lipids and Cholesterol as Regulators of Traffic in the Endomembrane System. *Annual Review of Biophysics* **39**, 559-578, (2010).
38 Slotte, J. P. The importance of hydrogen bonding in sphingomyelin's membrane interactions with co-lipids. *BBA - Biomembranes* **1858**, 304-310, (2016).
39 Brown, D. A. & London, E. Structure and function of sphingolipid- and cholesterol-rich membrane rafts. *J. Biol. Chem.* **275**, 17221-17224, (2000).
40 Pike, L. J. Rafts defined: a report on the Keystone Symposium on Lipid Rafts and Cell Function. *J. Lipid Res.* **47**, 1597-1598, (2006).
41 Lippoldt, J., Händel, C., Dietrich, U. & Käs, J. A. Dynamic membrane structure induces temporal pattern formation. *BBA - Biomembranes* **1838**, 2380-2390, (2014).
42 Sanders, C. R. & Mittendorf, K. F. Tolerance to Changes in Membrane Lipid Composition as a Selected Trait of Membrane Proteins. *Biochemistry* **50**, 7858-7867, (2011).
43 Coskun, U. & Simons, K. Cell membranes: the lipid perspective. *Structure* **19**, 1543-1548, (2011).
44 Escribá, P. V. *et al.* Membranes: a meeting point for lipids, proteins and therapies. *J. Cell. Mol. Med.* **12**, 829-875, (2008).

45 Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181**, 223, (1973).
46 Lodish, H. F. *et al.* in *Molecular cell biology* Ch. 3, (W.H. Freeman-Macmillan Learning, 2016).
47 Zhou, H.-X. & Cross, T. A. Influences of Membrane Mimetic Environments on Membrane Protein Structures. *Annual Review of Biophysics* **42**, 361-392, (2013).
48 Palsdottir, H. & Hunte, C. Lipids in membrane protein structures. *BBA - Biomembranes* **1666**, 2-18, (2004).
49 Dowhan, W. & Bogdanov, M. Lipid-Dependent Membrane Protein Topogenesis. *Annu. Rev. Biochem.* **78**, 515-540, (2009).
50 Phillips, R., Ursell, T., Wiggins, P. & Sens, P. Emerging roles for lipids in shaping membrane-protein function. *Nature* **459**, 379-385, (2009).
51 Cournia, Z. *et al.* Membrane Protein Structure, Function, and Dynamics: a Perspective from Experiments and Theory. *The Journal of Membrane Biology* **248**, 611-640, (2015).
52 Lee, A. G. How lipids affect the activities of integral membrane proteins. *BBA - Biomembranes* **1666**, 62-87, (2004).
53 Li, L., Shi, X., Guo, X., Li, H. & Xu, C. Ionic protein–lipid interaction at the plasma membrane: what can the charge do? *Trends Biochem. Sci.* **39**, 130-140, (2014).
54 Dawaliby, R. *et al.* Allosteric regulation of G protein-coupled receptor activity by phospholipids. *Nat. Chem. Biol.* **12**, 35-39, (2016).
55 Yamasaki, K., Daiho, T., Danko, S., Yasuda, S. & Suzuki, H. Nanodisc-based kinetic assays reveal distinct effects of phospholipid head groups on the phosphoenzyme transition of sarcoplasmic reticulum Ca2+-ATPase. *J. Biol. Chem.*, (2017).
56 Suh, B.-C. & Hille, B. PIP(2) is a necessary cofactor for ion channel function: How and why? *Annual review of biophysics* **37**, 175-195, (2008).
57 Whorton, M. R. & MacKinnon, R. Crystal structure of the mammalian GIRK2 K+ channel and gating regulation by G proteins, PIP2, and sodium. *Cell* **147**, 199-208, (2011).
58 Contreras, F. X. *et al.* Molecular recognition of a single sphingolipid species by a protein's transmembrane domain. *Nature* **481**, 525, (2012).
59 Nury, H. *et al.* Structural basis for lipid-mediated interactions between mitochondrial ADP/ATP carrier monomers. *FEBS Lett.* **579**, 6031-6036, (2005).
60 Goddard, A. D., Dijkman, P. M., Adamson, R. J. & Watts, A. Lipid-Dependent GPCR Dimerization. *Methods Cell Biol.* **117**, 341-357, (2013).
61 Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **541**, 421-424, (2017).
62 Schwall, C. T., Greenwood, V. L. & Alder, N. N. The stability and activity of respiratory Complex II is cardiolipin-dependent. *Biochim. Biophys. Acta* **1817**, 1588-1596, (2012).
63 Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172-175, (2014).
64 Nicolson, G. L. The Fluid-Mosaic Model of Membrane Structure: still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years. *Biochim. Biophys. Acta* **1838**, 1451-1466, (2014).
65 Kaiser, H.-J. *et al.* Lateral sorting in model membranes by cholesterol-mediated hydrophobic matching. *Proceedings of the National Academy of Sciences* **108**, 16628, (2011).
66 Sharpe, H. J., Stevens, T. J. & Munro, S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **142**, 158-169, (2010).
67 Prestwich, G. D. Phosphoinositide Signaling: From Affinity Probes to Pharmaceutical Targets. *Chem. Biol.* **11**, 619-637, (2004).

68 Balla, T., Szentpetery, Z. & Kim, Y. J. Phosphoinositide Signaling: New Tools and Insights. *Physiology* **24**, 231-244, (2009).
69 Montigny, C., Lyons, J., Champeil, P., Nissen, P. & Lenoir, G. On the molecular mechanism of flippase- and scramblase-mediated phospholipid transport. *Biochim. Biophys. Acta* **1861**, 767-783, (2016).
