Investigating the regulation of membrane protein assemblies by lipids and other small molecules using native mass spectrometry

Denis Shutin

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Physical and Theoretical Chemistry

06.11.2020

Abstract

In this thesis I describe the application of native mass spectrometry to investigate membrane protein interactions with lipids and other small molecules. This work is a combination of method development, resulting in novel ways to study membrane protein-lipid complexes by native mass spectrometry, and application of established techniques to important biological targets.

In chapter 2, I demonstrate for the first time the ability of surface-induced dissociation to distinguish between interfacial and non-interfacial phospholipid binding. Additionally, I show that combining this method with the modern high sensitivity instrumentation and statistical analysis enables the observation of subtle differences in oligomeric stabilities of membrane proteins in the presence of different species of interfacial lipids.

Chapter 3 describes application of native mass spectrometry to the investigation of a selected set of novel and established detergents. Controlled delipidation of membrane proteins is readily achieved by selecting a specific detergent from this set. Furthermore, a protocol for selective removal of either phospholipids or lipopolysaccharide with the aid of these detergents is presented. These findings contribute both towards the practical applications to membrane protein research and also towards a better fundamental understanding of detergents in general.

Chapter 4 describes the use of native mass spectrometry to probe the effects of lipid binding on a horse Na+/H+ transporter, NHE9. I show, that in the absence of specific lipid interactions NHE9 is unable to maintain its dimeric form. These results complement the recently obtained high-resolution structure of this protein.

Finally, chapter 5 illustrates investigation of the human MAPEG protein family. I use collision-induced dissociation experiments to demonstrate the preference of one member of this family, LTC4S, for particular phospholipid classes. I also identify appropriate native mass spectrometry conditions for another member of this family, MGST2, and link its instability to the removal of structurally important lipids by membrane mimetics.

Statement of Originality

This thesis describes research conducted between October 2016 and September 2020. The results reported are all tied to a common theme, with each chapter focusing on a distinct project. All the data presented were acquired and analysed by myself unless explicitly stated in the text. Each project involved some degree of collaboration, with a dedicated acknowledgments section at the end of each chapter further detailing specific contributions. All figures can be assumed to be of my own making, unless otherwise stated in the caption. I have opted for the singular pronoun 'I' in this thesis to highlight my own contributions to each of the projects, not to devalue the work performed by other people.

This work has not been submitted for a degree, diploma, or any other academic certification at any other institution, in whole or in part.

Denis Shutin

November 2020

Acknowledgments

I am grateful to the many people who have made the last four years such a wonderful experience.

First and foremost, I would like to thank Carol Robinson for giving me an opportunity to pursue my scientific endeavours as part of her group. She has always supported and encouraged me along the way and really helped me develop academically and personally over the years.

I am extremely grateful to Joseph Gault for mentoring me from my very first day as a green part II student that could not operate a biological pipette and all the way to the present. The roots of a large part of this work can be traced to the many insightful conversations we had over the years and I am forever indebted to him for his never ending support throughout my time as a graduate student (as well as being literally indebted to him for all the beer he bought me; I don't expect anyone to ever offer me a 6£ half-pint again in my life).

I would like to thank Leonhard Urner for his wisdom, positivity and great sense of humour, which made lab work a much more enjoyable experience. I have learnt a lot from him during our shared experiments as well as the many conversations both in and out of the lab; he is a great colleague and a true friend.

I thank the people at Waters, especially Kevin Giles and Jakub Ujma, for providing me the opportunity to use their development spectrometers for my research – a major part of this work would be impossible without them.

I would like to thank Vicki Wysocki and the members of her lab (in particular Sophie Harvey) for their help with surface induced dissociation methodology and instrumentation and for providing the SID cell and coating material.

I am grateful to Joseph Gault, Jani Bolla, Mark Agasid and Hsin-Yung (Jason) Yen for teaching me many aspects of molecular biology; to Andy Baldwin and Tim Allison for their help with software; to Kevin Giles, Jakub Ujma, Joseph Gault, Todd Mize and Andy Dolan for their practical and theoretical help with mass spectrometry instrumentation (one could say that these people were instrumental in this work). I would also like to thank Matt Bush for taking his time to give me advice on how to fix the 'linear' instrument during the uncertain times of the early pandemic – I am sure we would have got to the bottom of it if we were not interrupted by the lockdown.

I would like to thank the many collaborators that I have worked with over the years, including Patrick Rabe, Jos Kamps, Leah Taylor-Kearney, Madhuranayaki Thulasingam, Jesper Haeggström, Michael Landreh, David Drew and Iven Wilkenmann.

Thanks to all the fellow students and postdocs who made my DPhil such an enjoyable time both in and out of the lab, including (in no particular order) Joseph Gault, Leonhard Urner, Daniel Quetschlich, Francesco Fiorentino, Fabian Soltermann, Siyun Chen, Dominik Saman, Mark Agasid, Jani Bolla, Dror Chorev, Tarick El-Baba, Corinne Lutomski, Shane Chandler, Miranda Collier, Max Hantke, Matteo Degiacomi, Michael Landreh, Idlir Liko, Jon Hopper, Tim Allison, Anna Howes, Zihao Wang, Stephen Ambrose, Joseph Donlan, Xingyu (Phil) Qiu, Jason Yen, Ziwen Xu, Lars Sørensen, Haiping Tang, Andy Dolan, Joshua Sauer, Paul Fremdling and everyone else I had the pleasure working with. I would like to give special thanks to Vicky Cullen and Andy Dolan for their efforts in the lab organisation, especially during the chaos and turmoil of pandemic.

I would like to thank all of those friends outside of the lab; I especially thank Anatoly Smirnov for all the fun times and continuous moral support.

I owe a lot to my parents for their never ending support throughout my life. They have always created every opportunity for me to find success and I wouldn't be where I am without them (in more than one sense).

Finally, I would like to thank Mark Agasid, Francesco Fiorentino, Fabian Soltermann and Leonhard Urner for their review of this work – I am grateful beyond words for your help.

List of Abbreviations

AA	Ammonium acetate
AFM	Atomic force microscopy
AmtB	Ammonium transporter
ANOVA	Analysis of variance
AqpZ	Aquaporin Z
ATD	Arrival time distribution
BME	B-mercaptoethanol
BS	Biospin
C8E4	Tetraethylene glycol monooctyl ether
C8E6	Hexaethylene glycol monooctyl ether
CAPTR	Cation-to-anion proton transfer reactions
CCS	Collision cross section
CDL	Cardiolipin
CE	Collision energy
CEM	Chain ejection model
CG	Coarse-grained
CHS	Cholesteryl hemisuccinate
CID	Collision-induced dissociation
CIU	Collision-induced unfolding
CMC	Critical micelle concentration
CRM	Charged residue model
Cryo-EM	Cryogenic electron microscopy
CTD	C-terminal domain
CV	Column volume
DDM	*n*-dodecyl-β-d-maltopyranoside

DESI	Desorption electrospray ionisation
DO	1,2-dioleoyl
DOF	Degrees of freedom
DT	Drift tube
E_{lab}	Laboratory frame collision energy
E.Coli	*Escherichia coli*
ECD	Electron capture dissociation
EPR	Electron paramagnetic resonance
ERMS	Energy-resolved mass spectrometry
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FAB	Fast atom bombardment
FC	Foscholine
FLAP	Five-lipoxygenase-activating protein
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
GSH	Glutathione
HCD	Higher-energy C-trap dissociation
HDX	Hydrogen-deuterium exchange
IEM	Ion evaporation model
IM	Ion mobility
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β- d-1-thiogalactopyranoside
IST	In-source trapping
K_d	Equilibrium dissociation constant
KDO	3-deoxy-D-*manno*-oct-2-ulosonic acid

KLA	KDO_2-lipid A
LC	Liquid chromatography
LDAO	*N*,*N*-dimethyldodecylamine-*N*-oxide
LILBID	Laser induced liquid bead ion desorption
LPS	Lipopolysaccharide
LT	Leukotriene
LTC4S	Leukotriene C4 synthase
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionisation
MAPEG	Membrane-associated proteins in eicosanoid and glutathione metabolism
MD	Molecular dynamics
MGST2	Microsomal glutathione S-transferase 2
MP	Membrane protein
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MscL	Mechanosensitive channel of large conductance
MSP	Membrane scaffold protein
MW	Molecular weight
MWCO	Molecular weight cut-off
nESI	Nanoelectrospray ionisation
NG	n-nonyl-β-d-glucopyranoside
NHE	Na^+/H^+ exchanger
NMR	Nuclear magnetic resonance
OG	*n*-octyl-β-d-glucopyranoside
OGD	Oligoglycerol detergent
PA	Phosphatidic acid

PAGE	Polyacrylamide gel electrophoresis
PC	Phosphatidylcholine
PDB	Protein database
PE	Phosphatidylethanolamine
PEG	Polyethylene glycole
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PIP	Phosphoinositide
PL	Phospholipid
PO	1-Palmitoyl-2-Oleoyl
PS	Phosphatidylserine
PTM	Post-translational modification
RF	Radio frequency
S/N	Signal-to-noise ratio
SaBRe	Salt bridge rearrangement
SAM	Self-assembled monolayer
SASA	Solvent accessible surface area
SDS	Sodium dodecyl sulphate
SEC	Size-exclusion chromatography
SID	Surface-induced dissociation
SMALP	Styrene maleic acid lipid particle
SRIG	Stacked ring ion guide
SS	SemiSWEET
TM	Transmembrane
TOF	Time-of-flight
TSPO	Translocator protein (18 kDa)
TW	Travelling wave

UVPD	Ultra-violet photo-dissociation
WH	Wave height
WT	Wild type
WV	Wave velocity

List of Figures and Tables

Figure 1.1 Membrane protein types 3
Figure 1.2 Lipid types 5
Figure 1.3 Membrane protein-lipid interactions 8
Figure 1.4 Native mass spectrometry experiments 14
Figure 1.5 Membrane mimetics 21
Figure 1.6 Instruments 28
Figure 1.7 Electrospray fundamentals 30
Figure 1.8 Quadrupole mass filter 36
Figure 1.9 Time-of-flight mass analyser 39
Figure 1.10 Ion mobility spectrometry 42
Figure 1.11 Collision-induced unfolding 46
Figure 1.12 Surface induced dissociation 50
Figure 2.1 Collision-induced unfolding and surface-induced dissociation . 72
Figure 2.2 AmtB 73
Figure 2.3 SemiSWEET 74
Figure 2.4 Mass spectra of AmtB 77
Figure 2.5 SID of AmtB 80
Figure 2.6 SID of AmtB bound to POPG 81
Figure 2.7 SID-ERMS plot of AmtB 83
Figure 2.8 SID-ERMS of SS on Synapt G2 85
Figure 2.9 SID-ERMS plot of wild type and mutant SS on Synapt G2 87
Figure 2.10 CIU of SemiSWEET 89
Figure 2.11 Spectra of SemiSWEET obtained on the 'high-res' Synapt 92
Figure 2.12 In-source unfolding of SemiSWEET 93
Figure 2.13 SID of SemiSWEET on the 'high-res' Synapt 95
Figure 2.14 SID of SemiSWEET in the presence of various lipids 98
Figure 2.15 SID of SemiSWEET in the presence of cardiolipin 100
Figure 2.16 Statistical analysis of the SID-ERMS data of CDL-bound SS .. 103
Figure 2.17 Statistical analysis of the SID threshold energies of CDL-bound SS. 109
Table 2.1 Degrees of freedom correction factors 110
Figure 2.18 Molecular dynamics simulations of CDL-bound SS 111
Figure 2.19 Effect of impurities on SS spectra 115
Figure 2.20 Charge stripping considerations 118

Figure 2.21 Dimer percentage calculations. 128
Figure 3.1 Detergents 139
Table 3.1 Molecular weights of detergents 142
Figure 3.2 HLB and packing parameter 143
Figure 3.3 AmtB spectra 146
Figure 3.4 Phospholipid removal by detergents. 147
Figure 3.5 Lipopolysaccharide removal by detergents 151
Figure 3.6 Removal of LPS from AcrB by detergent 1 154
Figure 3.7 Removal of LPS from BtuCD by detergent 1 156
Figure 3.8 SDS-PAGE of AcrB and BtuCD 158
Figure 3.9 Effect of LPS binding on BtuCD 159
Table 3.2 Relative weight percentages of E4 in detergent structures 161
Figure 3.10 Average charge state of proteins in different detergents 163
Figure 3.11 Average charge state and delipidation of AmtB in different detergents 164
Figure 4.1 NHE orthologues 176
Figure 4.2 Cryo-EM structure of NHE9 178
Figure 4.3 Native MS of NHE9 in the presence of brain lipids 180
Figure 4.4 Native MS of NHE9 in the absence of brain lipids 181
Figure 4.5 Native MS of NHE9 with two monomeric species 185
Figure 4.6 Thermal shift assay of NHE9 187
Figure 4.7 Native MS of NHE9 with the addition of PIP2 189
Figure 4.8 Native MS of NHE9 triple mutant 192
Figure 4.9 NHE9 elevator access mechanism 194
Figure 5.1 Cysteinyl Leukotrienes 201
Figure 5.2 Mass spectrum of LTC4S in C8E4 obtained on UHMR 204
Figure 5.3 Crystal structure of LTC4S 205
Figure 5.4 Collision-induced unfolding of LTC4S 207
Figure 5.5 Solution-based activity assay of LTC4S 210
Figure 5.6 Mass spectra of MGST2 in C8E4 obtained on UHMR 211
Figure 5.7 Heterogeneity of MGST2 213
Figure 5.8 Mass spectra of MGST2 in C8E6 obtained on UHMR 215
Figure 5.9 Mass spectra of MGST2 in [G1]-mix obtained on UHMR 217
Figure 5.10 Mass spectra of MGST2 in DDM-C8E4 mixed micelle 218

Table of Contents

Abstract ..... **iii**
Statement of Originality ..... **v**
Acknowledgments ..... **vii**
List of Abbreviations ..... **ix**
List of Figures and Tables ..... **xiv**
Table of Contents ..... **xvii**
1 Introduction **1**
1.1 Membrane Proteins 1
1.2 Membrane Lipids 3
1.3 Membrane Protein – Lipid Interactions 6
1.4 Techniques Used to Investigate Membrane Protein – Lipid Interactions 9
1.5 Native Mass Spectrometry 12
1.6 Using Native Mass Spectrometry to Investigate Membrane Protein-Lipid Interactions 20
1.7 Native Mass Spectrometry Instrumentation 27
1.8 Ionisation 29
1.9 Transmission of High Molecular Weight Ions 34
1.10 Quadrupole Mass Filter and Tandem MS 35
1.11 TOF Mass Analyser 37
1.12 Orbitrap Mass Analyser 38
1.13 Ion Mobility Spectrometry 40
1.14 Collision Induced Unfolding and Dissociation 44
1.15 Surface Induced Dissociation 48
References 53
2 Application of Surface-Induced Dissociation to Detect Interfacial Lipids for Membrane Protein Complexes by Native Mass Spectrometry **68**
Abstract 68
2.1 Introduction 69
2.2 Experimental Outline 71
2.3 Surface-Induced Dissociation of AmtB 75
2.4 Surface-Induced Dissociation of SemiSWEET on Synapt G2 83
2.5 Collision-Induced Unfolding of SemiSWEET 86
2.6 Surface-Induced Dissociation of SemiSWEET on the ‘High Resolution’ Synapt 91
2.7 Molecular Dynamics Simulations of SemiSWEET 110

2.8 Potential Issues for SID-ERMS of Membrane Proteins 113
2.9 Summary and Conclusions 121
Materials and Methods 124
Acknowledgments 131
References 132
3 Development of Controlled Delipidation Methods for Native Mass Spectrometry Characterisation of Membrane Protein Complexes 136
Abstract 136
3.1 Introduction 137
3.2 Detergent Design 139
3.3 Phospholipid removal 142
3.4 Lipopolysaccharide removal 149
3.5 Charge Reducing Properties of Detergents 160
3.6 Summary and Conclusions 162
Materials and Methods 166
Acknowledgments 170
References 171
4 Investigating Lipid-Controlled Dimerisation of a Mammalian Ion Exchanger by Native Mass Spectrometry and Complementary Biophysical Methods 174
Abstract 174
4.1 Introduction 175
4.2 Obtaining High Resolution Cryo-EM Structure of NHE9 176
4.3 Native Mass Spectrometry Reveals the Role of Lipids for NHE9 Dimerisation 179
4.4 Summary and Conclusions 191
Materials and Methods 194
Acknowledgements 195
References 196
5 Identifying Specific Protein-Lipid Interactions for Human Leukotriene Biosynthetic Enzymes by Native Mass Spectrometry 199
Abstract 199
5.1 Introduction 200
5.2 Characterisation of LTC4S by Native MS 202
5.3 Characterisation of MGST2 by Native MS 209
Summary and Conclusions 217
Materials and Methods 219
Acknowledgments 221
References 222

Concluding Remarks **224**

1 Introduction

1.1 Membrane Proteins

Biological membranes are essential for life, enclosing and compartmentalising cells. All membranes have the same basic structure, with amphipathic lipid molecules spontaneously forming bilayers (with hydrophobic parts on the inside and hydrophilic segments facing the aqueous environment on the outside)[1]. Despite a similar overall structure, biological membranes can vary significantly in exact composition, due to a large diversity of constituent lipids[2]. In addition, membranes contain numerous proteins, with the ratio (by weight) of protein to lipid being close to 1 for most membranes[3].

Membrane proteins form around a quarter of the human proteome[4] and represent approximately 60% of drug targets[5]. They can be divided into two main groups in terms of their interaction with the lipid bilayer: integral membrane proteins have at least one segment inserted directly into the hydrophobic core of the membrane, while peripheral membrane proteins only interact with hydrophilic residues on the outside and are weakly associated with the membrane (**Figure 1.1**). When the term 'Membrane Protein (MP)' is encountered in this thesis it refers specifically to integral membrane proteins unless stated otherwise. Integral membrane proteins can be further separated into three groups: monotopic MPs are exposed to only one side of the lipid bilayer, bitopic MPs span the bilayer exactly once and polytopic MPs span the bilayer multiple times, with multiple hydrophilic segments on both sides of the bilayer[6]. Such membrane-spanning segments are most commonly hydrophobic α-helices, although the proteins located in the outer membranes of Gram-negative bacteria usually form β-barrels[7].

MPs can also be divided into several major families by their function, including transporters, channels, receptors and enzymes (such as proteases)[8-10]. Transporters are responsible for moving many different types of water-soluble molecules in and out of the cell. Transporters undergo a structural change upon the binding of the substrate, resulting in selective movement of molecules across the bilayer[11]. They are classified into three groups: uniporters transport a single substrate across the membrane along the diffusion gradient, while symporters and antiporters move a different ionic molecule in the energetically favoured direction to transport the substrate against the activity gradient (what is referred to as secondary active transport)[12]. Antiporters move such enabling molecules in the opposite direction to the substrate, while symporters do it in the same direction.

Ion channels and porins form pores that can be either closed completely or open to both sides of a membrane at same time. They can selectively move molecules at up to 10^5 times greater rate than the fastest of transporters, but can only mediate 'passive' transport (down the diffusion gradient)[13]. Ion channels often default to the closed state, with opening facilitated by a stimulus, such as a change in voltage across the membrane, mechanical stress or ligand binding.

The largest signalling protein family is G protein-coupled receptors (GPCRs), which respond to a large diversity of stimuli, including light, odorants, and small molecules and ions such as calcium[14]. GPCRs share the same basic structure with 7 trans-membrane helices, but have many diverse subfamilies and low-sequence identity[15,16]. Due to the numerous signalling pathways that they control, GPCRs are very attractive drug targets[17]. Another MP family which is appealing for drug targeting is membrane proteases, which can generate cleavage products on both sides of the bilayer to control in-cell and between-cells signalling[18].