70 Wymann, M. P. & Schneiter, R. Lipid signalling in disease. *Nat. Rev. Mol. Cell Biol.* **9**, 162-176, (2008).
71 Contreras, F.-X., Ernst, A. M., Wieland, F. & Brügger, B. Specificity of Intramembrane Protein–Lipid Interactions. *Cold Spring Harb. Perspect. Biol.* **3**, (2011).
72 Hille, B., Dickson, E. J., Kruse, M., Vivas, O. & Suh, B.-C. Phosphoinositides regulate ion channels. *Biochim. Biophys. Acta* **1851**, 844-856, (2015).
73 Hansen, S. B., Tao, X. & MacKinnon, R. Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2. *Nature* **477**, 495-498, (2011).
74 Lee, C.-H. *et al.* Non-basic amino acids in the ROMK1 channels via an appropriate distance modulate PIP2 regulated pHi-gating. *Biochem. Biophys. Res. Commun.* **473**, 303-310, (2016).
75 Hansen, S. B. Lipid agonism: The PIP2 paradigm of ligand-gated ion channels. *Biochim. Biophys. Acta* **1851**, 620-628, (2015).
76 Stangl, M. & Schneider, D. Functional competition within a membrane: Lipid recognition vs. transmembrane helix oligomerization. *BBA - Biomembranes* **1848**, 1886-1896, (2015).
77 Lee, S.-J. *et al.* Structural basis of control of inward rectifier Kir2 channel gating by bulk anionic phospholipids. *The Journal of General Physiology* **148**, 227, (2016).
78 Huang, C.-L., Feng, S. & Hilgemann, D. W. Direct activation of inward rectifier potassium channels by PIP2 and its stabilization by Gβγ. *Nature* **391**, 803, (1998).
79 Whorton, M. R. & MacKinnon, R. X-ray structure of the mammalian GIRK2–βγ G-protein complex. *Nature* **498**, 190, (2013).
80 Lee, A. G. Biological membranes: the importance of molecular detail. *Trends Biochem. Sci.* **36**, 493-500, (2011).
81 Lee, A. G. Lipid–protein interactions in biological membranes: a structural perspective. *BBA - Biomembranes* **1612**, 1-40, (2003).
82 Gonen, T. *et al.* Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* **438**, 633-638, (2005).
83 Marsh, D. Electron spin resonance in membrane research: Protein–lipid interactions. *Methods* **46**, 83-96, (2008).
84 East, J. M., Melville, D. & Lee, A. G. Exchange rates and numbers of annular lipids for the calcium and magnesium ion dependent adenosine triphosphatase. *Biochemistry* **24**, 2615-2623, (1985).
85 Hite, R. K., Li, Z. & Walz, T. Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals. *The EMBO Journal* **29**, 1652, (2010).
86 Bocquet, N. *et al.* X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* **457**, 111-114, (2009).
87 Bechara, C. *et al.* A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nat. Chem.* **7**, 255-262, (2015).
88 Eggensperger, S., Fisette, O., Parcej, D., Schäfer, L. V. & Tampé, R. An Annular Lipid Belt Is Essential for Allosteric Coupling and Viral Inhibition of the Antigen Translocation Complex TAP (Transporter Associated with Antigen Processing). *J. Biol. Chem.* **289**, 33098-33108, (2014).
89 Tang, D., Dean, W. L., Borchman, D. & Paterson, C. A. The influence of membrane lipid structure on plasma membrane Ca2+-ATPase activity. *Cell Calcium* **39**, 209-216, (2006).

90 Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347, (2016).
91 Starling, A. P., Dalton, K. A., East, J. M., Oliver, S. & Lee, A. G. Effects of phosphatidylethanolamines on the activity of the Ca(2+)-ATPase of sarcoplasmic reticulum. *Biochem. J.* **320**, 309-314, (1996).
92 Bennett, J. P., McGill, K. A. & Warren, G. B. Transbilayer disposition of the phospholipid annulus surrounding a calcium transport protein. *Nature* **274**, 823, (1978).
93 Starling, A. P., East, J. M. & Lee, A. G. Effects of phospholipid fatty acyl chain length on phosphorylation and dephosphorylation of the Ca2+-ATPase. *Biochem. J.* **310**, 875, (1995).
94 Gustavsson, M., Traaseth, N. J. & Veglia, G. Activating and Deactivating Roles of Lipid Bilayers on the Ca2+-ATPase/Phospholamban Complex. *Biochemistry* **50**, 10367-10374, (2011).
95 Filippov, A., Orädd, G. & Lindblom, G. The Effect of Cholesterol on the Lateral Diffusion of Phospholipids in Oriented Bilayers. *Biophys. J.* **84**, 3079-3086, (2003).
96 Lee, S.-J. *et al.* Secondary anionic phospholipid binding site and gating mechanism in Kir2.1 inward rectifier channels. *Nature Communications* **4**, 2786, (2013).
97 Cheng, W. W. L., D'Avanzo, N., Doyle, D. A. & Nichols, C. G. Dual-Mode Phospholipid Regulation of Human Inward Rectifying Potassium Channels. *Biophys. J.* **100**, 620-628, (2011).
98 Pliotas, C. *et al.* The role of lipids in mechanosensation. *Nature Structural &Amp; Molecular Biology* **22**, 991, (2015).
99 Yoshimura, K. & Sokabe, M. Mechanosensitivity of ion channels based on protein-lipid interactions. *Journal of the Royal Society, Interface* **7 Suppl 3**, S307-320, (2010).
100 Bavi, O., Vossoughi, M., Naghdabadi, R. & Jamali, Y. The Combined Effect of Hydrophobic Mismatch and Bilayer Local Bending on the Regulation of Mechanosensitive Ion Channels. *PLoS One* **11**, e0150578, (2016).
101 Collins, T., Young, G. T. & Millar, N. S. Competitive binding at a nicotinic receptor transmembrane site of two α7-selective positive allosteric modulators with differing effects on agonist-evoked desensitization. *Neuropharmacology* **61**, 1306-1313, (2011).
102 Ivetac, A. & Andrew McCammon, J. Mapping the Druggable Allosteric Space of G-Protein Coupled Receptors: a Fragment-Based Molecular Dynamics Approach. *Chem. Biol. Drug Des.* **76**, 201-217, (2010).
103 Ottosson, N. E. *et al.* A drug pocket at the lipid bilayer–potassium channel interface. *Science Advances* **3**, (2017).
104 Jazayeri, A. *et al.* Extra-helical binding site of a glucagon receptor antagonist. *Nature* **533**, 274, (2016).
105 Zhang, D. *et al.* Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature* **520**, 317, (2015).
106 Escribá, P. V. Membrane-lipid therapy: a new approach in molecular medicine. *Trends Mol. Med.* **12**, 34-43, (2006).
107 Escriba, P. V. *et al.* Membrane lipid therapy: Modulation of the cell membrane composition and structure as a molecular base for drug discovery and new disease treatment. *Prog. Lipid Res.* **59**, 38-53, (2015).
108 Tan, L. T. *et al.* Targeting Membrane Lipid a Potential Cancer Cure? *Front. Pharmacol.* **8**, 12, (2017).
109 Subramaniam, S. *et al.* Bioinformatics and Systems Biology of the Lipidome. *Chem. Rev.* **111**, 6452-6490, (2011).
110 Lodish, H. F. *et al.* in *Molecular cell biology* Ch. 15, (W.H. Freeman-Macmillan Learning, 2016).

111 Schöneberg, T. *et al.* Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacol. Ther.* **104**, 173-206, (2004).
112 Thompson, M. D., Hendy, G. N., Percy, M. E., Bichet, D. G. & Cole, D. E. C. in *Pharmacogenomics in Drug Discovery and Development* (ed Qing Yan) 153-187 (Springer New York, 2014).
113 Kepinska, M., Siara, M. & Sajewicz, W. *G protein-coupled receptors: Abnormalities in signal transmission, disease states and pharmacotherapy*. Vol. 71 (2014).
114 Tautermann, C. S. GPCR structures in drug design, emerging opportunities with new structures. *Bioorg. Med. Chem. Lett.* **24**, 4073-4079, (2014).
115 Hu, G.-M., Mai, T.-L. & Chen, C.-M. Visualizing the GPCR Network: Classification and Evolution. *Sci. Rep.* **7**, 15495, (2017).
116 Attwood, T. K. & Findlay, J. B. C. Fingerprinting G-protein-coupled receptors. *Protein Engineering, Design and Selection* **7**, 195-203, (1994).
117 Hilger, D., Masureel, M. & Kobilka, B. K. Structure and dynamics of GPCR signaling complexes. *Nat. Struct. Mol. Biol.* **25**, 4-12, (2018).
118 Ritter, S. L. & Hall, R. A. Fine-tuning of GPCR activity by receptor-interacting proteins. *Nat. Rev. Mol. Cell Biol.* **10**, 819-830, (2009).
119 Baltoumas, F. A., Theodoropoulou, M. C. & Hamodrakas, S. J. Interactions of the α-subunits of heterotrimeric G-proteins with GPCRs, effectors and RGS proteins: A critical review and analysis of interacting surfaces, conformational shifts, structural diversity and electrostatic potentials. *J. Struct. Biol.* **182**, 209-218, (2013).
120 Ferguson, S. S. G. Evolving Concepts in G Protein-Coupled Receptor Endocytosis: The Role in Receptor Desensitization and Signaling. *Pharmacol. Rev.* **53**, 1, (2001).
121 Thomsen, Alex R. B. *et al.* GPCR-G Protein-β-Arrestin Super-Complex Mediates Sustained G Protein Signaling. *Cell* **166**, 907-919, (2016).
122 Ferrandon, S. *et al.* Sustained cyclic AMP production by parathyroid hormone receptor endocytosis. *Nat. Chem. Biol.* **5**, 734, (2009).
123 Manglik, A. & Kobilka, B. The role of protein dynamics in GPCR function: insights from the β2AR and rhodopsin. *Curr. Opin. Cell Biol.* **27**, 136-143, (2014).
124 Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR Dynamics: Structures in Motion. *Chem. Rev.* **117**, 139-155, (2017).
125 Zhang, D., Zhao, Q. & Wu, B. Structural Studies of G Protein-Coupled Receptors. *Mol. Cells* **38**, 836-842, (2015).
126 Lu, M. & Wu, B. Structural studies of G protein-coupled receptors. *IUBMB Life* **68**, 894-903, (2016).
127 Okude, J. *et al.* Identification of a Conformational Equilibrium That Determines the Efficacy and Functional Selectivity of the μ-Opioid Receptor. *Angew. Chem. Int. Ed.* **54**, 15771-15776, (2015).
128 Gregorio, G. G. *et al.* Single-molecule analysis of ligand efficacy in β2AR–G-protein activation. *Nature* **547**, 68, (2017).
129 Galandrin, S. & Bouvier, M. Distinct Signaling Profiles of β1 and β2 Adrenergic Receptor Ligands toward Adenylyl Cyclase and Mitogen-Activated Protein Kinase Reveals the Pluridimensionality of Efficacy. *Mol. Pharmacol.* **70**, 1575, (2006).
130 Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549-555, (2011).
131 Venkatakrishnan, A. J. *et al.* Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature* **536**, 484, (2016).
132 Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature* **469**, 175-180, (2011).