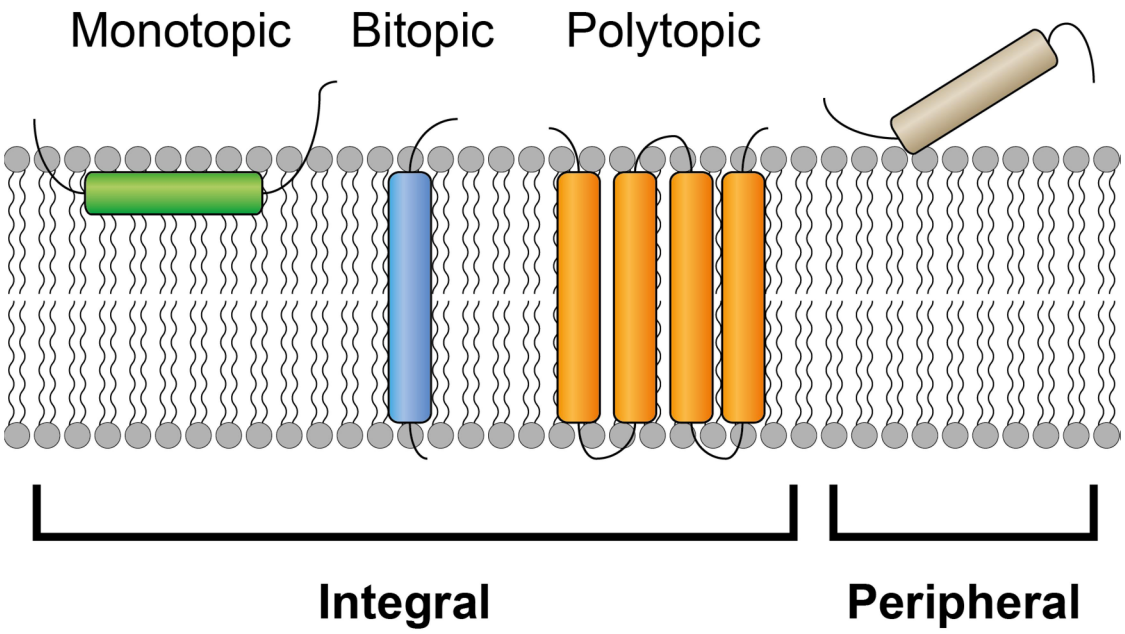


Figure 1.1 Membrane protein types. Integral membrane proteins are located inside the membrane and can be monotopic (green), bitopic (blue) or polytopic (orange). Peripheral membrane proteins (pink) only interact with the surface of the bilayer.

One unifying feature for all of these diverse MPs is the influence of their lipid environment on their structure and function. Various types of lipids and their interactions with MPs will be explored in the next sections.

1.2 Membrane Lipids

Lipids are loosely defined as molecules that are insoluble in water but soluble in many organic solvents, although more accurate and detailed definitions have been proposed based on their biological origin[19]. They can perform three general functions: be used as an efficient energy storage, act as messengers in diverse signalling pathways and, as was mentioned in the previous section, can be the building blocks of biological membranes. All membrane lipids must be ampipathic in order to self-assemble into bilayers, however there are many differences between the lipid species as a large diversity of them exists. Membrane lipids can

be classified into four main categories based on their structure: glycerophospholipids, sphingolipids, sterols and glycolipids[8,20].

Glycerophospholipids are the major structural lipids in most membranes, so they are often referred to simply as phospholipids (PLs)[8,21]. PLs are based on a glycerol moiety connected to two acyl fatty acid chains and a phosphate group, which can be modified by several functional groups (head groups) forming different PL classes (**Figure 1.2a**). The common membrane PLs include phosphatidic acid (PA), phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidylserine (PS) and phosphatidylinositol (PI). PI can also be phosphorylated on some or all of positions 3, 4 and 5 (**Figure 1.2a**), forming various phosphoinositides (PIPn, where n is the number of phosphorylation events). Another notable PL is cardiolipin (CDL), which has a PG molecule as a head group, resulting in a lipid with four acyl chains and two phosphate groups. There is even further PL diversity, because lipids with the same head group can differ by having different fatty acid chains at the tail. These chains can vary in length (the number of carbons) and also in the number and positions of double bonds (generally, saturated or mono-unsaturated fatty acids occupy the *sn*-1 position, while unsaturation is more common at the *sn*-2 position)[2,22].

Sphingolipids are based on a sphingosine molecule, which is connected to a fatty acid by an amide bond. They also have different head groups, which can be simple (such as hydroxyl for ceramide) or contain a modified phosphate group similar to PLs, (for example, sphingomyelin contains a phosphocholine head group) (**Figure 1.2b**).

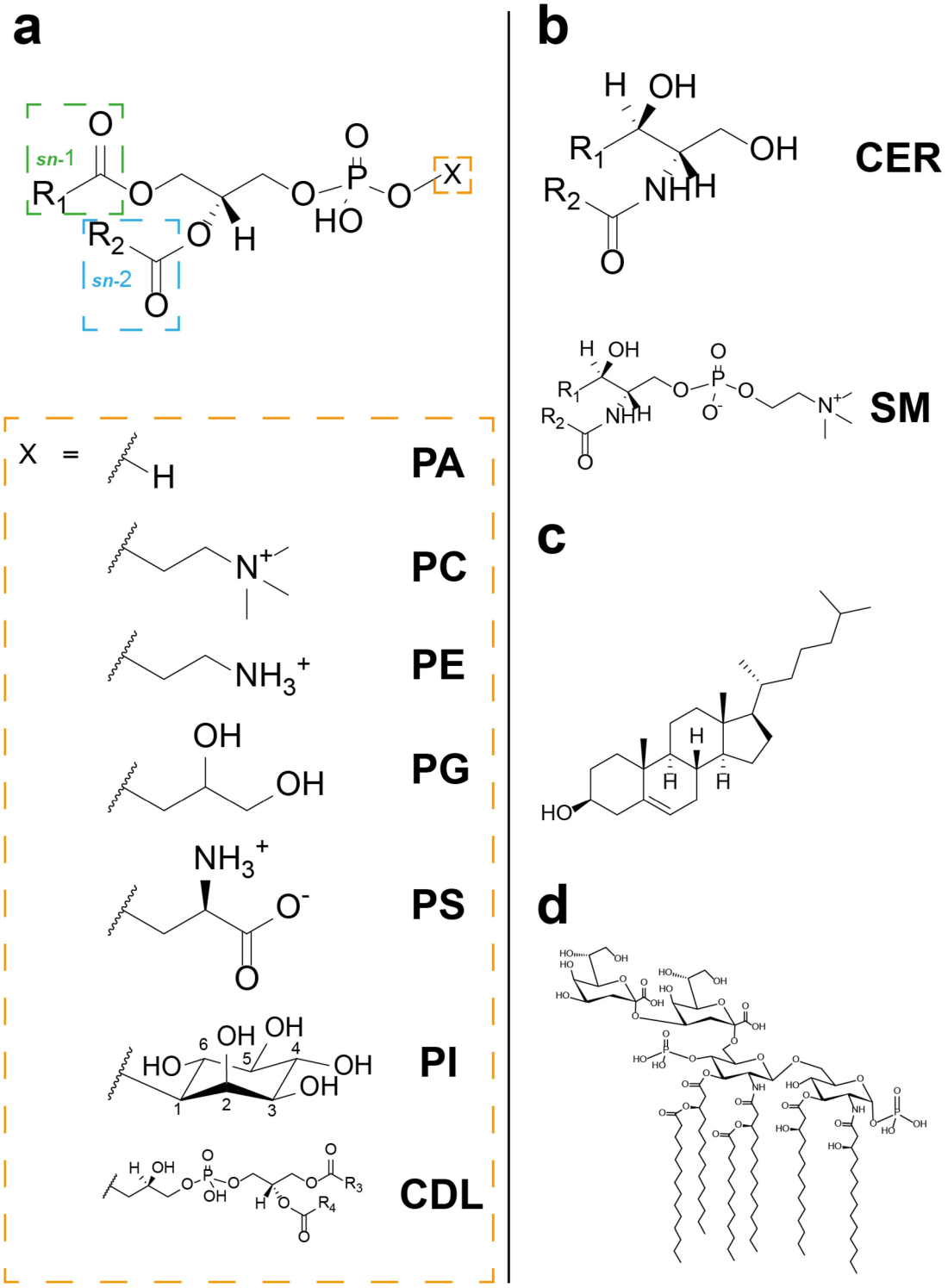


Figure 1.2 Lipid types. a, Phospholipids**. b,** Sphingolipids. **c,** Cholesterol. **d,** Core LPS (Lipid A-KDO_2 complex). Abbreviations: PA, phosphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PS, phosphatidylserine; PI, phosphatidylinositol; CDL, cardiolipin; CER, ceramide; SM, sphingomyelin; LPS, lipopolysaccharide; KDO, 3-deoxy-D-*manno*-oct-2-ulosonic acid .

Sterols consist of four fused hydrocarbon rings with an alcohol group and a side-chain. Cholesterol is the most common sterol in mammals (**Figure 1.2c**).

Glycolipids are lipids that contain sugar. There is a large number of different glycolipids performing many functions, including acting as signal transducers, protein anchors and receptors of bacteria and bacterial toxins[23]. One important family of glycolipids is lipopolysaccharides (LPSs), which are located in the outer leaflet of the outer membrane of many gram-negative bacteria, including *Escherichia coli* (*E.coli*)[24]. LPS induces immune response in humans and other animals and is largely responsible for septic shock[24,25]. The structure of LPS consists of three elements: Lipid A, core oligosaccharide and O-antigen, with Lipid A and the inner core oligosaccharide (3-deoxy-D-*manno*-oct-2-ulosonic acid, KDO) highly conserved and required for cellular growth of most gram-negative bacteria (**Figure 1.2d**)[25,26].

In addition to chemical diversity, membranes also exhibit compositional diversity, with different lipids present in different organisms, cells, organelles and even the two leaflets of the same membrane[2,21,27]. Moreover, a particular lipid environment is modulated by various factors such as cell growth conditions, cell division or disease[28-30]. The variety of existing lipids clearly indicates that they are not simple building blocks for biological membranes, but have more complex roles and interactions with MPs. The next section will describe how lipids can influence MP structure and function.

1.3 Membrane Protein – Lipid Interactions

Membrane proteins and lipids do not simply coexist in biological membranes, but influence each other in multiple ways[31]. In terms of interaction with proteins, lipids

can be assigned into three groups: bulk lipids, annular lipids and non-annular lipids (**Figure 1.3**)[32]. Bulk lipids are not in contact with proteins and can diffuse throughout the bilayer. Annular lipids are in direct contact with the transmembrane regions of a membrane protein and got their name for forming an annulus or ring around the protein. While individual annular lipid molecules are in fast exchange with the lipids in the bulk, they can still be important for the MP function as one molecule replaces another in the first lipid shell and the protein senses the same environment on average[33]. A recent study using molecular dynamics (MD) simulations observed unique lipid environments for different MPs, highlighting the high degree of selectivity for the surrounding lipid shells[34]. Both the hydrophobic chains and the polar head groups can be important for the annular lipid selection, with factors such as hydrophobic thickness or the distribution of charges affecting specificity of the protein-lipid interactions[35]. There is an energetic penalty for the mismatch between the hydrophobic transmembrane part of the protein and the acyl chains of the surrounding lipids, which can result in protein conformation being altered, with deformations such as tilts or bends[36]. One example of the importance of the acyl chain length for annular lipids is Ca^{2+}-ATPase protein, where the chain lengths of PC species affect both activity and the rates of phosphorylation and dephosphorylation[33,37]. The polar head groups of annular lipids can regulate membrane proteins by interacting with the amino acids of the opposite charge at the binding sites. An example of such regulation is mechanosensitive channel of large conductance (MscL) from *E.coli,* which requires anionic lipids for correct function[38,39].

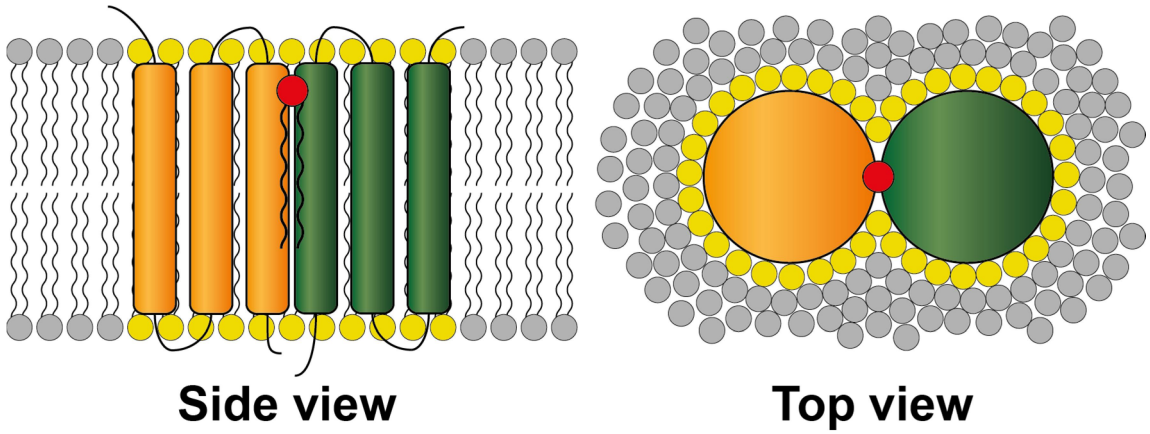


Figure 1.3 Membrane protein-lipid interactions. A dimeric membrane protein interacts with the surrounding lipids. Bulk lipids (grey) are not in direct contact with the protein. Annular lipids (yellow) form a protective ring around the protein and are in fast exchange with bulk lipids. Non-annular lipid (red) is stabilising the dimer formation.

Non-annular lipids bind more tightly to MPs, often at highly specific binding sites between transmembrane helices of a monomer or between subunits of an oligomeric complex[32]. These lipids perform multiple roles and can be essential for membrane protein structure or function. Non-annular lipids found between monomeric subunits in a membrane protein complex are often responsible for maintaining the correct oligomeric state[40]. While many diverse lipid species perform this function for different MPs, CDL is particularly well-suited for this purpose due to its unique dimeric structure, containing four acyl chains and two negative charges – CDL is sometimes referred to as 'molecular glue' for its ability to stick monomers together[41,42]. Non-annular lipids can also be necessary for MP function; for example, a bacterial potassium channel KcsA requires binding of anionic lipids to facilitate channel opening[43]. Another role of non-annular lipids is as allosteric effectors; for example, cholesterol modulates G-protein coupling and ligand binding for a member of GPCR family called serotonin$_{1A}$ receptor[44]. Similarly to annular lipids, MPs can select for either specific head groups or acyl chains of

non-annular lipids to fit into a particular binding site, with some examples of a single lipid species being required for MP oligomerisation[45].

There is also a distinct class of MP-lipid interactions, where the lipid acts as the actual substrate for a particular MP. An example of such interaction comes from lipid flippases, which transfer lipids across the membrane[46].

With all of these important MP-lipid interactions existing, a variety of biophysical techniques have been developed to study them and they will be reviewed in the next section.

1.4 Techniques Used to Investigate Membrane Protein – Lipid Interactions

Most of the methods applicable to MPs produce the best quality of data when performed on pure and homogeneous samples. As MPs are not water soluble, various membrane mimetics, such as detergent micelles, are employed to extract the protein of interest from its native membrane environments and solubilise it. Membrane mimetics compatible with native mass spectrometry are reviewed in more detail in **Section 1.6**.

One of the most widely used techniques for determining MP structures and their interactions with lipids is X-ray crystallography. The structures obtained at high resolution allow direct identification of lipid binding sites locations, as well as any structural rearrangements caused by lipid binding[47]. However, the number of MP crystal structures remains low compared to their soluble counterparts due to a number of challenges, such as low expression yields and instability outside of native membrane for MPs, as well as the difficulty of obtaining well-diffracting

crystals[48]. In addition, dynamic elements, such as soluble unstructured loops, are often not resolved, which could lead to missing structurally important elements[49]. Several challenges are also associated with identifying bound lipids, with weaker-bound species not surviving purification process or not being unambiguously characterised due to their highly dynamic nature[35]. This limits X-ray crystallography to probe only the most specific and strong binders in most cases (although even transiently-bound annular lipids have been co-crystallised with MPs in rare cases[32]).

Another method for high resolution protein structure determination is electron microscopy. Traditionally, electron 2D crystallography was the technique of choice[50], however single particle cryogenic electron microscopy (cryo-EM) is rapidly growing in popularity due to recent technological breakthroughs[51]. Single particle cryo-EM enables identification of high resolution protein structures without the need for crystallisation, which makes it the preferred structure determination method for many MP families[52]. Recent examples of single particle cryo-EM applied to MPs include determining structures of GPCRs coupled to G-proteins[53,54] and elucidating specific protein-lipid interactions for the heat- and capsaicin-activated ion channel, TRPV1[55]. The current challenges for obtaining high resolution structures with this method include the size limit (with the minimum protein size roughly defined to be 100kDa), sample heterogeneity and the high concentration of solutes in the buffer (for example, salt and glycerol which are often used during MP purification, as well as membrane mimetics such as detergent micelles)[56].

While the methods above are extremely useful for obtaining high resolution protein structures, they provide little or no information about protein dynamics. Nuclear

magnetic resonance (NMR) spectroscopy is an alternative method that can yield high-resolution protein structure as well as dynamics information. Solution NMR spectroscopy has until recently been limited to small proteins (roughly below 30-40 kDa in mass) due to slow tumbling times of big molecules, and, accordingly, also required the use of detergents that form small micelles[57,58]. In particular, foscholine (FC) 12, also known as dodecyl phosphocholine (DPC), has been extensively used for solution NMR, accounting for approximately 40% of the structures produced by this technique; however, there is substantial evidence of FC12 having destabilising and denaturing properties for many MPs[59]. The latest advances in technology and methods allow using solution NMR to probe dynamics of large proteins well above 100 kDa in mass[60,61]. Alternatively, solid state NMR can be employed to study MPs of various sizes in a detergent-free environment (for example, inserted into oriented or unoriented lipid bilayers)[57].

In addition to the aforementioned methods, various other biophysical techniques have been successfully applied to investigate protein-lipid interactions, including atomic force microscopy (AFM), Förster resonance energy transfer (FRET) and electron paramagnetic resonance (EPR). AFM uses a sharp tip to probe the surface of an MP and can be performed on the protein in its native membrane environment[62]. FRET allows investigating MP-lipid interactions by measuring the extent of energy transfer from the protein to different types of fluorescently labelled lipids[63]. EPR uses spin-labelled lipids to distinguish between bulk lipids and those in contact with the protein surface[64].

Computer technology is constantly improving at a rapid pace, meaning that MD simulations become increasingly popular in complementing experimental findings, or even performing original investigations of MP-lipid interactions[65-67]. Two main

types of force fields are used for the majority of MD simulations: atomistic scale and coarse-grain (CG) scale force fields. Atomistic scale force fields provide greater accuracy and resolution but are really computationally expensive and therefore limited in terms of the size of the system and the simulation time; CG scaling involves combing a group of atoms into a single 'bead', significantly reducing computation cost at the expense of resolution[68]. MD simulations can also yield quantitative information by estimating the energetics of protein-lipid interactions[69].

Native mass spectrometry (MS) is a complementary approach to the methods above that has made substantial contributions to the field of MP-lipid interactions over the past decade[8]. As native MS is the main technique used in this thesis, it will be explored more extensively in the following sections. The next section will provide an overview of native MS, describing common experiments and highlighting various strengths and limitations of this method.

1.5 Native Mass Spectrometry

Mass spectrometry is a technique that measures mass-to-charge ratios (*m/z*) of molecules in the gas phase. Throughout its early history, MS had been primarily used for the studies of small volatile molecules. In order to extend MS capabilities to larger molecules, such as peptides, novel ionisation techniques were developed, with fast atom bombardment (FAB) being an important early example[70,71]. Later, the revolutionary development of soft-ionisation techniques, most notably matrix-assisted laser desorption ionisation (MALDI)[72] and especially electrospray ionisation (ESI)[73] in the late 1980's, allowed transmission of large intact biomolecules into the gas phase, thus paving the way for the field of native MS[74].

Native MS differs from the other biological MS methods (such as bottom-up proteomics) by using non-denaturing buffers to maintain the native state of a protein prior to ionisation and then preserving any non-covalent interactions and (to a large extent) the native fold of the protein in the vacuum of a mass spectrometer[75]. This enables identification of stoichiometry of protein complexes, their binding partners, such as ligands or lipids, and also provides information on topology and dynamics[76]. Native MS does not require sample labelling or homogeneity and can be applied to proteins with a wide range of masses and structures, making it a valuable addition to a structural biologist's toolkit. Another advantage of native MS is its sensitivity, with as little as 2-3 μL of protein at 1-10 μM concentrations required for an experiment[77] – this can be especially important for proteins that are challenging to express in large quantities, such as many MPs.

Native MS experiments can be performed on instruments with various architectures (see **Section 1.7** for the types of spectrometers used in this work), but all of them share several key features (**Figure 1.4a**). The first step is ionisation, where the protein is transferred from solution into the gas phase. For native MS, the most commonly used ionisation method is a miniaturised version of ESI called nanoelectrospray ionisation (nESI)[78] (see **Section 1.8** for a more detailed description of nESI and alternative ionisation techniques). nESI produces a series of multiply charged ions, resulting in a characteristic Gaussian-like distribution of peaks in the spectra (**Figure 1.4b**). Some of the evidence for proteins maintaining their native-like structures during native MS comes from the observation that the average charge state for the protein is significantly lower than for the same protein sprayed from denaturing conditions[75,79]. This implies that the solvent accessible surface area (SASA) of the protein is smaller, leading to less charging and implying

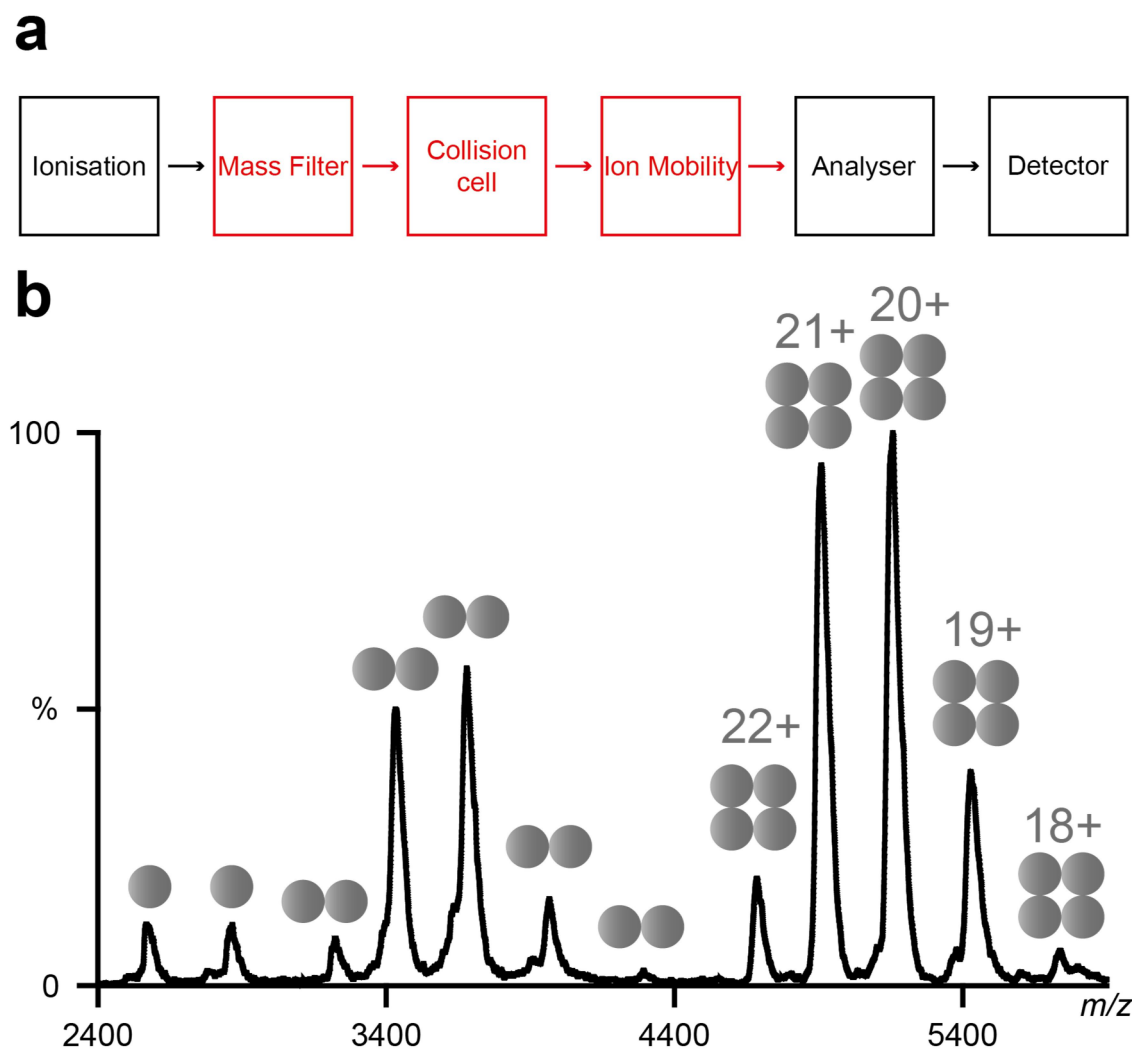


Figure 1.4 Native mass spectrometry experiments. a, A schematic representation of a general mass spectrometer, including essential (black) and optional (red) elements. **b,** An example of a native spectrum of a protein (concanavalin A), highlighting several native mass spectrometry features. Different oligomeric species can be observed in the same spectrum. For each species, a series of charge states is formed (indicated for the tetramer), resulting in a roughly Gaussian distribution – exact mass can be calculated from these patterns. The peaks have large widths due to incomplete removal of salt adducts.

that the protein remains in the folded state[80] (see **Section 1.8** for further discussion).

The most important stage of any mass spectrometer is the actual mass analyser, where the *m/z* of the molecule of interest is measured (**Figure 1.4a**). The two most commonly used types of analyser for intact protein analysis are time-of-flight (TOF)

and Orbitrap[81] (**Sections 1.11** and **1.12** respectively). Ideally, when the protein reaches the analyser, it has been stripped of any unwanted binders (such as solvent molecules, salt adducts or remaining detergent) in order to extract the accurate mass value. Moreover, the signal for a single proteoform is spread out over species with different number of such adducts, significantly reducing signal-to-noise ratio (S/N)[82]. Salt adducts in particular present a problem for native MS because they are not readily removed by gas-phase activation – as a result, proteins have to be buffer exchanged into a volatile salt solution, typically ammonium acetate (AA) at 0.1-1M concentration[77,83] – ammonia and acetic acid readily evaporate, leaving a 'naked' protein to be analysed[84]. AA has an additional advantage of having a pH value close to 7, which contributes to keeping most proteins in their folded state – although AA is technically not a buffer at pH 7, so it does not necessarily protect the protein throughout nESI process[85]. There are not many alternatives to AA for native MS, with examples like ammonium bicarbonate showing evidence of causing protein unfolding by bubble formation[86] – this can sometimes be an issue for proteins unstable in acetate. Moreover, proteins are typically exposed to high concentrations of various nonvolatile salts in their native environment, which can often be required for structural or functional reasons; these interactions are, therefore, lost during a typical native MS experiment[87]. Recently, small diameter nESI emitters have been introduced, that allow performing native MS experiments from nonvolatile buffers at the expense of resolution[87,88].

Detection of an intact native-like protein complex enables measuring its accurate mass – this could help with determining the number of post-translational modifications (PTMs) or bound ligands, but provides little information about internal composition. The structure and binding of a protein complex can be probed more

deeply by employing dissociation techniques that selectively disrupt some of the noncovalent interactions, while keeping covalent bonds intact. The most commonly used dissociation techniques are collision-induced dissociation (CID), which uses multiple collisions with small neutral gas molecules to slowly activate the protein complex of interest, and surface-induced dissociation (SID), where the protein complex is collided into a solid surface[81] (these methods are reviewed in detail in **Sections 1.14** and **1.15** respectively). Breaking non-covalent interactions of an oligomeric protein complex produces smaller 'fragment' subcomplexes; these then have their masses measured, providing information about the original stoichiometry[89]. SID is particularly informative, producing dissociation products that are representative of the complex assembly and thus providing insight about the internal architecture[90]. Alternatively, some information about protein assembly and topology can be obtained by disrupting non-covalent interactions in solution by using organic solvents[91,92].

As described above, nESI produces multiply charged ions, meaning that several species are simultaneously present in a spectrum. This can often pose a challenge for interpretation of dissociation experiments as it is unclear which dissociation products came from which precursor. Moreover, several oligomeric states can be present in solution adding to the difficulty – for example, monomers present in solution and formed as a result of gas-phase activation would be indistinguishable upon detection (although they might have different charge states in some cases). This problem can be alleviated by performing a tandem MS (MS/MS) experiment, where a particular *m/z* is selected by the quadrupole mass filter (**Section 1.10**) before any activation is applied[93].

Native MS by itself is limited to measuring the mass of a molecule, but not its size. Ion mobility (IM) spectrometry (**Section 1.13**) allows the measurement of the rotationally-averaged collision cross section (CCS) of a protein[94] and can be readily combined with native MS to obtain size and mass information in a single experiment[95]. The measured CCS value can be compared to a theoretical value calculated from a crystal structure[96] to verify that the protein has maintained a native-like conformation in the gas phase – in fact, some of the early evidence that the protein structure can be preserved in the gas phase comes from a similar experiment[97]. A protein can be intentionally unfolded inside a mass spectrometer by using gas phase activation in a collision-induced unfolding (CIU) experiment. Specific lipid or ligand binding can stabilise the protein conformation, resulting in extra energy required to cause an unfolding event – thus CIU can be used to investigate protein-ligand interactions[98,99]. Alternatively, various unfolded intermediates for a protein can be generated in solution and then investigated by IM without further gas-phase manipulation[100,101]. Care has to be taken when interpreting IM data, as the protein can be collapsed or unfolded if the selected charge state is too low or too high respectively[102,103].

While native MS is able to detect protein-protein and protein-ligand binding events across a wide range of masses, several potential problems have to be kept in mind when interpreting experimental results. When two species vary significantly in mass (for example, two different oligomeric states of the same protein) they will have different probabilities of ionisation, transmission, and detection, so their intensities cannot be compared directly without taking those factors into account[104]. This does not generally apply to a small ligand (such as lipid) bound to a much larger protein measured against the *apo* protein as their molecular weights are

similar[105]. Equilibrium dissociation constants (K_d) for protein-ligand interactions measured by native MS were found to be in agreement with values obtained by isothermal titration calorimetry in solution[81].

Another possible source of error comes from the fact that the binding strengths vary between solution and the gas phase. In solution, weak hydrophobic interactions are enhanced by the high cost of mixing non-polar compounds with water – this is called the hydrophobic effect[106]. The hydrophobic effect is a dominant energy contribution to a variety of important biological processes, including formation of lipid bilayers[107] (and detergent micelles[108]), protein folding[109-111] and protein-hydrophobic ligand association[112]. The hydrophobic effect predominantly arises due to large entropic factors, as the presence of a hydrophobic molecule in water results in ordering of water around that molecule in order to maximise the hydrogen bonding network, and this ordering comes at a high entropic penalty[113] (although enthalpic contributions to the hydrophobic effect are also important[114,115], as is evidenced, for example, by denaturing of proteins not only at high, but also at low temperatures[109,116]). Absence of solvent in vacuum leads to hydrophobic interactions being weakened, to the extent that some of them are unable to be observed by MS[117]; electrostatic interactions, conversely, become stronger in the gas phase[118]. However, it should be noted that despite the absence of the hydrophobic effect in vacuum, protein structures similar to those in solution are often retained upon the transfer into the gas phase due to kinetic trapping on the ms timescale of a typical MS experiment[119].

ESI polarity can also lead to observing false negatives, as has recently been shown for some MP-lipid interactions being detected exclusively in positive or negative polarity[120]. False positives can be observed due to non-specific binding; these

usually occur at very high concentrations of ligands or proteins, when multiple molecules end up in the same ESI droplet (see **Section 1.8** for more details on ESI mechanism) – for that reason, protein concentration is usually kept low for native MS (roughly below 20 μM) to avoid formation of unnatural oligomeric states[121].

Compared to many of structural biology methods mentioned in **Section 1.4**, native MS is a low structural resolution technique – some prior knowledge of the nature of the protein as well as any binders is required, as it is impossible to identify them by mass alone. Native MS can be complemented with other MS methods to obtain more information about protein and ligand identities and interactions. Gas phase activation can be used to break covalent bonds and analyse the resulting fragments to characterise the protein – this process is called top-down proteomics[122]. Traditionally, top-down proteomics were performed on denatured proteins, but recent instrumental and methodologic advances have enabled native top-down MS[123,124]. High energy CID can be used to fragment the protein backbone, but other techniques are also employed including electron based electron capture dissociation (ECD) and electron transfer dissociation (ETD)[125] and ultra-violet photo-dissociation (UVPD)[126,127]. To investigate protein-ligand interactions another MS-based method is used called hydrogen-deuterium exchange (HDX). In HDX, proteins get labelled with deuterium in solution and then are digested at low pH to preserve the labelling – this allows identifying solvent exposed regions of the protein, and any differences observed upon ligand binding can be used to determine location of binding sites and induced conformational changes[128].

Native MS can also be combined with other structural biology techniques, such as X-ray crystallography[129] and MD simulations[130,131]. These methods benefit from the

ability of native MS to obtain information on the nature of protein-ligand interactions, in turn providing a high structural resolution picture unavailable to native MS. Another interesting avenue is preparative MS, where native MS is used to select a single protein species and soft land it onto a surface, generating a homogeneous sample for analysis by high structural resolution methods[132]. A combination of complementary techniques can therefore lead to a better understanding of a biological system than any of those methods could provide by themselves; this is the principle behind the integrative structural biology approach[133]. Throughout this thesis I adhere to this approach, with native MS results being accompanied with data obtained by various other methods, including MD simulations, cryo EM and solution-phase activity assays.

In summary, native MS can be used to investigate protein-ligand interactions in a selective, sensitive and label-free manner. The next section will describe the application of native MS specifically to study MP-lipid interactions, highlighting the recent breakthroughs and achievements.

1.6 Using Native Mass Spectrometry to Investigate Membrane Protein-Lipid Interactions

Membrane proteins are not soluble in water, so they have to be solubilised by a membrane mimetic in order to be amenable to ESI (**Figure 1.5**). A big breakthrough in the native MS of MPs came from the use of detergents (**Figure 1.5a**) at 2x critical micelle concentration (CMC), which were shown to preserve intact MP complexes in the gas phase[134]. Detergent micelles continue being commonly used for this purpose[77]. Detergents are also often employed to extract MPs from their native bilayer environment[135]. Detergent selection has to be

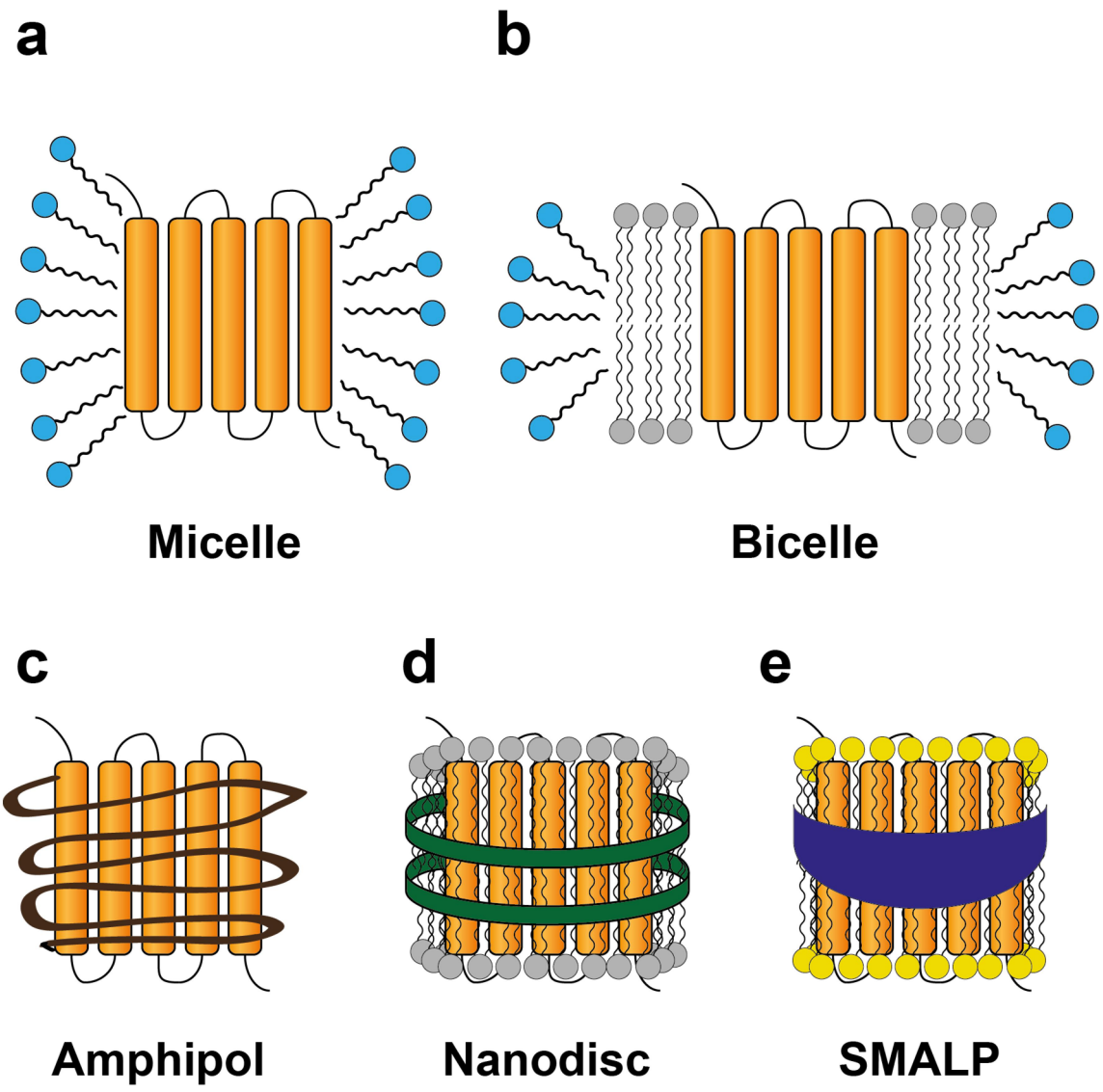


Figure 1.5 Membrane mimetics. A protein protected by various membrane mimetics: **a,** Detergent micelle (light blue); **b,** Bicelle, a combination of lipids (grey) and detergent; **c,** Amphipol; **d,** Nanodisc, two copies of MSP (green) surrounding lipids; **e,** SMALP, styrene maleic acid co-polymer (dark blue) extracts the protein together with the native lipids (yellow). MSP, membrane scaffold protein; SMALP, styrene maleic acid-lipid particle

carefully considered as it can have an effect on extraction efficiency and solution-phase or gas-phase stability. Saccharide-based detergents, such as *n*-dodecyl-β-d-maltopyranoside (DDM), are among the least denaturing ('mild') detergents and are extensively used for protein extraction and solubilisation[136]. During a native MS experiment, detergents have to be removed in order to measure the mass of the 'naked' protein; as a result, 'sticky' detergents like DDM require high amounts of

energy to be applied, which can in turn cause protein unfolding or even dissociation[47]. It can therefore be beneficial to use 'harsher' detergents for native MS, such as *N*,*N*-dimethyldodecylamine-*N*-oxide (LDAO), *n*-octyl-β-d-glucopyranoside (OG) and especially members of polyethylene glycole (PEG) family, particularly tetraethylene glycol monooctyl ether (C8E4)[47,137]. C8E4 can be easily removed with mild activation and also has charge reducing properties which can be beneficial for MP stability in the gas phase[137,138]. The large diversity of available detergents often leads to a time consuming screening step to identify the best detergent for a particular protein complex[139]. It is also not uncommon to extract the protein in one detergent (usually DDM) and then perform a detergent exchange into a 'native MS friendly' detergent by size-exclusion chromatography (SEC) or during a buffer exchange step using a Bio-Spin desalting column[77]. In some cases no single detergent exists that is simultaneously stabilising and native MS compatible – a mixed micelle approach can be used instead, where several detergents are combined (for example, a mixture containing DDM, cholesteryl hemisuccinate (CHS) and FC enabled native MS-based investigation of GPCRs[140]). A different way to address this issue is to design novel detergents with tailorable characteristics, such as the recently introduced oligoglycerol detergent (OGD) family[141]. In **Chapter 3** of this thesis, several novel detergents are investigated, focusing on their compatibility with protein purification and native MS.

One disadvantage of using detergent micelles is that they are very different from the native bilayer environment for MPs, missing factors such as lateral pressure as well as losing many important lipid interactions. Several 'more membrane-like' mimetics are compatible with native MS, including bicelles, amphipols, nanodiscs and styrene maleic acid lipid particles (SMALPs). Bicelles are bilayer discs made

up of lipids, with the hydrophobic edges of the bilayer protected by detergent molecules (**Figure 1.5b**)[142]. Bicelles were shown to preserve the protein oligomeric state in the gas phase[143], however they have not been extensively used in native MS since the original report[8].

Amphipols are flexible ampipathic polymers with high density of hydrophobic chains that strongly interact with MPs (**Figure 1.5c**)[144]. They require high energy to be removed in the gas-phase (compared to detergent micelles) which can cause dissociation of oligomeric complexes[143]. However, amphipols were shown to be useful for native MS of monomeric MPs[145].