133 Kang, Y. *et al.* Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* **523**, 561-567, (2015).

134 Carpenter, B., Nehmé, R., Warne, T., Leslie, A. G. W. & Tate, C. G. Structure of the adenosine A2A receptor bound to an engineered G protein. *Nature* **536**, 104-107, (2016).
135 Furness, S. G. B. *et al.* Ligand-Dependent Modulation of G Protein Conformation Alters Drug Efficacy. *Cell* **167**, 739-749.e711, (2016).
136 DeVree, B. T. *et al.* Allosteric coupling from G protein to the agonist-binding pocket in GPCRs. *Nature* **535**, 182-186, (2016).
137 Hanson, M. A. *et al.* A specific cholesterol binding site is established by the 2.8 Å structure of the human β(2)-adrenergic receptor in an alternate crystal form. *Structure* **16**, 897-905, (2008).
138 Oates, J. & Watts, A. Uncovering the intimate relationship between lipids, cholesterol and GPCR activation. *Curr. Opin. Struct. Biol.* **21**, 802-807, (2011).
139 Prasanna, X., Sengupta, D. & Chattopadhyay, A. Cholesterol-dependent Conformational Plasticity in GPCR Dimers. *Sci. Rep.* **6**, 31858, (2016).
140 Pucadyil, T. J. & Chattopadhyay, A. Cholesterol modulates ligand binding and G-protein coupling to serotonin1A receptors from bovine hippocampus. *BBA - Biomembranes* **1663**, 188-200, (2004).
141 Byrne, E. F. X. *et al.* Structural basis of Smoothened regulation by its extracellular domains. *Nature* **535**, 517-522, (2016).
142 Inagaki, S. *et al.* Modulation of the Interaction between Neurotensin Receptor NTS1 and Gq Protein by Lipid. *J. Mol. Biol.* **417**, 95-111, (2012).
143 Vögler, O. *et al.* The Gβγ Dimer Drives the Interaction of Heterotrimeric Gi Proteins with Nonlamellar Membrane Structures. *J. Biol. Chem.* **279**, 36540-36545, (2004).
144 Paila, Y. D. & Chattopadhyay, A. in *Cholesterol Binding and Cholesterol Transport Proteins: Structure and Function in Health and Disease* (ed J. Robin Harris) 439-466 (Springer Netherlands, 2010).
145 Saliba, A.-E., Vonkova, I. & Gavin, A.-C. The systematic analysis of protein-lipid interactions comes of age. *Nat. Rev. Mol. Cell Biol.* **16**, 753-761, (2015).
146 Francy, C. A., Clinton, R. W., Fröhlich, C., Murphy, C. & Mears, J. A. Cryo-EM Studies of Drp1 Reveal Cardiolipin Interactions that Activate the Helical Oligomer. *Sci. Rep.* **7**, 10744, (2017).
147 Montenegro, F. A., Cantero, J. R. & Barrera, N. P. Combining Mass Spectrometry and X-Ray Crystallography for Analyzing Native-Like Membrane Protein Lipid Complexes. *Front. Physiol.* **8**, 892, (2017).
148 Murata, T., Yamato, I., Kakinuma, Y., Leslie, A. G. W. & Walker, J. E. Structure of the Rotor of the V-Type Na+-ATPase from Enterococcus hirae. *Science* **308**, 654, (2005).
149 Zhou, M. *et al.* Mass Spectrometry of Intact V-Type ATPases Reveals Bound Lipids and the Effects of Nucleotide Binding. *Science* **334**, 380, (2011).
150 Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular Interaction Studies Using Microscale Thermophoresis. *Assay Drug Dev. Technol.* **9**, 342-353, (2011).
151 Seidel, S. A. I. *et al.* Label-Free Microscale Thermophoresis Discriminates Sites and Affinity of Protein–Ligand Binding. *Angew. Chem. Int. Ed.* **51**, 10656-10659, (2012).
152 Robinson, C. V. From molecular chaperones to membrane motors: through the lens of a mass spectrometrist. *Biochem. Soc. Trans.* **45**, 251, (2017).
153 Lössl, P., van de Waterbeemd, M. & Heck, A. J. R. The diverse and expanding role of mass spectrometry in structural and molecular biology. *The EMBO Journal* **35**, 2634, (2016).
154 Smith, R., Mathis, A. D., Ventura, D. & Prince, J. T. Proteomics, lipidomics, metabolomics: a mass spectrometry tutorial from a computer scientist's point of view. *BMC Bioinformatics* **15**, S9, (2014).

155 Katta, V. & Chait, B. T. Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J. Am. Chem. Soc.* **113**, 8534-8535, (1991).
156 Ganem, B., Li, Y. T. & Henion, J. D. Detection of noncovalent receptor-ligand complexes by mass spectrometry. *J. Am. Chem. Soc.* **113**, 6294-6296, (1991).
157 Ruotolo Brandon , T. *et al.* Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes. *Angew. Chem. Int. Ed.* **46**, 8001-8004, (2007).
158 Leney, A. C. & Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **28**, 5-13, (2017).
159 Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chem. Rev.* **107**, 3544-3567, (2007).
160 Landreh, M. & Robinson, C. V. A new window into the molecular physiology of membrane proteins. *The Journal of Physiology* **593**, 355-362, (2015).
161 van den Heuvel, R. H. & Heck, A. J. Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* **8**, 519-526, (2004).
162 Boeri Erba, E. & Petosa, C. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Sci.* **24**, 1176-1192, (2015).
163 Mehmood, S., Allison, T. M. & Robinson, C. V. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu. Rev. Phys. Chem.* **66**, 453-474, (2015).
164 Lanucara, F., Holman, S. W., Gray, C. J. & Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nat. Chem.* **6**, 281-294, (2014).
165 Karas, M., Bahr, U. & Dülcks, T. Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. *Fresenius J. Anal. Chem.* **366**, 669-676, (2000).
166 Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64, (1989).
167 Dole, M. *et al.* Molecular Beams of Macroions. *The Journal of Chemical Physics* **49**, 2240-2249, (1968).
168 Hilton, G. R. & Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of The Royal Society Interface* **9**, 801-816, (2012).
169 Mädler, S., Erba, E. B. & Zenobi, R. in *Applications of MALDI-TOF Spectroscopy* (eds Zongwei Cai & Shuying Liu) 1-36 (Springer Berlin Heidelberg, 2013).
170 Frauke, M., Lucie, S., Julia, L., Bernhard, B. & Stephan, N. LILBID-mass spectrometry of the mitochondrial preprotein translocase TOM. *J. Phys.: Condens. Matter* **22**, 454132, (2010).
171 Ambrose, S. *et al.* Native Desorption Electrospray Ionization Liberates Soluble and Membrane Protein Complexes from Surfaces. *Angew. Chem. Int. Ed.* **56**, 14463-14468, (2017).
172 Dunn, G. H. in *Electron Impact Ionization* (eds Tilmann D. Märk & Gordon H. Dunn) 277-319 (Springer Vienna, 1985).
173 Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protocols* **8**, 639-651, (2013).
174 Kebarle, P. & Verkerk Udo, H. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898-917, (2009).
175 Fernández de la Mora, J. The Fluid Dynamics of Taylor Cones. *Annual Review of Fluid Mechanics* **39**, 217-243, (2006).

176 Tang, K. & Smith, R. D. Physical/chemical separations in the break-up of highly charged droplets from electrosprays. *J. Am. Soc. Mass Spectrom.* **12**, 343-347, (2001).
177 Banerjee, S. & Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012**, (2012).
178 Wilm, M. & Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **68**, 1-8, (1996).
179 Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **85**, 2-9, (2013).
180 Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics* **64**, 2287-2294, (1976).
181 Iavarone, A. T. & Williams, E. R. Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *J. Am. Chem. Soc.* **125**, 2319-2327, (2003).
182 Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *Int. J. Mass spectrom.* **219**, 39-62, (2002).
183 Gamero-Castaño, M. & Fernández de la Mora, J. Mechanisms of electrospray ionization of singly and multiply charged salt clusters. *Anal. Chim. Acta* **406**, 67-91, (2000).
184 Konermann, L., Rodriguez, A. D. & Liu, J. On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. *Anal. Chem.* **84**, 6798-6804, (2012).
185 Gumerov, D. R., Dobo, A. & Kaltashov, I. A. Protein—Ion Charge-State Distributions in Electrospray Ionization Mass Spectrometry: Distinguishing Conformational Contributions from Masking Effects. *Eur. J. Mass Spectrom.* **8**, 123-129, (2002).
186 Rostom, A. A. & Robinson, C. V. Detection of the Intact GroEL Chaperonin Assembly by Mass Spectrometry. *J. Am. Chem. Soc.* **121**, 4718-4719, (1999).
187 Borysik, A. J., Hewitt, D. J. & Robinson, C. V. Detergent Release Prolongs the Lifetime of Native-like Membrane Protein Conformations in the Gas-Phase. *J. Am. Chem. Soc.* **135**, 6078-6083, (2013).
188 Barrera, N. P., Di Bartolo, N., Booth, P. J. & Robinson, C. V. Micelles Protect Membrane Complexes from Solution to Vacuum. *Science* **321**, 243, (2008).
189 Shvartsburg, A. A. & Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **80**, 9689-9699, (2008).
190 Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. & Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* **3**, 1139-1152, (2008).
191 Hofstadler, S. A., Bakhtiar, R. & Smith, R. D. Electrospray Ionization Mass Spectroscopy: Part I. Instrumentation and Spectral Interpretation. *J. Chem. Educ.* **73**, A82, (1996).
192 Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A. & Robinson, C. V. A Tandem Mass Spectrometer for Improved Transmission and Analysis of Large Macromolecular Assemblies. *Anal. Chem.* **74**, 1402-1407, (2002).
193 Douglas Donald, J., Frank Aaron, J. & Mao, D. Linear ion traps in mass spectrometry. *Mass Spectrom. Rev.* **24**, 1-29, (2004).
194 Cotter, R. J. Time-of-flight mass spectrometry for the structural analysis of biological molecules. *Anal. Chem.* **64**, 1027A-1039A, (1992).
195 Benesch, J. L. P. *et al.* Quadrupole-Time-of-Flight Mass Spectrometer Modified for Higher-Energy Dissociation Reduces Protein Assemblies to Peptide Fragments. *Anal. Chem.* **81**, 1270-1274, (2009).
196 Doroshenko, V. M. & Cotter, R. J. Ideal velocity focusing in a reflectron time-of-flight mass spectrometer. *J. Am. Soc. Mass Spectrom.* **10**, 992-999, (1999).

197 Mamyrin, B. A. K., V. I.; Shmikk, D. V.; Zagulin, V. A. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Journal of Experimental and Theoretical Physics* **37**, 45, (1973).
198 Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **72**, 1156-1162, (2000).
199 Eliuk, S. & Makarov, A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Review of Analytical Chemistry* **8**, 61-80, (2015).
200 Kingdon, K. H. A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. *Physical Review* **21**, 408-418, (1923).
201 Scigelova, M. & Makarov, A. Orbitrap Mass Analyzer – Overview and Applications in Proteomics. *Proteomics* **6**, 16-21, (2006).
202 Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* **40**, 430-443, (2005).
203 van de Waterbeemd, M. *et al.* High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles. *Nat. Methods* **14**, 283, (2017).
204 Gys, T. Micro-channel plates and vacuum detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **787**, 254-260, (2015).