Nanodiscs are nanoscale artificial lipid bilayers encircled by two membrane scaffold protein (MSP) belts (**Figure 1.5d**)[146]. This membrane mimetic has recently become a relatively common alternative to detergent micelles for native MS, with instrumental and software advances enabling analysis of MPs inside intact nanodiscs[147]. SMALPs form similar encircled bilayers, but they use SMA polymer particles to extract MPs directly from the membrane surrounded by native lipids (**Figure 1.5e**)[148]. Laser induced liquid bead ion desorption (LILBID)-MS has been used to remove nanodiscs and SMALPs in the gas phase while preserving native oligomeric states of proteins[149]. In addition to the methods described above, very recently ejection of MPs into the gas phase was achieved directly from native membranes, with no membrane mimetics required[150].

Choosing the appropriate membrane mimetic is not the only challenge for native MS of MPs. A large number of structurally diverse lipids surrounds MPs in their native environment (**Section 2**). Many of these species are extremely close in mass or completely isobaric, meaning that endogenous lipids observed bound to

MPs during a native MS experiment cannot be identified simply based on their masses. This problem is intensified by the fact that native MS has the potential to preserve even the weakest annular lipid binding events, leading to a large number of peaks which are extraordinarily challenging to resolve[151]. Therefore, most strategies involve at least partial delipidation of protein in solution. Delipidation is performed by using detergent molecules to displace lipids, so one way to achieve this is to increase the exposure time of proteins to detergent. This delipidation method was applied to an ABC-transporter protein TmrAB, with individual lipid binding events resolved in mass spectra after 48-hour delipidation in DDM at 37 ^{o}C[152]. Progressive delipidation can be also achieved by using increasing detergent concentrations or by employing naturally delipidating detergents such as OG, C8E4 or n-nonyl-β-d-glucopyranoside (NG)[153]. By increasing the concentration of OG or NG in a stepwise manner weakly-bound lipids can be removed, while selectively preserving non-annular lipids[154]. Delipidation stages can be combined with liquid chromatography-MS (LC-MS) based quantitative lipidomics on the extracted lipids – this allows identifying lipid species co-purified with the protein and then matching their masses to the binding lipids observed by native MS[153]. Alternatively, modern high-resolution mass spectrometers allow identifying lipids in the gas-phase by dissociating them from the protein with CID activation[155]. Especially impressive is the recently reported method of nativeomics, where a complete top-down characterisation of endogenously binding lipids and ligands is performed starting from a native protein complex[156]. At other times complete delipidation can be desirable, with individual lipid species added back to observe the effect they have on the membrane protein[139]. Overall, delipidation considerations combined with requirements for protein stability lead to long and

tedious detergent screening experiments having to be performed for every MP. In **Chapter 3**, I investigate a novel group of detergents (designed and synthesised by Leonhard Urner) with gradually varying delipidation strengths that are compatible with diverse MPs, so the desired degree of delipidation can be achieved simply by choosing the appropriate detergent from this group.

One of the ways in which lipids can influence MP structure is by promoting the formation of the functionally relevant oligomeric state. The ability of native MS to simultaneously observe different oligomeric states of the protein as well as individual binding events makes it a great method for investigating this type of MP-lipid interactions. For example, a bacterial leucine transporter, LeuT, was shown to be a dimer with an extra mass of about 7.4 kDa attached to it; but when LeuT was incubated with a delipidating NG detergent, exclusively monomeric species were observed[40]. Using MS/MS, the extra 7.4 kDa mass was assigned to 3 PL molecules and 1 CDL molecule per monomer, highlighting the role of lipids for maintaining the dimeric structure of LeuT. Native MS data contributed to the creation of the interfacial strength oligomerisation model, which proposes that the MPs with weak (low surface area, few or no salt bridges) subunit interfaces require lipids to form stable oligomers[40]. An example of this behaviour can be observed for the members of Na^+/H^+ antiporter family, where NapA and NHA2 have strong interfaces and form stable dimers in the absence of lipids, while NhaA forms a weak interface and requires lipids for dimerisation[40,131]. In **Chapter 4** of this thesis I use native MS to show lipid-induced dimerisation of another Na^+/H^+ exchanger, NHE9. Native MS has also been used to discover previously unknown oligomeric states of MPs facilitated by lipid binding. Beta-barrel assembly machinery (BAM) is a heterooligomer consisting of five subunits (A to E); structural studies performed in

detergent identified its stoichiometry as 1:1:1:1:1[157]. Native MS experiments performed in the native membranes revealed a hexamer containing a second BamE subunit with CDL facilitating its formation[150]. The native MS methods described above provide useful insight into the role of lipids in MP oligomerisation, but they do so indirectly through observing the changes in ratios of different oligomeric species upon addition or removal of lipids. In **Chapter 2** of this thesis, I employ SID to probe the effect of lipids on the subunit interface more directly. I show that this approach is sufficiently sensitive to discern very subtle differences to the oligomeric stability caused by lipids of the same class but different lengths of their hydrophobic tails.

Native MS can also be used to identify lipids that stabilise a particular folded state of a protein. One way to perform this type of study is to subject MP-lipid complexes to CIU. The lipids that have a strong stabilising effect on the protein unfolding are likely to be structurally important. One striking example is the interaction between ammonium transporter AmtB and PG lipid which was identified as specific by CIU and then lead to the crystal structure of PG-bound AmtB which confirmed a significant structural difference of this protein with and without the lipid[47]. In **Chapter 5** of this thesis, I use CIU to investigate the effect of lipids on a leukotriene synthase protein LTC4S. AmtB is one of the most studied MPs by native MS, with allosteric effects of lipid binding on protein-protein[158] and protein-lipid[159] interactions observed; even binding thermodynamics of lipids to AmtB were calculated using a specially designed temperature-controlled source[160]. Native MS has also been used to investigate allosteric effects of lipid binding to other MPs, including important drug targets. For example, a recent study revealed that PIP_2 enhances selectivity of G-protein coupling to GPCRs[140].

In summary, in just over a decade native MS has become an indispensable tool for probing MP-lipid interactions, with a wide range of possible applications. The next sections will provide a brief technical overview of this method, focusing on instrumentation and experimental techniques employed in this work.

1.7 Native Mass Spectrometry Instrumentation

Various spectrometer designs have been shown to be compatible with native MS. Here, I will only outline the instruments that have been extensively used in this work. These can be separated into two broad categories: TOF-based spectrometers and Orbitrap-based spectrometers. Synapt instruments (Waters Corporation) are based on the TOF mass analyser; in this work the Synapt G2[161] (**Figure 1.6a**), the prototype high-resolution instrument based upon the Synapt G2-Si system (referred to as the 'high-res' Synapt in this work)[162] (**Figure 1.6b**) and the Synapt G1 modified with a linear drift tube[163] (**Figure 1.6c**) were used. The Orbitrap system used in this work is the Q Exactive UHMR (Thermo Scentific)[164] (**Figure 1.6d**). The Synapt G2 and the 'high-res' Synapt were modified with an SID cell in the trap region as described previously[165]. In the following sections, I will describe various aspects of native MS and I will highlight the features of those instruments that make each of them suitable for particular experiments.

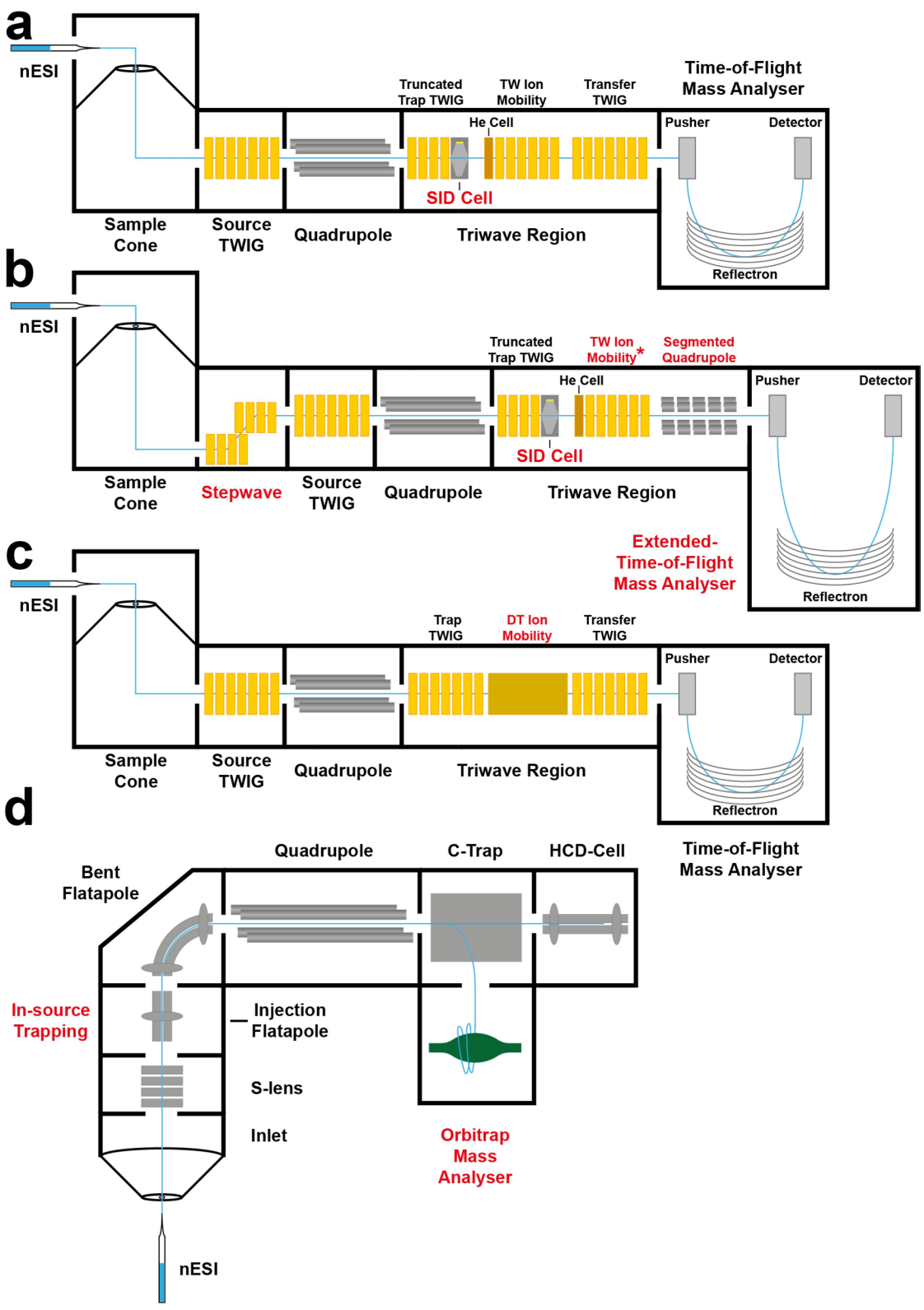


Figure 1.6 Instruments. Schematic representations of the main mass spectrometer types used in this work. Important features of each are highlighted in red. **a,** Synapt G2, modified with an SID cell. **b,** High resolution Synapt G2, with a longer time-of-flight and improved ion optics (red). (Note: a version of this instrument with a cyclic ion mobility cell was also used, see **Section 1.13** for more details). **c**, Synapt G1, modified with a linear drift tube ion mobility cell. **d,** Q Exactive UHMR, with an Orbitrap mass analyser and in-source trapping capability. Ions in the C-trap can be sent to the HCD cell for activation or straight to Orbitrap for mass analysis.

1.8 Ionisation

As was mentioned in **Section 5**, nESI is the ionisation method of choice for most native MS studies; all of the instruments used in this thesis (**Figure 1.6**) were equipped with a nESI source. nESI is a smaller version of ESI so they share the same general mechanism[166]. In order to initiate electrospray, a high (typically 1-1.5 kV for nESI) potential difference is applied between the spectrometer inlet and the conductive vessel containing the aqueous solution of the molecule of interest (**Figure 1.7a**). nESI typically employs borosilicate glass capillaries for this purpose, which are pulled into needle-like shapes, clipped to make an orifice of roughly 1-12 μm in diameter and gold-coated to make them conductive[83,167]. The imposed electric field penetrates into the liquid and causes the build-up of positive charges (in the commonly used positive ionisation mode) at the tip of the needle. The influence of the like-charge repulsion causes the extension of the meniscus, which is resisted by the surface tension of the liquid[168].The two opposing forces result in formation of the cone with a half-angle of approximately 49.3° called the Taylor cone[169]. If the applied voltage is sufficiently high, a permanent jet emerges from the tip of the cone that eventually breaks up into small charged droplets due to charge repulsion[170]. The solvent gradually evaporates from these droplets, until the surface tension becomes equal to the repulsion of positive charges in the droplet at the Rayleigh limit[171], given by the equation:

$$Q = 8\pi(\varepsilon_0 \gamma R^3)^{1/2}, \qquad (1.1)$$

where Q is the droplet charge, ε_0 is electrical permittivity, γ is the surface tension of the solvent and R is the droplet radius. Beyond this point the droplets are no longer stable, causing coulomb fission of the droplet via a similar cone-jet

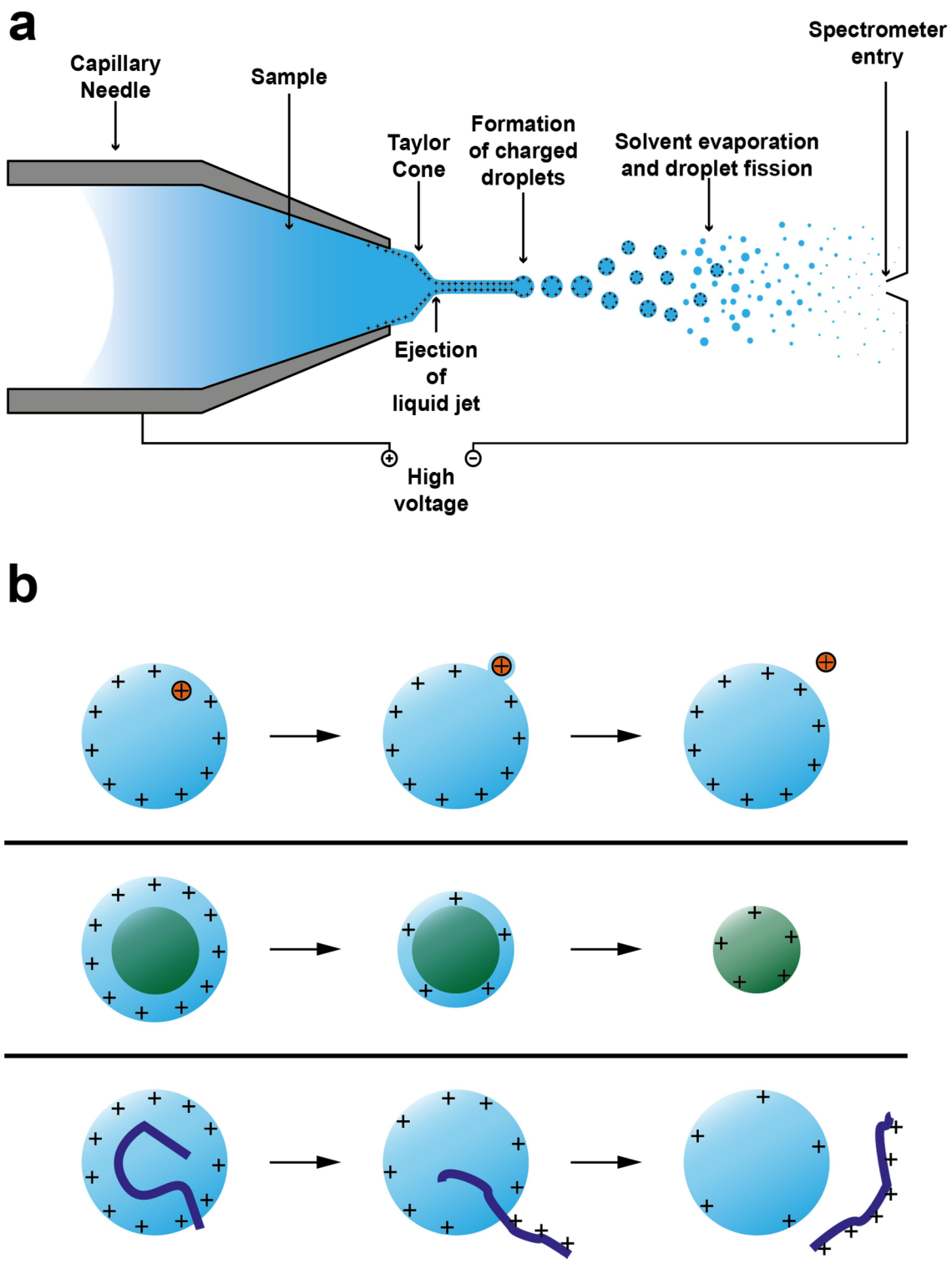


Figure 1.7 Electrospray fundamentals. a, Droplet formation. The application of high voltage causes formation of a cone which breaks down into charged droplets. Droplets decrease in size via evaporation and fission until free gaseous ions are created. **b,** Three models for gaseous ion production. Ion evaporation model (top): small charged ions break free from the droplet surface. Charged residue model (middle): globular proteins remain inside droplets until complete evaporation occurs and retain any residual charge. Chain injection model (bottom): unfolded proteins leave the droplet starting at one of the termini, picking up charge in the process.

mechanism and formation of smaller progeny droplets[172]. Repeated evaporation-fission cycles occur, until gaseous ions emerge from the final generation of droplets. Three mechanisms have been proposed for this last step, depending on the nature of the ion: ion evaporation model (IEM), charged residue model (CRM) and chain ejection model (CEM) (**Figure 1.7b**)[166]. It should be kept in mind that although these models are generally accepted, many aspects are still subject of ongoing research.

IEM proposes that once a droplet radius drops below a certain critical size (estimated to be around 10nm) ions can be ejected from the droplet surface (**Figure 1.7b, top panel**)[173]. This removal of charge replaces a fission event as the means of repulsion relief; in fact, the ion typically leaves with its own small solvation sphere and, therefore, resembles a progeny droplet, blurring the distinction between the two processes[166,174]. IEM is almost exclusively applicable to small ions, although an example of a small protein ion (ubiquitin) being formed via this mechanism has been recently reported[175]. In order for this to happen, the protein has to be supercharged in solution by lowering the pH and is required to maintain its folded state in the highly acidic environment[175]. Consequently, IEM is expected to be applicable to a very small number of proteins under specific conditions.

A much more common way to form gaseous protein ions is by CRM[176,177]. A large and roughly spherical molecule (such as a folded native-like protein) remains located in the centre of the droplet until all the solvent evaporates, leaving a naked molecule with any residual charges attached to it (**Figure 1.7b, middle panel**). Throughout the droplet evolution excess charge is removed by a combination of coulomb fission and IEM, meaning that CRM predicts the final charge on the protein to be consistent with a Rayleigh limit for a droplet with a radius of a protein

– this observation was experimentally confirmed for both soluble[178,179] and membrane proteins (solvated by non-charge reducing detergents, such as DDM)[137]. Since the process has a degree of randomness associated to it, a Gaussian distribution of charges is produced around the main ion. The mass of the species with a particular charge, z, is a sum of the mass of the protein and the mass of z charge carries, typically H^+ ions (in negative mode, the mass of z protons needs to be subtracted instead). It is important to note that the final charge on the protein is solely determined by the ESI process and has no relation to its charge in solution[166].

CEM describes the formation of ions from proteins that are denatured and unfolded in solution (**Figure 1.7b, bottom panel**)[166,180]. Unfolding causes the hydrophobic parts of the protein to be exposed to water, greatly reducing solubility. One of the termini then gets ejected from the droplet, followed by the rest of the molecule, which picks up charges from the surface of the droplet as it leaves. The large surface area of the denatured protein means that both the average charge state and the width of charge state distribution are larger than they are for a globular protein produced via CRM.

Since the protein charge is determined during the ESI stage, it is also the stage where the charge manipulations are usually performed. Charge reduction can be beneficial for maintaining the native structures of proteins in the gas phase[138]. One way the charge reduction can be achieved is by a combined CRM-IEM mechanism[181]. Small molecules act as charge carriers and remove excess charge through multiple IEM events. As a result, fewer than usual charges are left for the protein ion during the last evaporation stage of CRM. A good charge reducing agent has to be good at binding charge carriers and at the same time have a

sufficient hydrophobic character to reside close to the droplet surface. PEG polymers were shown to exhibit these properties[182]; C8E4 detergent, which is based on PEG, is also assumed to be charge reducing by this mechanism[137]. Another way to reduce the charge state is by introducing an organic solvent (typically acetonitrile) into the source of a mass spectrometer in order to expose the ESI droplets to the vapour; this mechanism most likely works by affecting the surface tension and, therefore, the Rayleigh limit of the droplets[138,183]. For MPs, charge can also be reduced by switching to negative ion mode, however, significant reduction is only observed for saccharide-based detergents, which are not charge reducing in the positive mode[184]. Further reduction is generally hard to achieve for MPs solubilised in charge reducing detergents like C8E4, with the recently described trimethylamine oxide (TMAO) being a notable exception[185]. In addition to the above methods, charge reduction can be achieved after the ESI stage of an MS experiment, by introducing custom modifications to the spectrometer; for example, cation-to-anion proton transfer reactions (CAPTR) have been recently used to investigate the effects of charge state on the protein conformation in the gas phase[186-189].

As was mentioned above, nESI is used over standard ESI for native MS. In addition to the obvious benefit of a roughly ten-fold reduction in the amount of sample required for one experiment, nESI also produces smaller initial droplets with 100-1000 times lower volumes[167]. The smaller droplets require less energy to desolvate, so the final protein ion is more likely to retain its folding state due to less internal energy[121]. The fewer desolvation events also mean that the needle can be placed closer to the mass spectrometer inlet, resulting in fewer ions missing the target[168]. Finally, the presence of multiple ions in the same ESI droplet can lead to

a non-specific interaction between them, which can be misleading. These include protein-protein interactions (which means that protein concentrations above approximately 20 μM should be avoided[121]) and protein-ligand interactions, such as salt adducts. The smaller initial droplet sizes reduce the probability of such non-specific binding events, meaning that nESI is more tolerant to the presence of salt than ESI and, together with the use of high concentrations of volatile AA which outcompetes sodium ions (**Section 1.5**), produces well-resolved protein peaks[168]. Taking this idea further, needles with sub-micrometer diameters were designed, which could produce resolved mass spectra starting with high concentrations of physiological salts[87,88].

Other ionisation methods have also been employed for native MS of membrane proteins. Desorption electrospray ionisation (DESI) allows to investigate samples from solid surfaces[190]. DESI can be used for ligand screening against MP targets[191]. LILBID uses a laser to generate ions from liquid droplets[192]. LILBID can produce sufficient energy to remove 'sticky' membrane mimetics, such as nanodiscs[193] or SMALPs[149].

1.9 Transmission of High Molecular Weight Ions

After the gaseous protein ions have been formed, it remains challenging to transfer them to the analyser and detector stages of a mass spectrometer. ESI occurs at atmospheric pressures, whereas mass analysers require very high vacuum for accurate mass measurements. Therefore, transition between the high and low pressure regions is inevitable for all the ions. The first hurdle occurs when the freshly generated protein ions go through the inlet of a mass spectrometer. The sudden pressure drop causes the ion plume to expand rapidly, and the high

velocity of ions makes refocusing them very challenging. Surprisingly, increasing the pressure in the source region greatly improves transmission of high molecular weight species[194]. This effect arises from the greater number of collisions between the ions and the neutral background gas molecules which reduces kinetic energy of the proteins, making them easier to refocus and transmit to the downstream sections of a mass spectrometer[195]. This energy dampening is called collisional focusing or collisional cooling and was traditionally practically achieved by throttling the pump with a valve[196]. For example, Synapts G1 and G2 used in this work have such a valve that allows increasing the backing pressure in the source. The modern versions of Synapts, including the 'high-res' Synapt used in this work, employ an off-axis ion guide called a StepWave (**Figure 1.6b**), which enables sufficient ion focusing without the need for a backing pressure valve[197]. Some degree of collisional cooling is provided by a larger sample cone used with StepWave systems, which increases the backing pressure by leaking more atmospheric gases inside the spectrometer[197,198]. Orbitrap-based Q Exactive systems also require no backing pressure, with a gentle voltage gradient being sufficient for ion transmission[155]; in the novel UHMR spectrometer in-source trapping (IST) can also be used to dampen excess energy of ions[164]. For all of these instruments, various radio frequency (RF) ion guides are employed to transport the ions from the source region all the way to the mass analyser; these include stacked ring ion guides (SRIGs) and multipole ion guides[164,199,200].

1.10 Quadrupole Mass Filter and Tandem MS

A quadrupole is usually located immediately after the source ion guides and is an essential part of MS/MS experiments (**Figure 1.4a**). The quadrupole consists of

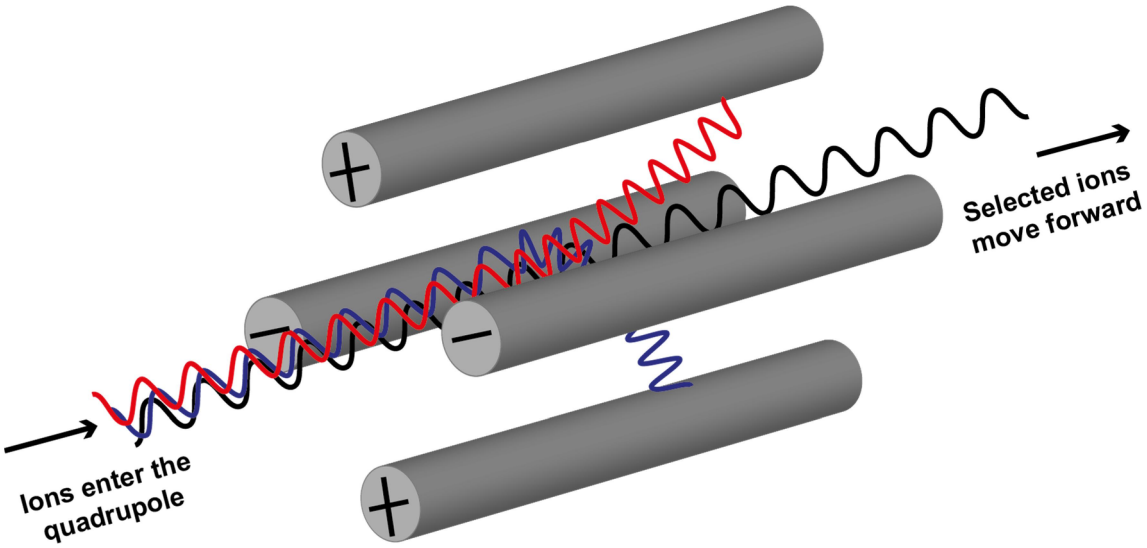


Figure 1.8 Quadrupole mass filter. Quadrupole operated in selection mode; only ions with a certain mass-to-charge ratio (black) have a stable trajectory.

four rod electrodes, often circular or hyperbolic in cross-section (**Figure 1.8**). A suitable electric field for ion transmission and selection can be created by applying the same potential to opposing rods and an opposing potential to adjacent rods. Both continuous voltage, with value *U*, and oscillating RF voltage, with amplitude *V* and frequency *f*, are applied. For each combination of *U* and *V*, only ions with particular *m/z* values have a stable oscillating trajectory and are able to traverse the quadrupole without colliding into one of the electrodes. This *m/z* can be calculated by the second order differential equation called the Mathieu equation, solutions to which are well known. Thus, the quadrupole can act as a mass filter during an MS/MS experiment, selecting only a single species for further interrogation. Alternatively, it is possible to use the quadrupole as an ion guide by applying only RF voltage[196].

The maximum *m/z* that can be resolved by a particular quadrupole, M_{max}, is given by the equation:

$$M_{\max} = \frac{7 \times 10^6 V_{\max}}{f^2 r_0^2}, \quad (1.2)$$

where V_{max} is the RF amplitude, f is the RF frequency and r_0 is the inner radius between the rods in meters[201]. In order to achieve high values of *m/z* required for native MS, either f can be reduced, r_0 can be reduced or V_{max} can be increased. The latter two are practically challenging due to a possible voltage breakdown, so low RF frequency quadrupoles are usually employed for native MS, which function up to 32000 *m/z*[196]. Low RF frequencies significantly compromise the resolving power of quadrupoles, so they are no longer used as mass analysers[202], despite being popular during the early native MS history[203].

It should be noted that for an MS/MS experiment a completely desolvated ion has to be generated in the source region before reaching the quadrupole since an exact *m/z* value needs to be determined. This can pose a challenge for many MPs protected by membrane mimetics, which often require high amounts of energy to dissociate. On UHMR instruments the IST capability presents an opportunity for high energy deposition in the source region without significant transmission losses[164]. For Synapts, the source activation is generally limited to 200V in the sample cone, although a high energy source modification has been demonstrated[153].

1.11 TOF Mass Analyser

Unlike the quadrupole, the TOF mass analyser has only small losses of resolution at high *m/z* values, with the mass range limited only by detector sensitivity[204]. An initial packet of ions is given a burst of kinetic energy by an acceleration pulse with potential, W. The ions then traverse a known distance, L, through a field-free region

with a constant velocity. The time it takes to travel this distance, *t*, can be measured and then used to calculate *m/z* by the equation:

$$m/z = 2eW\left(\frac{t}{L}\right)^2, \quad (1.3)$$

where *e* is the electron charge.

A continuous ion source, such as ESI, is not naturally compatible with the TOF, since a packet of ions has to be released at a well-defined time. This problem is alleviated by using orthogonal acceleration TOF, where the analyser is positioned perpendicularly to the ion flow (**Figure 1.9**). This allows collecting a small packet which is then released by the pusher electrode. To correct for imperfect focusing of the ions resulting in a distribution of initial orthogonal velocities, a reflectron device is introduced[205]. An electric field opposes the movement of the ions; more energetic ions penetrate deeper into the field before being turned around, so the ions with the same mass will arrive at the detector simultaneously[206].

TOF requires very high vacuum (around 10^{-7}-10^{-6} mbar) to minimise the number of collisions occurring as they reduce both sensitivity and resolution[207]. The 'high-res' Synapt modifications include a novel pre-TOF transfer optic in the form of an axial field segmented quadrupole and a longer pathway for the TOF itself (**Figure 1.6b**), which in combination significantly improve the mass resolution and sensitivity compared to the older Synapt models[162].

1.12 Orbitrap Mass Analyser

Orbitrap is a relatively recent addition to the native MS toolbox[200], but it has continuously grown in popularity during the last decade. In an Orbitrap ions orbit around the central electrode; while orbiting, they oscillate along the electrode, with

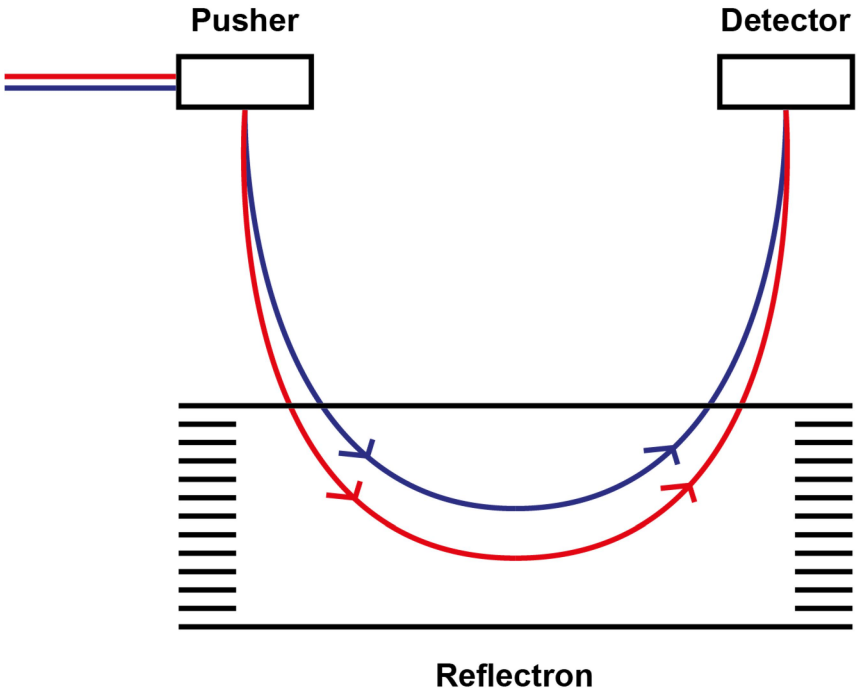


Figure 1.9 Time-of-flight mass analyser. Blue and red ions have identical masses, but differ in initial velocity (red ions are faster, blue – slower). The reflectron corrects the difference, resulting in simultaneous arrival of both ions to the detector.

axial frequencies dependent on their *m/z* values. Moving ion packets induce an image charge on pairs of outer electrodes and this time domain waveform is converted into a frequency (*m/z*) domain by Fourier transform[208]. Orbitrap has found increasing use in native MS due to its ability to produce very high mass resolution spectra, even in the high *m/z* range[209]. One example of the utility of the Orbitrap is the power to resolve endogenously bound lipids bound to MPs, with only a few Da difference between the species[155]. The main disadvantage of the Orbitrap is the fact that it cannot be readily coupled to IM, due to the similar analysis time of both of these methods – some prototype IM-Orbitrap systems have been developed, but they lead to much more time consuming experiments than the existing IM-TOF architectures[210].

1.13 Ion Mobility Spectrometry

IM is a very powerful method that provides the information about the size of an ion rather than its mass. Inside an IM cell, ions are separated based on their mobility as they move through inert gas under the influence of a weak electric field[211]. As was mentioned in **Section 1.5**, mobility can then be converted into a CCS value, which is indicative of the protein size. Two main types of IM cells are combined with native MS: drift tube (DT) and travelling wave (TW).

Inside the DT the ions undergo numerous collisions with neutral gas molecules (usually helium or nitrogen) driven by a uniform, static electric field, *E* (**Figure 1.10a**). The ions move at constant velocity, which is directly related to *E* by a constant, *K*, also called mobility. Mobility can be calculated from the time taken for ions to traverse the tube, t_D, by the equation:

$$K = \frac{D}{t_D E} + t_0, \tag{1.4}$$

where *D* is the distance the ion has to travel through the drift tube and t_0 is the time spent between the end of the drift tube and detection. The CCS value, Ω can then be obtained using the Mason-Schamp[212] equation:

$$\Omega = \frac{3ze}{16N}\left(\frac{2\pi}{\mu k_B T}\right)^{0.5}\frac{1}{K_0}, \tag{1.5}$$

where *z* is the ion charge state, *e* is the electron charge, *N* is drift gas number density, k_B is the Boltzmann constant, μ is the reduced mass of the ion-drift gas pair, T is the gas temperature and K_0 is the reduced mobility (mobility corrected to 273.2 K and 760 Torr). For the two species with the same mass and charge, the one with the larger CCS will undergo more collisions and thus will have a longer drift time.

TW is an alternative type of an IM cell that has been incorporated into commercially available Synapt systems due to its high transmission efficiency, sensitivity and resolving power[199,213,214]. TWIM consists of a stacked-ring ion guide[213,215], with opposite phases of RF voltage applied to consecutive ring electrodes to provide radial confinement[199]. A direct current potential is then sequentially applied to each pair[199] (or two pairs[215]) of electrodes downstream, providing a travelling wave that pushes the ions forward. The IM separation is achieved by using conditions at which the ions are propelled forward by the waves instead of "surfing" them, continuously rolling over in the process[199]. Ions with higher mobility values are overtaken by the waves fewer times, resulting in a smaller drift time value (**Figure 1.10b**). Nitrogen gas is preferred over helium for TWIM, because higher mobility of ions in helium would require using lower travelling wave heights in order to avoid the surfing scenario, at the cost of resolution[215]. However, filling the cell with nitrogen at high pressures is problematic, as very high potentials would be required to allow ions to enter the cell, leading to ion losses through scattering and fragmentation. As a consequence, starting from Synapt G2 a helium cell precedes the nitrogen-filled IM cell (**Figure 1.6a**), as significantly lower fields are required to introduce ions into a helium-filled chamber due to higher ion mobilities in helium[215].

One major disadvantage of TWIM compared to DTIM is the inability to directly convert the observed arrival time distributions (ATDs) into CCS values. Instead, a calibration curve is required, using calibrants with known CCS values (previously measured on DT instruments)[163,216]. Soluble protein calibrants do not work for MPs of similar mass due to a significant difference in mobility values[217]. A strategy using soluble calibrants of strictly higher mass than the MP of interest has been

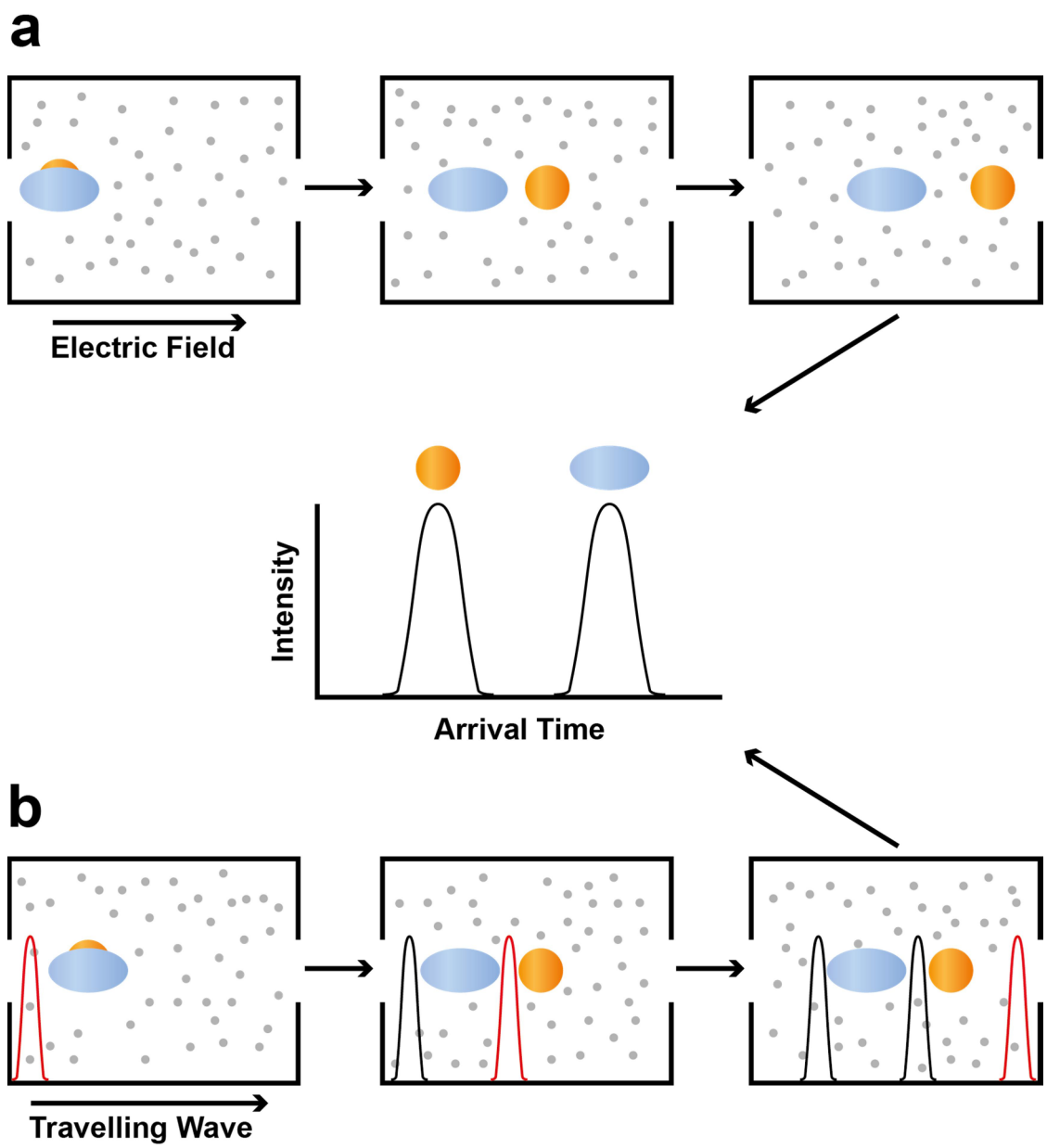


Figure 1.10 Ion mobility spectrometry. Blue and orange represent two species with identical mass and charge, but different sizes. **a,** Drift tube ion mobility spectrometry. Proteins move under the effect of a weak electric field. The larger protein undergoes more collisions with neutral gas molecules (grey), resulting in a greater arrival time. **b,** Travelling wave ion mobility spectrometry. Ions are propelled forward by pulses of potential, the larger protein 'rolls over' the wave more times, also resulting in a greater arriving time.

proposed, however, significant variations in the accuracy were observed depending on the TW conditions[217]. With these considerations in mind, I have chosen not to attempt calculating the CCS values of MPs from TW data and report

them as ATDs instead. Where the CCS values were desired, a DT instrument was used to calculate those directly using the equation (1.5).