205 Ladislas Wiza, J. Microchannel plate detectors. *Nuclear Instruments and Methods* **162**, 587-601, (1979).
206 Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **389**, 1017-1031, (2007).
207 Winger, B. E., Hofstadler, S. A., Bruce, J. E., Udseth, H. R. & Smith, R. D. High-resolution accurate mass measurements of biomolecules using a new electrospray ionization ion cyclotron resonance mass spectrometer. *J. Am. Soc. Mass Spectrom.* **4**, 566-577, (1993).
208 Reading, E. *et al.* The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase. *Angew. Chem. Int. Ed.* **54**, 4577-4581, (2015).
209 Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat. Methods* **13**, 333-336, (2016).
210 Ruotolo, B. T. & Robinson, C. V. Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.* **10**, 402-408, (2006).
211 Benesch, J. L. & Ruotolo, B. T. Mass spectrometry: come of age for structural and dynamical biology. *Curr. Opin. Struct. Biol.* **21**, 641-649, (2011).
212 Shvartsburg, A. A. & Jarrold, M. F. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86-91, (1996).
213 Bechara, C. & Robinson, C. V. Different Modes of Lipid Binding to Membrane Proteins Probed by Mass Spectrometry. *J. Am. Chem. Soc.* **137**, 5240-5247, (2015).
214 Bolla, J. R. *et al.* Direct observation of the influence of cardiolipin and antibiotics on lipid II binding to MurJ. *Nat. Chem.* **10**, 363, (2018).
215 Marcoux, J. *et al.* Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump. *Proceedings of the National Academy of Sciences* **110**, 9704-9709, (2013).
216 Schmidt, C. & Robinson, C. V. Dynamic protein ligand interactions – insights from MS. *FEBS J.* **281**, 1950-1964, (2014).
217 Kim, J. *et al.* Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. *Nature* **517**, 396, (2014).
218 Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185-194, (2013).

219 Marty, M. T., Hoi, K. K. & Robinson, C. V. Interfacing Membrane Mimetics with Mass Spectrometry. *Acc. Chem. Res.* **49**, 2459-2467, (2016).
220 Hopper, J. T. *et al.* Detergent-free mass spectrometry of membrane protein complexes. *Nat. Methods* **10**, 1206-1208, (2013).
221 Calabrese, A. N., Watkinson, T. G., Henderson, P. J. F., Radford, S. E. & Ashcroft, A. E. Amphipols Outperform Dodecylmaltoside Micelles in Stabilizing Membrane Protein Structure in the Gas Phase. *Anal. Chem.* **87**, 1118-1126, (2015).
222 Watkinson, T. G. *et al.* Systematic analysis of the use of amphipathic polymers for studies of outer membrane proteins using mass spectrometry. *Int. J. Mass spectrom.* **391**, 54-61, (2015).
223 Dürr, U. H. N., Gildenberg, M. & Ramamoorthy, A. The Magic of Bicelles Lights Up Membrane Protein Structure. *Chem. Rev.* **112**, 6054-6074, (2012).
224 Lee, S. C. *et al.* A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat. Protocols* **11**, 1149-1162, (2016).
225 Dörr, J. M. *et al.* The styrene–maleic acid copolymer: a versatile tool in membrane research. *Eur. Biophys. J.* **45**, 3-21, (2016).
226 Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2**, 853-856, (2002).
227 Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J. Am. Soc. Mass. Spectrom.* **126**, 3477-3487, (2004).
228 Bibow, S. *et al.* Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I. *Nat. Struct. Mol. Biol.* **24**, 187-193, (2017).
229 Skar-Gislinge, N. *et al.* Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *J. Am. Chem. Soc.* **132**, 13713-13722, (2010).
230 Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. *FEBS Lett.* **584**, 1721-1727, (2010).
231 Marty, M., Zhang, H., Cui, W., Gross, M. & Sligar, S. Interpretation and Deconvolution of Nanodisc Native Mass Spectra. *J. Am. Soc. Mass Spectrom.* **25**, 269-277, (2014).
232 Shaw, A. W., Pureza, V. S., Sligar, S. G. & Morrissey, J. H. The Local Phospholipid Environment Modulates the Activation of Blood Clotting. *J. Biol. Chem.* **282**, 6556-6563, (2007).
233 Denisov, I. G. & Sligar, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chem. Rev.* **117**, 4669-4713, (2017).
234 Denisov, I. G. & Sligar, S. G. Nanodiscs for structural and functional studies of membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 481-486, (2016).
235 Han, L. *et al.* Protein-glycolipid interactions studied in vitro using ESI-MS and nanodiscs: insights into the mechanisms and energetics of binding. *Anal. Chem.* **87**, 4888-4896, (2015).
236 Leney, A. C., Fan, X., Kitova, E. N. & Klassen, J. S. Nanodiscs and Electrospray Ionization Mass Spectrometry: A Tool for Screening Glycolipids Against Proteins. *Anal. Chem.* **86**, 5271-5277, (2014).
237 Zhang, Y. *et al.* Protein–Glycosphingolipid Interactions Revealed Using Catch-and-Release Mass Spectrometry. *Anal. Chem.* **84**, 7618-7621, (2012).
238 Marty, M. T. *et al.* Native Mass Spectrometry Characterization of Intact Nanodisc Lipoprotein Complexes. *Anal. Chem.* **84**, 8957-8960, (2012).
239 Marty, M. T. *et al.* Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **87**, 4370-4376, (2015).
240 Yen, H.-Y. *et al.* Ligand binding to a G protein–coupled receptor captured in a mass spectrometer. *Science Advances* **3**, (2017).

241 Ritchie, T. K. *et al.* in *Methods Enzymol.* Vol. Volume 464 (ed Düzgünes Nejat) 211-231 (Academic Press, 2009).
242 Schuler, M., Denisov, I. & Sligar, S. in *Lipid-Protein Interactions* Vol. 974 *Methods in Molecular Biology* (ed Jörg H. Kleinschmidt) Ch. 18, 415-433 (Humana Press, 2013).
243 Chen, P. S., Toribara, T. Y. & Warner, H. Microdetermination of Phosphorus. *Anal. Chem.* **28**, 1756-1758, (1956).
244 Dyachenko, A. *et al.* Tandem Native Mass-Spectrometry on Antibody–Drug Conjugates and Submillion Da Antibody–Antigen Protein Assemblies on an Orbitrap EMR Equipped with a High-Mass Quadrupole Mass Selector. *Anal. Chem.* **87**, 6095-6102, (2015).
245 Rose, R. J., Damoc, E., Denisov, E., Makarov, A. & Heck, A. J. R. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* **9**, 1084, (2012).
246 Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2**, 715-726, (2007).
247 Bald, T. *et al.* pymzML—Python module for high-throughput bioinformatics on mass spectrometry data. *Bioinformatics* **28**, 1052-1053, (2012).
248 Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534-2536, (2008).
249 Morgner, N. & Robinson, C. V. Massign: An Assignment Strategy for Maximizing Information from the Mass Spectra of Heterogeneous Protein Assemblies. *Anal. Chem.* **84**, 2939-2948, (2012).
250 Oliphant, T. E. Python for Scientific Computing. *COMPUT. SCI. ENG.* **9**, 10-20, (2007).
251 Millman, K. J. & Aivazis, M. Python for Scientists and Engineers. *COMPUT. SCI. ENG.* **13**, 9-12, (2011).
252 Cleary, S. P., Thompson, A. M. & Prell, J. S. Fourier Analysis Method for Analyzing Highly Congested Mass Spectra of Ion Populations with Repeated Subunits. *Anal. Chem.* **88**, 6205-6213, (2016).
253 Tsui, F. C., Ojcius, D. M. & Hubbell, W. L. The intrinsic pKa values for phosphatidylserine and phosphatidylethanolamine in phosphatidylcholine host bilayers. *Biophys. J.* **49**, 459-468, (1986).
254 Marsh, D. in *Handbook of Lipid Bilayers* Ch. 2, 43-860 (CRC Press, 2013).
255 Brügger, B., Erben, G., Sandhoff, R., Wieland, F. T. & Lehmann, W. D. Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 2339-2344, (1997).
256 Ivanova, P. T. *et al.* Electrospray ionization mass spectrometry analysis of changes in phospholipids in RBL-2H3 mastocytoma cells during degranulation. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7152-7157, (2001).
257 Almén, M. S., Nordström, K. J. V., Fredriksson, R. & Schiöth, H. B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* **7**, 50, (2009).
258 Hübner, C. A. & Jentsch, T. J. Ion channel diseases. *Hum. Mol. Genet.* **11**, 2435-2445, (2002).
259 Nath, A., Atkins, W. M. & Sligar, S. G. Applications of Phospholipid Bilayer Nanodiscs in the Study of Membranes and Membrane Proteins. *Biochemistry* **46**, 2059-2069, (2007).
260 Borgnia, M. J., Kozono, D., Calamita, G., Maloney, P. C. & Agre, P. Functional reconstitution and characterization of AqpZ, the E. coli water channel protein11Edited by W. Baumeister. *J. Mol. Biol.* **291**, 1169-1179, (1999).
261 Blakey, D. *et al.* Purification of the Escherichia coli ammonium transporter AmtB reveals a trimeric stoichiometry. *Biochem. J.* **364**, 527, (2002).

262 Javelle, A. *et al.* An Unusual Twin-His Arrangement in the Pore of Ammonia Channels Is Essential for Substrate Conductance. *J. Biol. Chem.* **281**, 39492-39498, (2006).
263 Savage, D. F., Egea, P. F., Robles-Colmenares, Y., Iii, J. D. O. C. & Stroud, R. M. Architecture and Selectivity in Aquaporins: 2.5 Å X-Ray Structure of Aquaporin Z. *PLoS Biol.* **1**, e72, (2003).
264 Grinkova, Y. V., Denisov, I. G. & Sligar, S. G. Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Engineering, Design and Selection* **23**, 843-848, (2010).
265 Midtgaard, S. R. *et al.* Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter* **10**, 738-752, (2014).
266 Kendrick, E. A Mass Scale Based on CH2 = 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. *Anal. Chem.* **35**, 2146-2154, (1963).
267 Sleno, L. The use of mass defect in modern mass spectrometry. *J. Mass Spectrom.* **47**, 226-236, (2012).
268 McKay, A. R., Ruotolo, B. T., Ilag, L. L. & Robinson, C. V. Mass Measurements of Increased Accuracy Resolve Heterogeneous Populations of Intact Ribosomes. *J. Am. Chem. Soc.* **128**, 11433-11442, (2006).
269 Mehmood, S. *et al.* Charge Reduction Stabilizes Intact Membrane Protein Complexes for Mass Spectrometry. *J. Am. Chem. Soc.* **136**, 17010-17012, (2014).
270 Daniel, J. M., Friess, S. D., Rajagopalan, S., Wendt, S. & Zenobi, R. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *Int. J. Mass spectrom.* **216**, 1-27, (2002).
271 Rajagopal, S., Rajagopal, K. & Lefkowitz, R. J. Teaching old receptors new tricks: biasing seven-transmembrane receptors. *Nature Reviews Drug Discovery* **9**, 373, (2010).
272 Chini, B. & Parenti, M. G-protein-coupled receptors, cholesterol and palmitoylation: facts about fats. *J. Mol. Endocrinol.* **42**, 371-379, (2009).