The main utility of experimentally obtained CCS values comes from being able to compare them against the theoretical ones, calculated from crystal structures – this allows verifying that the protein maintains its native-like conformation in the gas phase. Numerous computational methods for estimating a theoretical CCS value exist, differing in accuracy and computing power required[218]. One relatively simple method is projection approximation (PA), equating the CCS to the average projection area over all possible orientations[219]. This method ignores the long-range interactions and the scattering of protein by the gas molecules, so the calculated value is underestimated[103]. However, a simple multiplication by an empirically derived factor of 1.14 corrects the PA CCS value[220]. In addition, any differences in the exact sequence composition of the protein between the crystal structure and the experiment can also be corrected by equation[102]:

$$CCS_{calc} = 1.14 CCS_{PA} \left(\frac{M_{exp}}{M_{PDB}}\right)^{\frac{2}{3}}, \quad (1.6)$$

Where M_{exp} is the molecular weight of the protein used in the experiment and M_{PDB} is the molecular weight of the protein in the crystal structure.

The resolving power of either a DT or a TW cell increases proportionally to the square root of its length[221]. While several extremely long IM cells have been constructed[222,223], an elegant practical approach is to build a cyclic TW IM cell, where the ions can be sent for multiple passes essentially creating an arbitrarily long path[162]. The IM resolution of the new cyclic device developed by Waters Corporation is higher than the previous generation of the TWIM used in Synapts

even for a single pass around the 'track'[162]. The Cyclic IM cell was shown to be compatible with big macromolecules, such as proteins, with no significant structural rearrangements observed from prolonged exposure to vacuum[224].

The way IM experiments are practically performed on Synapt instruments is by first collecting a packet of ions and storing it in the Trap ion guide before the IM cell (**Figure 1.6a-c**). The ions are then released at the same time and are separated according to their mobility in the DT or the TW IM chamber. This separation is maintained in the Transfer ion optic all the way to the pusher electrode, where the first 'bin' of ions is collected and released into TOF (as was described in **Section 1.11**). A total of 200 such bins constitute a single IM cycle; incorrectly chosen settings can result in a 'roll-over' effect, if the ions take longer than 200 bins and appear to be in the first few bins of the next IM cycle instead. This method relies on the TOF operating at a much faster time scale than IM; so it is not readily extended to Orbitrap as was mentioned in the previous section.

1.14 Collision Induced Unfolding and Dissociation

A simple way to impart energy into protein ions inside a mass spectrometer is by repeated collisions with background gas molecules. The energy of such collisions is controlled by electrode potential, which increases the velocity of ions. This activation can happen at any stage of a mass spectrometer due to the presence of some amount of neutral gas all the way until the high vacuum of a mass analyser stage. Native MS instruments usually also have dedicated collision cells filled with neutral gas (typically nitrogen or argon) at higher pressures and with ability to apply high voltages – for example, the Trap and Transfer ion guides on Synapts can act

as collision cells (**Figure 1.6a-c**), and so does the HCD-cell (higher-energy C-trap dissociation, a variant of CID) on UHMR (**Figure 1.6d**).

The degree of activation controls the effect on the protein ion. At minimal collision energies collisional cooling helps with focusing and transmission of high molecular weight species (**Section 1.9**), while higher energies are required to dissociate unwanted solvent adducts[89] (or membrane mimetics in the case of MPs)[139]. Increasing the potential even higher can lead to structural rearrangements, such as unfolding and even dissociation – this is the basis of CIU and CID. Their mechanism is discussed below.

Repeated collisions gradually increase the internal energy of a protein, disrupting some of the interactions that stabilise the folded state and eventually resulting in the unfolding of one or more subunits[225]. The unfolding event is accompanied by the transfer of protons from the other subunits which move to the newly available surface area to minimise charge repulsion; this causes further destabilisation of the unfolding subunit creating a 'positive-feedback' loop[226-230]. The unfolding can be preceded by a compaction for protein complexes containing a large cavity, especially if they have a low overall charge[102]. In a CIU experiment, this unfolding is happening in a controlled manner, with gradually increasing collision voltage starting from the native-like conformation and observing the changes in CCS by IM. The increase in CCS with voltage is not gradual, but instead happens in one or more sharp characteristic transitions, which at least partially mimic the unfolding of individual domains in solution[231,232]. The change in the transition voltage (measured at the point where the intensity of the original conformation drops to 50%) upon addition of ligands or lipids corresponds to induced structural rearrangements (**Figure 1.11**) (some examples of specific MP-lipid interactions

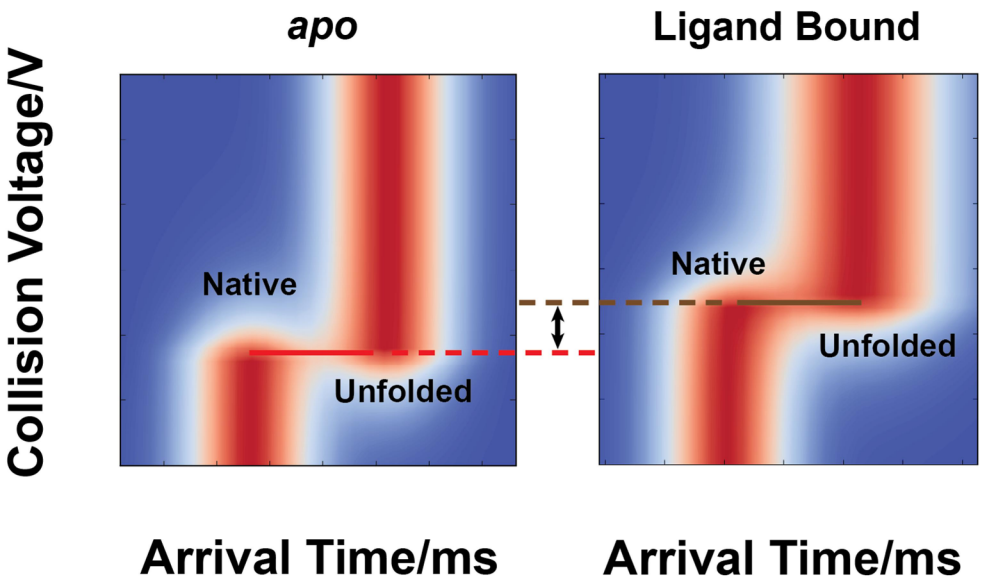


Figure 1.11 Collision-induced unfolding. Model data of a protein unfolding, showing one characteristic unfolding transition. Upon ligand binding, this transition occurs at a higher collision voltage, reflecting induced stabilisation.

discovered this way were discussed in **Section 1.5**). This voltage increase can be converted into the stabilisation value[233]. Stabilisation values should be compared against a standard for every protein, because even non-specific binding events can cause a small increase in CIU stabilisation values[234]. While theoretical models exist relating the stabilisation value of a protein to actual physical energy values[47], this is still the subject of ongoing research; therefore, in his work such values are only used to observe the differences in stabilisation induced by different binding partners to the same protein, with no deeper physical meaning attached.

Increasing the energy even further results in complete dissociation of one of the subunits – CID occurs[89]. If possible, this happens at the point where both fragments possess equal amounts of charge (for large complexes) as this is where the energy barrier for dissociation is minimised due to the maximum charge repulsion[226]. This results in an 'asymmetric' charge pattern, where a monomer leaves with half of the total charge, despite representing a much smaller fraction of the mass of the

complex[227]. After the first subunit has departed, the process can be repeated at higher energies, with more monomer subunits dissociating sequentially from the stripped oligomer[227].

The release of a single unfolded subunit is the CID pathway for the vast majority of protein complexes regardless of internal architecture, meaning that very little structural information can be extracted[89]. A small number of exceptions are known, however, where CID results in the release of compact, low-charged fragments. These are thought to occur when dissociation at the subunit interface is more energetically favourable than unfolding, for example, if the monomer is stabilised towards unfolding by intramolecular disulphide bonds[235] or if the interface is particularly weak[236] or destabilised by repulsion due to a high amount of charge on the complex[102]. In addition, unfolding does not occur if there is insufficient charge in the system[229], as is often the case for small charge-reduced dimers. SID produces folded dissociation products reflecting internal arrangement of subunits in a much more predictable and reliable way and will be reviewed in the next section.

An alternative mechanism has recently been proposed to explain the experimentally observed charged asymmetry of CID products. Salt bridge rearrangement (SaBRe) model explains the charge asymmetry by heterolytic scission of ion pairs in salt bridges, which naturally results in the majority of charges remaining on the same subunit after dissociation[237]. However, recent data obtained by gas-phase HDX experiments revealed migration of a large number of deuterium labels into the dissociated subunit; this would not be expected in SaBRe mechanism as the interfaces containing the salt bridges are shielded and, therefore, should be deuterium-free[238]. While this observation does not completely

rule out some contribution of SaBRe to protein complex dissociation, I will be using a more traditional subunit unfolding interpretation of CIU, CID and SID data in this thesis based on current evidence.

1.15 Surface Induced Dissociation

In SID, ions acquire energy via a single high-energy collision with a solid surface, in contrast to the multitude of low-energy collisions with small neutral gas molecules during CID. SID as an MS method was first introduced in 1975, when the ion beam was shrouded with a stainless steel tube that was used to dissociate small organic molecules[239]. It was followed up a decade later, with the surface introduced into a modified magnetic sector-quadrupole spectrometer for a more routine application of SID[240]. Since then, SID devices have been introduced into many different instrumental designs, including spectrometers used for native MS[241].

One import factor for SID devices is the surface material. A large variety of materials have been used for SID of gaseous ions, including diamond, graphite, contaminated and clean metals, Langmuir-Blodgett films and alkanethiolate self-assembled monolayer (SAM) films[242]. The latter, especially the fluorinated forms, have a number of properties that make them particularly useful as SID collision targets. SAM films spontaneously form by adsorption of the sulphur head group of alkanethiols to a piece of gold immersed in solution[243], resulting in well-organised monolayer structures[244]. Fluorinated SAMs are relatively easy to prepare[244], give decreased neutralisation of the ions due to high kinetic energies[245] (neutralisation is a major source of signal loss for metallic surfaces, as the electrons are easily ejected from the surface by the energy transferred during collision) and provide

more efficient kinetic-to-internal energy conversion compared to non-fluorinated SAMs[246]. In addition, they are stable under vacuum, resist contamination due to their low reactivity and can be very easily regenerated[246].

For a long time, the application of SID was mostly limited to small molecules[247]. In the latter half of the last decade SID was extended to native-like protein complexes. The activation of a non-specific dimer of cytochrome C by the surface resulted in a symmetric charge partitioning for the two monomers, in contrast to CID[248]. Similar results were observed for higher order oligomers, with the monomer leaving with the fraction of charge proportional to its fraction of mass; interestingly, larger dissociation products, such as dimers, were also produced[249]. The difference in the charge patterns was hypothesised to be dependent on the different time scales of the two dissociation processes: while the 'slow-heating' of multiple gas collisions results in subunit unfolding, the fast (picoseconds) nature of surface activation can provide access to alternative reaction pathways where multiple interfacial bonds are broken simultaneously, resulting in dissociation without unfolding[242,248]. However, these initial experiments were performed on a Q-TOF (quadrupole-TOF) instrument without an IM cell[250], so it was not possible to observe the unfolding events directly. In order to address this limitation, an SID device was installed into Synapt G2 to enable CCS measurements: this showed both that the low charge dissociation products were native-like (SID-IM configuration[251]) and that more compact precursors give more compact SID products (IM-SID configuration[165]). The design of this SID cell (**Figure 1.12a**) is the same as the ones used in this work and is described below.

The cell consists of 10 stainless steel electrode lenses: two entrance lenses, two exit lenses, five deflector lenses and the surface lens; the voltage can be applied

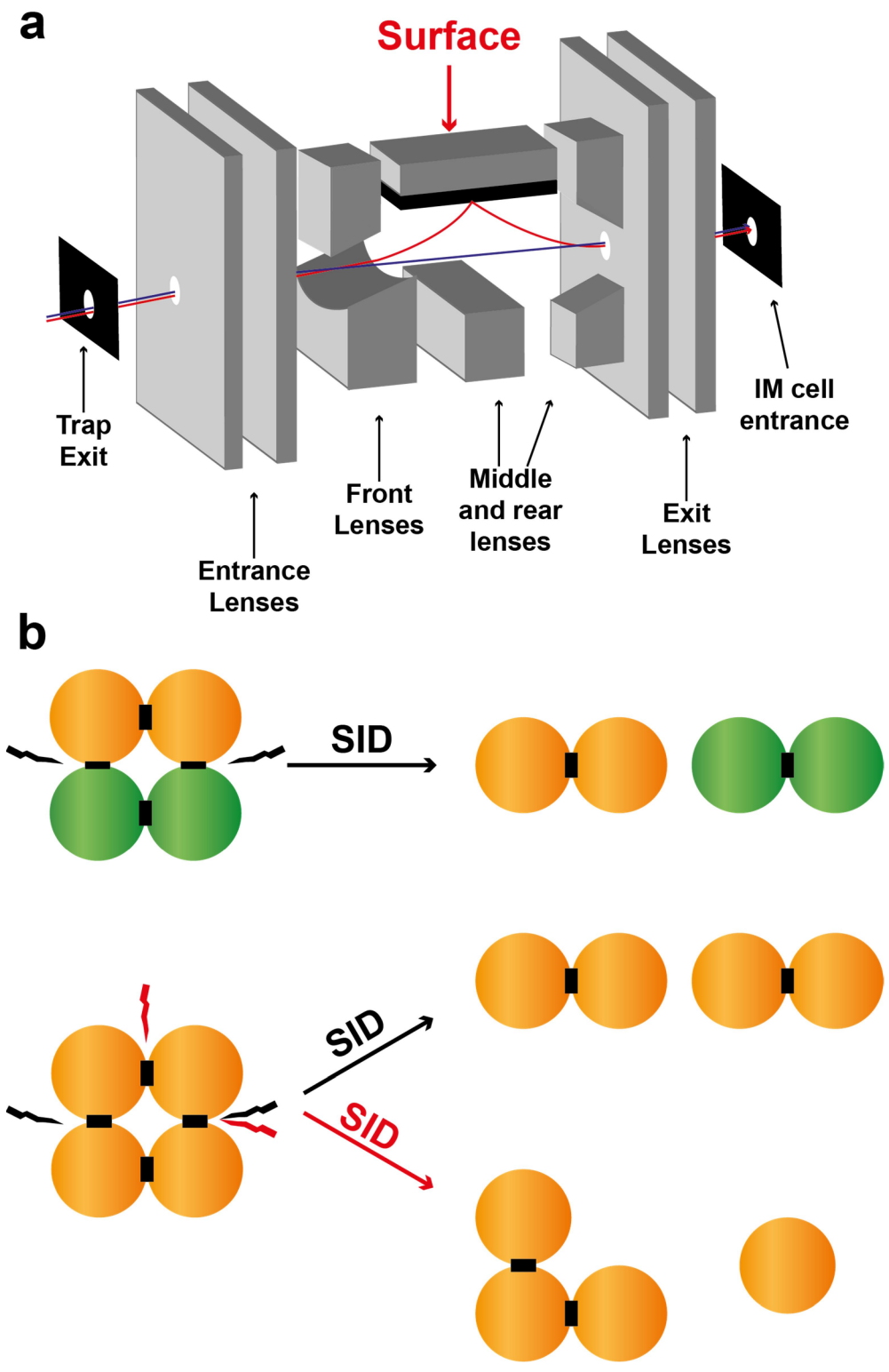


Figure 1.12 Surface induced dissociation. a, SID cell, operated in fly-through mode (blue) and SID mode (red). Each lens is an individually controlled electrode. **b,** SID of homotetramers with D2 (top) and C4 (bottom) symmetry. For a D2 tetramer the weaker interfaces break first, resulting in formation of dimers; for a C4 tetramer all of the interfaces are identical, so monomers, dimers and trimers can form.

independently to each lens. The surface is coated with gold, which is in turn coated with 2-(perfluorodecyl)-ethanethiol to form a fluorinated SAM[250]. The device can be operated in two modes: in fly through mode (ions transmitted through the cell without steering them into the surface) all the lenses have approximately the same voltages as the instrumental inputs directly before and after the SID cell in order to transmit ions through; in SID mode the front end deflectors are used to steer the ions into the surface and the back end deflectors are tuned to draw the products of the surface and transmit them into the helium cell. The SID acceleration voltage is determined by the potential difference between the trap exit and the SID surface and can be manipulated using the trap bias. Another important parameter is the injection voltage, which is defined by the difference in potentials between the SID surface and the helium cell entrance. This determines the amount of energy used to pull the ions from the cell into the next section of the spectrometer after collision.

One of the main advantages of SID is its ability to provide information on the internal subunit architecture and assembly pathways of oligomeric protein complexes. One example of this ability comes from investigating tetrameric proteins with D2 (dimer of dimers) symmetry. It had been previously shown that the disassembly pathway in solution is the reverse of the assembly pathway and that the larger (stronger) interfaces are formed earlier during assembly[252]. The expected assembly pathway for D2 tetramers is monomer-dimer (C2)-tetramer (D2)[253]. SID dissociation showed clear preference for dimers, with CCS values consistent with the cleavage of the smallest interface[254]. Importantly, for tetramers with C4 symmetry (having four identical interfaces) dissociation into monomers, dimers and trimers was observed as expected (**Figure 1.12b**). This finding demonstrates the ability of SID to provide direct information on the strength of

subunit interfaces[255]. This fact makes SID an extremely useful technique for improving protein modelling[256], generating novel structural models when high-resolution structures cannot be obtained[257,258], and confirming subunit connectivity and topology of designed protein complexes[259]. To access the most informative SID pathways (maintaining the connectivity of fragments and minimal unfolding) the use of low charge states is recommended[260].

SID can also be used to investigate the effects of ligand binding on the interfacial cleavage energy. This is usually achieved by the means of energy-resolved MS (ERMS) plot, where the amount of the surviving precursor is plotted against the applied SID energy. Ligands that are expected to stabilise the interface resulted in higher energies required for dissociation compared to the *apo* precursor, whereas non-interfacial ligands had no significant effect[254]. The nature of the dissociation products (for example, whether monomers or dimers are formed and how many ligands remain bound) can also provide information about the ligand localisation[261,262]. SID was also applied to MP-lipid interactions, however no stabilisation effects have been observed so far[263]. In **Chapter 2**, I detect stabilising interfacial lipids for the first time and develop a method to observe small differences in stabilisation induced by different lipid species.

Despite the wealth of structural information that can be obtained using SID, its application has been limited to a small number of research groups. This can be partly attributed to the complicated SID device, with 10 manually tuneable voltages (**Figure 1.12a**). The use of such devices requires a highly experienced operator and can result in significant signal losses even after hours of optimisation. In order to fix this problem, novel SID cells have recently been designed and implemented. A generation 2 device cuts down the number of individually-tuned voltages from 10

to 7 [264]. Intriguingly, a generation 3 device with only three electrodes was demonstrated earlier this year – it uses a novel split lens geometry that naturally replaces existing instrumental lenses and significantly improves sensitivity[241]. Both generation 2 and generation 3 devices work with stainless steel surfaces instead of the fluorinated SAMs described above; surprisingly, this does not result in major charge neutralisation by electrons from the low ionisation energy metal[241,264]. Together with combining SID with high-resolution mass analysers, such as Orbitrap[265] and Fourier-transform ion cyclotron resonance (FTICR)[266], these developments are paving the way for a wide-spread application of native SID-MS in the near future.

References

1 Lodish, H. *et al.* in *Molecular Cell Biology. 4th Edition. Section 3.4, Membrane Proteins* (W.H.Freeman, 2000).

2 Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Biol* **19**, 281-296, doi:10.1038/nrm.2017.138 (2018).

3 Muller, D. J., Wu, N. & Palczewski, K. Vertebrate membrane proteins: structure, function, and insights from biophysical approaches. *Pharmacol Rev* **60**, 43-78, doi:10.1124/pr.107.07111 (2008).

4 Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M. & Berglund, L. Prediction of the human membrane proteome. *Proteomics* **10**, 1141-1149, doi:10.1002/pmic.200900258 (2010).

5 Bakheet, T. M. & Doig, A. J. Properties and identification of human protein drug targets. *Bioinformatics* **25**, 451-457, doi:10.1093/bioinformatics/btp002 (2009).

6 Blobel, G. Intracellular protein topogenesis. *Proc Natl Acad Sci U S A* **77**, 1496-1500, doi:10.1073/pnas.77.3.1496 (1980).