273 Rankovic, Z., Brust, T. F. & Bohn, L. M. Biased agonism: An emerging paradigm in GPCR drug discovery. *Bioorg. Med. Chem. Lett.* **26**, 241-250, (2016).
274 Gentry, P. R., Sexton, P. M. & Christopoulos, A. Novel Allosteric Modulators of G Protein-coupled Receptors. *J. Biol. Chem.* **290**, 19478-19488, (2015).
275 van der Westhuizen, E. T., Valant, C., Sexton, P. M. & Christopoulos, A. Endogenous allosteric modulators of G protein-coupled receptors. *J. Pharmacol. Exp. Ther.* **353**, 246-260, (2015).
276 Prasad, R., Paila, Y. D. & Chattopadhyay, A. Membrane cholesterol depletion enhances ligand binding function of human serotonin1A receptors in neuronal cells. *Biochem. Biophys. Res. Commun.* **390**, 93-96, (2009).
277 Chattopadhyay, A. GPCRs: Lipid-Dependent Membrane Receptors That Act as Drug Targets. *Advances in Biology* **2014**, 12, (2014).
278 Troupiotis-Tsaïlaki, A. *et al.* Ligand chain length drives activation of lipid G protein-coupled receptors. *Sci. Rep.* **7**, 2020, (2017).
279 Zocher, M., Zhang, C., Rasmussen, S. G. F., Kobilka, B. K. & Müller, D. J. Cholesterol increases kinetic, energetic, and mechanical stability of the human β2-adrenergic receptor. *Proceedings of the National Academy of Sciences* **109**, E3463, (2012).
280 Warne, T., Serrano-Vega, M. J., Tate, C. G. & Schertler, G. F. X. Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expression Purif.* **65**, 204-213, (2009).
281 Warne, T., Chirnside, J. & Schertler, G. F. X. Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization. *BBA - Biomembranes* **1610**, 133-140, (2003).
282 Warne, T. *et al.* The structural basis for agonist and partial agonist action on a β1-adrenergic receptor. *Nature* **469**, 241-244, (2011).

283 Nehmé, R. *et al.* Mini-G proteins: Novel tools for studying GPCRs in their active conformation. *PLoS One* **12**, e0175642, (2017).
284 Carpenter, B. & Tate, C. G. Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation. *Protein Engineering, Design and Selection* **29**, 583-594, (2016).
285 Ring, A. M. *et al.* Adrenaline-activated structure of β2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575, (2013).
286 Carpenter, B. & Tate, C. G. Expression and Purification of Mini G Proteins from Escherichia coli. *Bio-protocol* **7**, e2235, (2017).
287 Scott, D. J., Kummer, L., Egloff, P., Bathgate, R. A. D. & Plückthun, A. Improving the apo-state detergent stability of NTS1 with CHESS for pharmacological and structural studies. *BBA - Biomembranes* **1838**, 2817-2824, (2014).
288 Egloff, P. *et al.* Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli. *Proceedings of the National Academy of Sciences* **111**, E655-E662, (2014).
289 Landreh, M. *et al.* Controlling release, unfolding and dissociation of membrane protein complexes in the gas phase through collisional cooling. *Chem. Commun.* **51**, 15582-15584, (2015).
290 Sud, M., Fahy, E., Cotter, D., Dennis, E. A. & Subramaniam, S. LIPID MAPS-Nature Lipidomics Gateway: An Online Resource for Students and Educators Interested in Lipids. *J. Chem. Educ.* **89**, 291-292, (2012).
291 Dowhan, W. A retrospective: Use of Escherichia coli as a vehicle to study phospholipid synthesis and function. *Biochim. Biophys. Acta* **1831**, 471-494, (2013).
292 Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol. Rev.* **40**, 133-159, (2016).
293 Dawaliby, R. *et al.* Phosphatidylethanolamine Is a Key Regulator of Membrane Fluidity in Eukaryotic Cells. *J. Biol. Chem.* **291**, 3658-3667, (2016).
294 Traynor-Kaplan, A. *et al.* Fatty-acyl chain profiles of cellular phosphoinositides. *Biochim. Biophys. Acta* **1862**, 513-522, (2017).
295 Skwarek, L. C. & Boulianne, G. L. Great expectations for PIP: phosphoinositides as regulators of signaling during development and disease. *Dev. Cell* **16**, 12-20, (2009).
296 Schlinkmann, K. M. *et al.* Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proceedings of the National Academy of Sciences* **109**, 9810, (2012).
297 Ben-Nissan, G., Chotiner, A., Tarnavsky, M. & Sharon, M. Structural Characterization of Missense Mutations Using High Resolution Mass Spectrometry: A Case Study of the Parkinson's-Related Protein, DJ-1. *J. Am. Soc. Mass Spectrom.* **27**, 1062-1070, (2016).
298 Roth, C. B., Hanson, M. A. & Stevens, R. C. Stabilization of the Human β2-Adrenergic Receptor TM4–TM3–TM5 Helix Interface by Mutagenesis of Glu122(3.41), A Critical Residue in GPCR Structure. *J. Mol. Biol.* **376**, 1305-1319, (2008).
299 Che, T. *et al.* Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. *Cell* **172**, 55-67 e15, (2018).
300 Miller-Gallacher, J. L. *et al.* The 2.1 Å Resolution Structure of Cyanopindolol-Bound β1-Adrenoceptor Identifies an Intramembrane Na+ Ion that Stabilises the Ligand-Free Receptor. *PLoS One* **9**, e92727, (2014).
301 Injin, B. & and Hee-Jung, C. Structural Features of β2 Adrenergic Receptor: Crystal Structures and Beyond. *Mol. Cells* **38**, 105-111, (2015).