7 Fairman, J. W., Noinaj, N. & Buchanan, S. K. The structural biology of beta-barrel membrane proteins: a summary of recent reports. *Curr Opin Struct Biol* **21**, 523-531, doi:10.1016/j.sbi.2011.05.005 (2011).

8 Bolla, J. R., Agasid, M. T., Mehmood, S. & Robinson, C. V. Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. *Annu Rev Biochem* **88**, 85-111, doi:10.1146/annurev-biochem-013118-111508 (2019).

9 Remm, M. & Sonnhammer, E. Classification of transmembrane protein families in the Caenorhabditis elegans genome and identification of human orthologs. *Genome Res* **10**, 1679-1689, doi:10.1101/gr.gr-1491r (2000).

10 Oberai, A., Ihm, Y., Kim, S. & Bowie, J. U. A limited universe of membrane protein families and folds. *Protein Sci* **15**, 1723-1734, doi:10.1110/ps.062109706 (2006).

11 Lodish, H. *et al.* in *Molecular Cell Biology. 4th Edition. Section 15.2, Overview of Membrane Transport Proteins* (W.H.Freeman, 2000).

12 Wolfersberger, M. G. Uniporters, symporters and antiporters. *J Exp Biol* **196**, 5-6 (1994).

13 Alberts, B. *et al.* in *Molecular Biology of the Cell. 4th Edition. Ion Channels and the Electrical Properties of Membranes* (Garland Science, 2002).
14 Bockaert, J. & Pin, J. P. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *Embo J* **18**, 1723-1729, doi:10.1093/emboj/18.7.1723 (1999).
15 Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356-363, doi:10.1038/nature08144 (2009).
16 Basith, S. *et al.* Exploring G Protein-Coupled Receptors (GPCRs) Ligand Space via Cheminformatics Approaches: Impact on Rational Drug Design. *Front Pharmacol* **9**, 128, doi:10.3389/fphar.2018.00128 (2018).
17 Shoichet, B. K. & Kobilka, B. K. Structure-based drug screening for G-protein-coupled receptors. *Trends Pharmacol Sci* **33**, 268-272, doi:10.1016/j.tips.2012.03.007 (2012).
18 Paschkowsky, S., Oestereich, F. & Munter, L. M. Embedded in the Membrane: How Lipids Confer Activity and Specificity to Intramembrane Proteases. *J Membrane Biol* **251**, 369-378, doi:10.1007/s00232-017-0008-5 (2018).
19 Fahy, E. *et al.* A comprehensive classification system for lipids. *J Lipid Res* **46**, 839-861, doi:10.1194/jlr.E400004-JLR200 (2005).
20 Berg, J. M., Tymoczko, J. L. & Stryer, L. in *Biochemistry. 5th Edition. Section 12.3 There Are Three Common Types of Membrane Lipids* (W H Freeman, 2002).
21 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, doi:10.1038/nrm2330 (2008).
22 Yamashita, A. *et al.* Acyltransferases and transacylases that determine the fatty acid composition of glycerolipids and the metabolism of bioactive lipid mediators in mammalian cells and model organisms. *Prog Lipid Res* **53**, 18-81, doi:10.1016/j.plipres.2013.10.001 (2014).
23 Malhotra, R. Membrane Glycolipids: Functional Heterogeneity: A Review. *Biochem Anal Biochem* **1**, doi:10.4172/2161-1009.1000108 (2012).
24 Emiola, A., George, J. & Andrews, S. S. A Complete Pathway Model for Lipid A Biosynthesis in Escherichia coli. *PLoS One* **10**, e0121216, doi:10.1371/journal.pone.0121216 (2014).
25 Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635-700, doi:10.1146/annurev.biochem.71.110601.135414 (2002).
26 Whitfield, C. & Trent, M. S. Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* **83**, 99-128, doi:10.1146/annurev-biochem-060713-035600 (2014).
27 Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol Rev* **40**, 133-159, doi:10.1093/femsre/fuv008 (2016).
28 Ernst, R., Ballweg, S. & Levental, I. Cellular mechanisms of physicochemical membrane homeostasis. *Curr Opin Cell Biol* **53**, 44-51, doi:10.1016/j.ceb.2018.04.013 (2018).
29 Atilla-Gokcumen, G. E. *et al.* Dividing cells regulate their lipid composition and localization. *Cell* **156**, 428-439, doi:10.1016/j.cell.2013.12.015 (2014).
30 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, doi:10.1016/j.plipres.2015.04.003 (2015).
31 Marsh, D. Protein modulation of lipids, and vice-versa, in membranes. *Biochim Biophys Acta* **1778**, 1545-1575, doi:10.1016/j.bbamem.2008.01.015 (2008).
32 Lee, A. G. How lipids affect the activities of integral membrane proteins. *Biochim Biophys Acta* **1666**, 62-87, doi:10.1016/j.bbamem.2004.05.012 (2004).
33 Lee, A. G. Lipid-protein interactions. *Biochem Soc Trans* **39**, 761-766, doi:10.1042/BST0390761 (2011).
34 Corradi, V. *et al.* Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS Cent Sci* **4**, 709-717, doi:10.1021/acscentsci.8b00143 (2018).

35 Contreras, F. X., Ernst, A. M., Wieland, F. & Brugger, B. Specificity of intramembrane protein-lipid interactions. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a004705 (2011).
36 Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochim Biophys Acta* **1838**, 1548-1559, doi:10.1016/j.bbamem.2013.11.009 (2014).
37 Starling, A. P., East, J. M. & Lee, A. G. Effects of phospholipid fatty acyl chain length on phosphorylation and dephosphorylation of the Ca(2+)-ATPase. *Biochem J* **310 (Pt 3)**, 875-879, doi:10.1042/bj3100875 (1995).
38 Powl, A. M., East, J. M. & Lee, A. G. Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL. *Biochemistry-Us* **47**, 12175-12184, doi:10.1021/bi801352a (2008).
39 Powl, A. M., East, J. M. & Lee, A. G. Anionic phospholipids affect the rate and extent of flux through the mechanosensitive channel of large conductance MscL. *Biochemistry-Us* **47**, 4317-4328, doi:10.1021/bi702409t (2008).
40 Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **541**, 421-424, doi:10.1038/nature20820 (2017).
41 Musatov, A. & Sedlak, E. Role of cardiolipin in stability of integral membrane proteins. *Biochimie* **142**, 102-111, doi:10.1016/j.biochi.2017.08.013 (2017).
42 van Gestel, R. A. *et al.* The influence of the acyl chain composition of cardiolipin on the stability of mitochondrial complexes; an unexpected effect of cardiolipin in alpha-ketoglutarate dehydrogenase and prohibitin complexes. *J Proteomics* **73**, 806-814, doi:10.1016/j.jprot.2009.11.009 (2010).
43 Marius, P. *et al.* Binding of anionic lipids to at least three nonannular sites on the potassium channel KcsA is required for channel opening. *Biophys J* **94**, 1689-1698, doi:10.1529/biophysj.107.117507 (2008).
44 Pucadyil, T. J. & Chattopadhyay, A. Cholesterol modulates ligand binding and G-protein coupling to serotonin(1A) receptors from bovine hippocampus. *Biochim Biophys Acta* **1663**, 188-200, doi:10.1016/j.bbamem.2004.03.010 (2004).
45 Contreras, F. X. *et al.* Molecular recognition of a single sphingolipid species by a protein's transmembrane domain. *Nature* **481**, 525-529, doi:10.1038/nature10742 (2012).
46 King, G. & Sharom, F. J. Proteins that bind and move lipids: MsbA and NPC1. *Crit Rev Biochem Mol Biol* **47**, 75-95, doi:10.3109/10409238.2011.636505 (2012).
47 Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172-+, doi:10.1038/nature13419 (2014).
48 Birch, J. *et al.* The fine art of integral membrane protein crystallisation. *Methods* **147**, 150-162, doi:10.1016/j.ymeth.2018.05.014 (2018).
49 Harkey, T. *et al.* The Role of a Crystallographically Unresolved Cytoplasmic Loop in Stabilizing the Bacterial Membrane Insertase YidC2. *Sci Rep* **9**, 14451, doi:10.1038/s41598-019-51052-9 (2019).
50 Walz, T. & Grigorieff, N. Electron Crystallography of Two-Dimensional Crystals of Membrane Proteins. *J Struct Biol* **121**, 142-161, doi:10.1006/jsbi.1998.3945 (1998).
51 Kuhlbrandt, W. Biochemistry. The resolution revolution. *Science* **343**, 1443-1444, doi:10.1126/science.1251652 (2014).
52 Cheng, Y. Membrane protein structural biology in the era of single particle cryo-EM. *Curr Opin Struct Biol* **52**, 58-63, doi:10.1016/j.sbi.2018.08.008 (2018).
53 Zhang, Y. *et al.* Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature* **546**, 248-253, doi:10.1038/nature22394 (2017).
54 Garcia-Nafria, J., Lee, Y., Bai, X., Carpenter, B. & Tate, C. G. Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein. *Elife* **7**, doi:10.7554/eLife.35946 (2018).

55 Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347-351, doi:10.1038/nature17964 (2016).
56 Thonghin, N., Kargas, V., Clews, J. & Ford, R. C. Cryo-electron microscopy of membrane proteins. *Methods* **147**, 176-186, doi:10.1016/j.ymeth.2018.04.018 (2018).
57 Liang, B. & Tamm, L. K. NMR as a tool to investigate the structure, dynamics and function of membrane proteins. *Nat Struct Mol Biol* **23**, 468-474, doi:10.1038/nsmb.3226 (2016).
58 Danmaliki, G. I. & Hwang, P. M. Solution NMR spectroscopy of membrane proteins. *Biochim Biophys Acta Biomembr* **1862**, 183356, doi:10.1016/j.bbamem.2020.183356 (2020).
59 Chipot, C. *et al.* Perturbations of Native Membrane Protein Structure in Alkyl Phosphocholine Detergents: A Critical Assessment of NMR and Biophysical Studies. *Chem Rev* **118**, 3559-3607, doi:10.1021/acs.chemrev.7b00570 (2018).
60 Huang, Y. *et al.* Use of paramagnetic (19)F NMR to monitor domain movement in a glutamate transporter homolog. *Nat Chem Biol* **16**, 1006-1012, doi:10.1038/s41589-020-0561-6 (2020).
61 Bibow, S. & Hiller, S. A guide to quantifying membrane protein dynamics in lipids and other native-like environments by solution-state NMR spectroscopy. *FEBS J* **286**, 1610-1623, doi:10.1111/febs.14639 (2019).
62 Engel, A. & Gaub, H. E. Structure and mechanics of membrane proteins. *Annu Rev Biochem* **77**, 127-148, doi:10.1146/annurev.biochem.77.062706.154450 (2008).
63 Loura, L. M. & Prieto, M. FRET in Membrane Biophysics: An Overview. *Front Physiol* **2**, 82, doi:10.3389/fphys.2011.00082 (2011).
64 Marsh, D. Electron spin resonance in membrane research: protein-lipid interactions from challenging beginnings to state of the art. *Eur Biophys J* **39**, 513-525, doi:10.1007/s00249-009-0512-3 (2010).
65 Chavent, M., Duncan, A. L. & Sansom, M. S. Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Curr Opin Struct Biol* **40**, 8-16, doi:10.1016/j.sbi.2016.06.007 (2016).
66 Duncan, A. L., Song, W. & Sansom, M. S. P. Lipid-Dependent Regulation of Ion Channels and G Protein-Coupled Receptors: Insights from Structures and Simulations. *Annu Rev Pharmacol Toxicol* **60**, 31-50, doi:10.1146/annurev-pharmtox-010919-023411 (2020).
67 Hedger, G. & Sansom, M. S. P. Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations. *Biochim Biophys Acta* **1858**, 2390-2400, doi:10.1016/j.bbamem.2016.02.037 (2016).
68 Goossens, K. & De Winter, H. Molecular Dynamics Simulations of Membrane Proteins: An Overview. *J Chem Inf Model* **58**, 2193-2202, doi:10.1021/acs.jcim.8b00639 (2018).
69 Corey, R. A., Stansfeld, P. J. & Sansom, M. S. P. The energetics of protein-lipid interactions as viewed by molecular simulations. *Biochem Soc Trans* **48**, 25-37, doi:10.1042/BST20190149 (2020).
70 Barber, M., Bordoli, R. S., Sedgwick, R. D. & Tyler, A. N. Fast Atom Bombardment of Solids as an Ion-Source in Mass-Spectrometry. *Nature* **293**, 270-275, doi:DOI 10.1038/293270a0 (1981).
71 Barber, M., Bordoli, R. S., Elliott, G. J., Sedgwick, R. D. & Tyler, A. N. Fast Atom Bombardment Mass-Spectrometry. *Anal Chem* **54**, A645-& (1982).
72 Karas, M. & Hillenkamp, F. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons. *Anal Chem* **60**, 2299-2301, doi:DOI 10.1021/ac00171a028 (1988).
73 Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray Ionization for Mass-Spectrometry of Large Biomolecules. *Science* **246**, 64-71, doi:DOI 10.1126/science.2675315 (1989).

74 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, doi:10.1016/j.cbpa.2004.08.006 (2004).
75 Leney, A. C. & Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectr* **28**, 5-13, doi:10.1007/s13361-016-1545-3 (2017).
76 Heck, A. J. R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Methods* **5**, 927-933, doi:10.1038/nmeth.1265 (2008).
77 Allison, T. M. & Agasid, M. T. Native Protein Mass Spectrometry. *Methods Mol Biol* **2073**, 287-299, doi:10.1007/978-1-4939-9869-2_15 (2020).
78 Juraschek, R., Dulcks, T. & Karas, M. Nanoelectrospray - More than just a minimized-flow electrospray ionization source. *J Am Soc Mass Spectr* **10**, 300-308, doi:Doi 10.1016/S1044-0305(98)00157-3 (1999).
79 Vahidi, S., Stocks, B. B. & Konermann, L. Partially disordered proteins studied by ion mobility-mass spectrometry: implications for the preservation of solution phase structure in the gas phase. *Anal Chem* **85**, 10471-10478, doi:10.1021/ac402490r (2013).
80 Testa, L., Brocca, S. & Grandori, R. Charge-surface correlation in electrospray ionization of folded and unfolded proteins. *Anal Chem* **83**, 6459-6463, doi:10.1021/ac201740z (2011).
81 Boeri Erba, E., Signor, L. & Petosa, C. Exploring the structure and dynamics of macromolecular complexes by native mass spectrometry. *J Proteomics* **222**, 103799, doi:10.1016/j.jprot.2020.103799 (2020).
82 Donnelly, D. P. *et al.* Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat Methods* **16**, 587-594, doi:10.1038/s41592-019-0457-0 (2019).
83 Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* **2**, 715-726, doi:10.1038/nprot.2007.73 (2007).
84 Verkerk, U. H. & Kebarle, P. Ion-ion and ion-molecule reactions at the surface of proteins produced by nanospray. Information on the number of acidic residues and control of the number of ionized acidic and basic residues. *J Am Soc Mass Spectrom* **16**, 1325-1341, doi:10.1016/j.jasms.2005.03.018 (2005).
85 Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry. *J Am Soc Mass Spectrom* **28**, 1827-1835, doi:10.1007/s13361-017-1739-3 (2017).
86 Hedges, J. B., Vahidi, S., Yue, X. & Konermann, L. Effects of ammonium bicarbonate on the electrospray mass spectra of proteins: evidence for bubble-induced unfolding. *Anal Chem* **85**, 6469-6476, doi:10.1021/ac401020s (2013).
87 Xia, Z., DeGrandchamp, J. B. & Williams, E. R. Native mass spectrometry beyond ammonium acetate: effects of nonvolatile salts on protein stability and structure. *Analyst* **144**, 2565-2573, doi:10.1039/c9an00266a (2019).
88 Susa, A. C., Xia, Z. & Williams, E. R. Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment. *Anal Chem* **89**, 3116-3122, doi:10.1021/acs.analchem.6b04897 (2017).
89 Benesch, J. L. P. Collisional Activation of Protein Complexes: Picking Up the Pieces. *J Am Soc Mass Spectr* **20**, 341-348, doi:10.1016/j.jasms.2008.11.014 (2009).
90 Stiving, A. Q. *et al.* Surface-Induced Dissociation: An Effective Method for Characterization of Protein Quaternary Structure. *Anal Chem*, doi:10.1021/acs.analchem.8b05071 (2018).
91 Marsh, J. A. *et al.* Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell* **153**, 461-470, doi:10.1016/j.cell.2013.02.044 (2013).
92 Ahnert, S. E., Marsh, J. A., Hernandez, H., Robinson, C. V. & Teichmann, S. A. Principles of assembly reveal a periodic table of protein complexes. *Science* **350**, aaa2245, doi:10.1126/science.aaa2245 (2015).

93 Barth, M. & Schmidt, C. Native mass spectrometry-A valuable tool in structural biology. *J Mass Spectrom* **55**, e4578, doi:10.1002/jms.4578 (2020).
94 Gabelica, V. & Marklund, E. Fundamentals of ion mobility spectrometry. *Curr Opin Chem Biol* **42**, 51-59, doi:10.1016/j.cbpa.2017.10.022 (2018).
95 Ben-Nissan, G. & Sharon, M. The application of ion-mobility mass spectrometry for structure/function investigation of protein complexes. *Curr Opin Chem Biol* **42**, 25-33, doi:10.1016/j.cbpa.2017.10.026 (2018).
96 Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J. & Benesch, J. L. Collision cross sections for structural proteomics. *Structure* **23**, 791-799, doi:10.1016/j.str.2015.02.010 (2015).
97 Ruotolo, B. T. *et al.* Evidence for macromolecular protein rings in the absence of bulk water. *Science* **310**, 1658-1661, doi:10.1126/science.1120177 (2005).
98 Hopper, J. T. & Oldham, N. J. Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability. *J Am Soc Mass Spectrom* **20**, 1851-1858, doi:10.1016/j.jasms.2009.06.010 (2009).
99 Dixit, S. M., Polasky, D. A. & Ruotolo, B. T. Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Curr Opin Chem Biol* **42**, 93-100, doi:10.1016/j.cbpa.2017.11.010 (2018).
100 El-Baba, T. J. *et al.* Melting Proteins: Evidence for Multiple Stable Structures upon Thermal Denaturation of Native Ubiquitin from Ion Mobility Spectrometry-Mass Spectrometry Measurements. *J Am Chem Soc* **139**, 6306-6309, doi:10.1021/jacs.7b02774 (2017).
101 El-Baba, T. J. & Clemmer, D. E. Solution thermochemistry of concanavalin A tetramer conformers measured by variable-temperature ESI-IMS-MS. *Int J Mass Spectrom* **443**, 93-100, doi:10.1016/j.ijms.2019.06.004 (2019).
102 Hall, Z., Politis, A., Bush, M. F., Smith, L. J. & Robinson, C. V. Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics. *J Am Chem Soc* **134**, 3429-3438, doi:10.1021/ja2096859 (2012).
103 Jurneczko, E. & Barran, P. E. How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst* **136**, 20-28, doi:10.1039/c0an00373e (2011).
104 Boeri Erba, E., Barylyuk, K., Yang, Y. & Zenobi, R. Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal Chem* **83**, 9251-9259, doi:10.1021/ac201576e (2011).
105 Kitova, E. N., El-Hawiet, A., Schnier, P. D. & Klassen, J. S. Reliable determinations of protein-ligand interactions by direct ESI-MS measurements. Are we there yet? *J Am Soc Mass Spectrom* **23**, 431-441, doi:10.1007/s13361-011-0311-9 (2012).
106 Southall, N. T., Dill, K. A. & Haymet, A. D. J. A view of the hydrophobic effect. *Journal of Physical Chemistry B* **106**, 521-533, doi:10.1021/jp015514e (2002).
107 Tanford, C. The hydrophobic effect and the organization of living matter. *Science* **200**, 1012-1018, doi:10.1126/science.653353 (1978).
108 Maibaum, L., Dinner, A. R. & Chandler, D. Micelle formation and the hydrophobic effect. *Journal of Physical Chemistry B* **108**, 6778-6781, doi:10.1021/jp037487t (2004).
109 Matysiak, S., Debenedetti, P. G. & Rossky, P. J. Role of hydrophobic hydration in protein stability: a 3D water-explicit protein model exhibiting cold and heat denaturation. *J Phys Chem B* **116**, 8095-8104, doi:10.1021/jp3039175 (2012).
110 van Gils, J. H. M. *et al.* The hydrophobic effect characterises the thermodynamic signature of amyloid fibril growth. *Plos Comput Biol* **16**, doi:ARTN e1007767
111 Camilloni, C. *et al.* Towards a structural biology of the hydrophobic effect in protein folding. *Sci Rep-Uk* **6**, doi:ARTN 28285

112 Snyder, P. W. *et al.* Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proc Natl Acad Sci U S A* **108**, 17889-17894, doi:10.1073/pnas.1114107108 (2011).
113 Frank, H. S. & Evans, M. W. Free Volume and Entropy in Condensed Systems .3. Entropy in Binary Liquid Mixtures - Partial Molal Entropy in Dilute Solutions - Structure and Thermodynamics in Aqueous Electrolytes. *J Chem Phys* **13**, 507-532, doi:Doi 10.1063/1.1723985 (1945).
114 Schauperl, M., Podewitz, M., Waldner, B. J. & Liedl, K. R. Enthalpic and Entropic Contributions to Hydrophobicity. *J Chem Theory Comput* **12**, 4600-4610, doi:10.1021/acs.jctc.6b00422 (2016).
115 Setny, P., Baron, R. & McCammon, J. A. How Can Hydrophobic Association Be Enthalpy Driven? *Journal of Chemical Theory and Computation* **6**, 2866-2871, doi:10.1021/ct1003077 (2010).
116 Dias, C. L. *et al.* The hydrophobic effect and its role in cold denaturation. *Cryobiology* **60**, 91-99, doi:10.1016/j.cryobiol.2009.07.005 (2010).
117 Bich, C., Baer, S., Jecklin, M. C. & Zenobi, R. Probing the hydrophobic effect of noncovalent complexes by mass spectrometry. *J Am Soc Mass Spectrom* **21**, 286-289, doi:10.1016/j.jasms.2009.10.012 (2010).
118 Daniel, J. M., Friess, S. D., Rajagopalan, S., Wendt, S. & Zenobi, R. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *International Journal of Mass Spectrometry* **216**, 1-27, doi:10.1016/S1387-3806(02)00585-7 (2002).
119 Bakhtiari, M. & Konermann, L. Protein Ions Generated by Native Electrospray Ionization: Comparison of Gas Phase, Solution, and Crystal Structures. *J Phys Chem B* **123**, 1784-1796, doi:10.1021/acs.jpcb.8b12173 (2019).
120 Liko, I. *et al.* Lipid binding attenuates channel closure of the outer membrane protein OmpF. *Proc Natl Acad Sci U S A* **115**, 6691-6696, doi:10.1073/pnas.1721152115 (2018).
121 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, doi:10.1021/cr068289b (2007).
122 Catherman, A. D., Skinner, O. S. & Kelleher, N. L. Top Down proteomics: facts and perspectives. *Biochem Biophys Res Commun* **445**, 683-693, doi:10.1016/j.bbrc.2014.02.041 (2014).
123 Belov, M. E. *et al.* From Protein Complexes to Subunit Backbone Fragments: A Multi-stage Approach to Native Mass Spectrometry. *Anal Chem* **85**, 11163-11173, doi:10.1021/ac4029328 (2013).
124 Skinner, O. S. *et al.* Top-down characterization of endogenous protein complexes with native proteomics. *Nat Chem Biol* **14**, 36-41, doi:10.1038/nchembio.2515 (2018).
125 Lermyte, F., Valkenborg, D., Loo, J. A. & Sobott, F. Radical solutions: Principles and application of electron-based dissociation in mass spectrometry-based analysis of protein structure. *Mass Spectrom Rev* **37**, 750-771, doi:10.1002/mas.21560 (2018).
126 O'Brien, J. P., Li, W. Z., Zhang, Y. & Brodbelt, J. S. Characterization of Native Protein Complexes Using Ultraviolet Photodissociation Mass Spectrometry. *J Am Chem Soc* **136**, 12920-12928, doi:10.1021/ja505217w (2014).
127 Sipe, S. N., Patrick, J. W., Laganowsky, A. & Brodbelt, J. S. Enhanced Characterization of Membrane Protein Complexes by Ultraviolet Photodissociation Mass Spectrometry. *Anal Chem* **92**, 899-907, doi:10.1021/acs.analchem.9b03689 (2020).
128 Giladi, M. & Khananshvili, D. Hydrogen-Deuterium Exchange Mass-Spectrometry of Secondary Active Transporters: From Structural Dynamics to Molecular Mechanisms. *Front Pharmacol* **11**, 70, doi:10.3389/fphar.2020.00070 (2020).

129 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, doi:10.3389/fphys.2017.00892 (2017).
130 Marklund, E. G. & Benesch, J. L. Weighing-up protein dynamics: the combination of native mass spectrometry and molecular dynamics simulations. *Curr Opin Struct Biol* **54**, 50-58, doi:10.1016/j.sbi.2018.12.011 (2019).
131 Landreh, M. *et al.* Integrating mass spectrometry with MD simulations reveals the role of lipids in Na+/H+ antiporters. *Nat Commun* **8**, 13993, doi:10.1038/ncomms13993 (2017).
132 Rauschenbach, S., Ternes, M., Harnau, L. & Kern, K. Mass Spectrometry as a Preparative Tool for the Surface Science of Large Molecules. *Annu Rev Anal Chem (Palo Alto Calif)* **9**, 473-498, doi:10.1146/annurev-anchem-071015-041633 (2016).
133 Rout, M. P. & Sali, A. Principles for Integrative Structural Biology Studies. *Cell* **177**, 1384-1403, doi:10.1016/j.cell.2019.05.016 (2019).
134 Barrera, N. P., Di Bartolo, N., Booth, P. J. & Robinson, C. V. Micelles protect membrane complexes from solution to vacuum. *Science* **321**, 243-246, doi:10.1126/science.1159292 (2008).
135 le Maire, M., Champeil, P. & Moller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta* **1508**, 86-111, doi:10.1016/s0304-4157(00)00010-1 (2000).
136 Prive, G. G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397, doi:10.1016/j.ymeth.2007.01.007 (2007).
137 Reading, E. *et al.* The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase. *Angew Chem Int Edit* **54**, 4577-4581, doi:10.1002/anie.201411622 (2015).
138 Mehmood, S. *et al.* Charge Reduction Stabilizes Intact Membrane Protein Complexes for Mass Spectrometry. *J Am Chem Soc* **136**, 17010-17012, doi:10.1021/ja510283g (2014).
139 Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat Protoc* **8**, 639-651, doi:10.1038/nprot.2013.024 (2013).
140 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, doi:10.1038/s41586-018-0325-6 (2018).
141 Urner, L. H. *et al.* Modular detergents tailor the purification and structural analysis of membrane proteins including G-protein coupled receptors. *Nat Commun* **11**, 564, doi:10.1038/s41467-020-14424-8 (2020).
142 Agah, S. & Faham, S. Crystallization of membrane proteins in bicelles. *Methods Mol Biol* **914**, 3-16, doi:10.1007/978-1-62703-023-6_1 (2012).
143 Hopper, J. T. S. *et al.* Detergent-free mass spectrometry of membrane protein complexes. *Nat Methods* **10**, 1206-+, doi:10.1038/Nmeth.2691 (2013).
144 Popot, J. L. Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu Rev Biochem* **79**, 737-775, doi:10.1146/annurev.biochem.052208.114057 (2010).
145 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, doi:10.1016/j.ijms.2015.06.017 (2015).
146 Bayburt, T. H. & Sligar, S. G. Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Sci* **12**, 2476-2481, doi:10.1110/ps.03267503 (2003).
147 Marty, M. T. Nanodiscs and mass spectrometry: Making membranes fly. *Int J Mass Spectrom* **458**, doi:10.1016/j.ijms.2020.116436 (2020).
148 Lee, S. C. *et al.* A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat Protoc* **11**, 1149-1162, doi:10.1038/nprot.2016.070 (2016).

149 Hellwig, N. *et al.* Native mass spectrometry goes more native: investigation of membrane protein complexes directly from SMALPs. *Chem Commun (Camb)* **54**, 13702-13705, doi:10.1039/c8cc06284f (2018).
150 Chorev, D. S. *et al.* Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science* **362**, 829-834, doi:10.1126/science.aau0976 (2018).
151 Landreh, M., Marty, M. T., Gault, J. & Robinson, C. V. A sliding selectivity scale for lipid binding to membrane proteins. *Current Opinion in Structural Biology* **39**, 54-60, doi:10.1016/j.sbi.2016.04.005 (2016).
152 Bechara, C. *et al.* A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nat Chem* **7**, 255-262, doi:10.1038/Nchem.2172 (2015).
153 Gupta, K. *et al.* Identifying key membrane protein lipid interactions using mass spectrometry. *Nat Protoc* **13**, 1106-1120, doi:10.1038/nprot.2018.014 (2018).
154 Bolla, J. R. *et al.* A Mass-Spectrometry-Based Approach to Distinguish Annular and Specific Lipid Binding to Membrane Proteins. *Angew Chem Int Ed Engl* **59**, 3523-3528, doi:10.1002/anie.201914411 (2020).
155 Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat Methods* **13**, 333-336, doi:10.1038/nmeth.3771 (2016).
156 Gault, J. *et al.* Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. *Nat Methods* **17**, 505-508, doi:10.1038/s41592-020-0821-0 (2020).
157 Bakelar, J., Buchanan, S. K. & Noinaj, N. The structure of the beta-barrel assembly machinery complex. *Science* **351**, 180-186, doi:10.1126/science.aad3460 (2016).
158 Cong, X., Liu, Y., Liu, W., Liang, X. & Laganowsky, A. Allosteric modulation of protein-protein interactions by individual lipid binding events. *Nat Commun* **8**, 2203, doi:10.1038/s41467-017-02397-0 (2017).
159 Patrick, J. W. *et al.* Allostery revealed within lipid binding events to membrane proteins. *Proc Natl Acad Sci U S A* **115**, 2976-2981, doi:10.1073/pnas.1719813115 (2018).
160 Cong, X. *et al.* Determining Membrane Protein-Lipid Binding Thermodynamics Using Native Mass Spectrometry. *J Am Chem Soc* **138**, 4346-4349, doi:10.1021/jacs.6b01771 (2016).
161 Zhong, Y. Y., Hyung, S. J. & Ruotolo, B. T. Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions. *Analyst* **136**, 3534-3541, doi:10.1039/c0an00987c (2011).
162 Giles, K. *et al.* A Cyclic Ion Mobility-Mass Spectrometry System. *Anal Chem* **91**, 8564-8573, doi:10.1021/acs.analchem.9b01838 (2019).
163 Bush, M. F. *et al.* Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal Chem* **82**, 9557-9565, doi:10.1021/ac1022953 (2010).
164 Fort, K. L. *et al.* Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. *Analyst* **143**, 100-105, doi:10.1039/c7an01629h (2017).
165 Zhou, M. W., Huang, C. S. & Wysocki, V. H. Surface-Induced Dissociation of Ion Mobility-Separated Noncovalent Complexes in a Quadrupole/Time-of-Flight Mass Spectrometer. *Anal Chem* **84**, 6016-6023, doi:10.1021/ac300810u (2012).
166 Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal Chem* **85**, 2-9, doi:10.1021/ac302789c (2013).
167 Wilm, M. & Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal Chem* **68**, 1-8, doi:DOI 10.1021/ac9509519 (1996).
168 Kebarle, P. & Verkerk, U. H. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev* **28**, 898-917, doi:10.1002/mas.20247 (2009).
169 Taylor, G. Disintegration of Water Drops in Electric Field. *Proc R Soc Lon Ser-A* **280**, 383-+, doi:DOI 10.1098/rspa.1964.0151 (1964).

170 Cloupeau, M. & Prunetfoch, B. Electrohydrodynamic Spraying Functioning Modes - a Critical-Review. *J Aerosol Sci* **25**, 1021-1036, doi:Doi 10.1016/0021-8502(94)90199-6 (1994).
171 Rayleigh, L. XX.On the equilibrium of liquid conducting masses charged with electricity. *Philosophical Magazine Series 5* **14**, 184-186, doi:10.1080/14786448208628425 (1882).
172 de la Mora, J. F. The fluid dynamics of Taylor cones. *Annu Rev Fluid Mech* **39**, 217-243, doi:10.1146/annurev.fluid.39.050905.110159 (2007).
173 Iribarne, J. V. & Thomson, B. A. Evaporation of Small Ions from Charged Droplets. *J Chem Phys* **64**, 2287-2294, doi:Doi 10.1063/1.432536 (1976).
174 Konermann, L., Metwally, H., Duez, Q. & Peters, I. Charging and supercharging of proteins for mass spectrometry: recent insights into the mechanisms of electrospray ionization. *Analyst*, doi:10.1039/c9an01201j (2019).
175 Aliyari, E. & Konermann, L. Formation of Gaseous Proteins via the Ion Evaporation Model (IEM) in Electrospray Mass Spectrometry. *Anal Chem* **92**, 10807-10814, doi:10.1021/acs.analchem.0c02290 (2020).
176 Dole, M. *et al.* Molecular Beams of Macroions. *J Chem Phys* **49**, 2240-&, doi:Doi 10.1063/1.1670391 (1968).
177 de la Mora, J. F. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal Chim Acta* **406**, 93-104 (2000).
178 Kaltashov, I. A. & Mohimen, A. Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal Chem* **77**, 5370-5379, doi:10.1021/ac050511 (2005).
179 Heck, A. J. & Van Den Heuvel, R. H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom Rev* **23**, 368-389, doi:10.1002/mas.10081 (2004).
180 Metwally, H., Duez, Q. & Konermann, L. Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry. *Anal Chem* **90**, 10069-10077, doi:10.1021/acs.analchem.8b02926 (2018).
181 Hogan, C. J., Carroll, J. A., Rohrs, H. W., Biswas, P. & Gross, M. L. Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Anal Chem* **81**, 369-377, doi:10.1021/ac8016532 (2009).
182 Consta, S. & Malevanets, A. Manifestations of charge induced instability in droplets effected by charged macromolecules. *Phys Rev Lett* **109**, 148301, doi:10.1103/PhysRevLett.109.148301 (2012).
183 Hopper, J. T. & Oldham, N. J. Alkali metal cation-induced destabilization of gas-phase protein-ligand complexes: consequences and prevention. *Anal Chem* **83**, 7472-7479, doi:10.1021/ac201686f (2011).
184 Liko, I., Hopper, J. T., Allison, T. M., Benesch, J. L. & Robinson, C. V. Negative Ions Enhance Survival of Membrane Protein Complexes. *J Am Soc Mass Spectrom* **27**, 1099-1104, doi:10.1007/s13361-016-1381-5 (2016).
185 Patrick, J. W. & Laganowsky, A. Generation of Charge-Reduced Ions of Membrane Protein Complexes for Native Ion Mobility Mass Spectrometry Studies. *J Am Soc Mass Spectrom*, doi:10.1007/s13361-019-02187-6 (2019).
186 Laszlo, K. J., Munger, E. B. & Bush, M. F. Folding of Protein Ions in the Gas Phase after Cation -to-Anion Proton-Transfer Reactions. *J Am Chem Soc* **138**, 9581-9588, doi:10.1021/Jacs.6b04282 (2016).
187 Laszlo, K. J., Buckner, J. H., Munger, E. B. & Bush, M. F. Native-Like and Denatured Cytochrome c Ions Yield Cation-to-Anion Proton Transfer Reaction Products with Similar Collision Cross-Sections. *J Am Soc Mass Spectrom*, doi:10.1007/s13361-017-1620-4 (2017).
188 Gadzuk-Shea, M. M. & Bush, M. F. Effects of Charge State on the Structures of Serum Albumin Ions in the Gas Phase: Insights from Cation-to-Anion Proton-Transfer Reactions, Ion Mobility, and Mass Spectrometry. *J Phys Chem B* **122**, 9947-9955, doi:10.1021/acs.jpcb.8b08427 (2018).

189 Laszlo, K. J. & Bush, M. F. Interpreting the Collision Cross Sections of Native-like Protein Ions: Insights from Cation-to-Anion Proton-Transfer Reactions. *Anal Chem* **89**, 7607-7614, doi:10.1021/acs.analchem.7b01474 (2017).
190 Takats, Z., Wiseman, J. M. & Cooks, R. G. Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *Journal of Mass Spectrometry* **40**, 1261-1275, doi:10.1002/jms.922 (2005).
191 Ambrose, S. *et al.* Native Desorption Electrospray Ionization Liberates Soluble and Membrane Protein Complexes from Surfaces. *Angew Chem Int Ed Engl* **56**, 14463-14468, doi:10.1002/anie.201704849 (2017).
192 Wattenberg, A., Sobott, F., Barth, H. D. & Brutschy, B. Studying noncovalent protein complexes in aqueous solution with laser desorption mass spectrometry. *International Journal of Mass Spectrometry* **203**, 49-57, doi:Doi 10.1016/S1387-3806(00)00261-X (2000).
193 Henrich, E. *et al.* Analyzing native membrane protein assembly in nanodiscs by combined non-covalent mass spectrometry and synthetic biology. *Elife* **6**, doi:10.7554/eLife.20954 (2017).
194 Douglas, D. J. & French, J. B. Collisional Focusing Effects in Radio-Frequency Quadrupoles. *J Am Soc Mass Spectr* **3**, 398-408, doi:Doi 10.1016/1044-0305(92)87067-9 (1992).
195 Chernushevich, I. V. & Thomson, B. A. Collisional cooling of large ions in electrospray mass spectrometry. *Anal Chem* **76**, 1754-1760, doi:10.1021/ac035406j (2004).
196 Sobott, F., Hernandez, 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, doi:10.1021/ac0110552 (2002).
197 Waters. Dramatically enhanced analytical sensitivity with the use of novel StepWave ion transfer technology in the SYNAPT G2-S system. *Waters Application Note Literature No. 720003964EN* (2011).
198 Wilson, J. W., Donor, M. T., Shepherd, S. O. & Prell, J. S. Increasing Collisional Activation of Protein Complexes Using Smaller Aperture Source Sampling Cones on a Synapt Q-IM-TOF Instrument with a Stepwave Source. *J Am Soc Mass Spectr* **31**, 1751-1754, doi:10.1021/jasms.0c00117 (2020).
199 Giles, K. *et al.* Applications of a travelling wave-based radio-frequencyonly stacked ring ion guide. *Rapid Commun Mass Sp* **18**, 2401-2414, doi:10.1002/rcm.1641 (2004).
200 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-+, doi:10.1038/Nmeth.2208 (2012).
201 Dawson, P. H. Quadrupole Mass Analyzers - Performance, Design and Some Recent Applications. *Mass Spectrometry Reviews* **5**, 1-37, doi:DOI 10.1002/mas.1280050102 (1986).
202 Hilton, G. R. & Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *J R Soc Interface* **9**, 801-816, doi:10.1098/rsif.2011.0823 (2012).
203 Loo, J. A. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews* **16**, 1-23 (1997).
204 Lossl, P., Snijder, J. & Heck, A. J. R. Boundaries of Mass Resolution in Native Mass Spectrometry. *J Am Soc Mass Spectr* **25**, 906-917, doi:10.1007/s13361-014-0874-3 (2014).
205 Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. & Zagulin, V. A. Mass-Reflectron a New Nonmagnetic Time-of-Flight High-Resolution Mass-Spectrometer. *Zh Eksp Teor Fiz+* **64**, 82-89 (1973).
206 Wollnik, H. Time-of-Flight Mass Analyzers. *Mass Spectrometry Reviews* **12**, 89-114, doi:DOI 10.1002/mas.1280120202 (1993).
207 Chernushevich, I. V., Verentchikov, A. N., Ens, W. & Standing, K. G. Effect of ion-molecule collisions in the vacuum chamber of an electrospray time-of-flight mass spectrometer on mass

spectra of proteins. *J Am Soc Mass Spectrom* **7**, 342-349, doi:10.1016/1044-0305(95)00681-8 (1996).
208 Makarov, A. Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Anal Chem* **72**, 1156-1162, doi:DOI 10.1021/ac991131p (2000).
209 Zubarev, R. A. & Makarov, A. Orbitrap Mass Spectrometry. *Anal Chem* **85**, 5288-5296, doi:10.1021/ac4001223 (2013).
210 Poltash, M. L. *et al.* Fourier Transform-Ion Mobility-Orbitrap Mass Spectrometer: A Next-generation Instrument for Native Mass Spectrometry. *Anal Chem*, doi:10.1021/acs.analchem.8b02463 (2018).
211 Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S. J. & Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc* **3**, 1139-1152, doi:10.1038/nprot.2008.78 (2008).
212 Mason, E. A. & Schamp, H. W. Mobility of Gaseous Ions in Weak Electric Fields. *Ann Phys-New York* **4**, 233-270, doi:Doi 10.1016/0003-4916(58)90049-6 (1958).
213 Pringle, S. D. *et al.* An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry* **261**, 1-12, doi:10.1016/j.ijms.2006.07.021 (2007).
214 Dodds, J. N., May, J. C. & McLean, J. A. Correlating Resolving Power, Resolution, and Collision Cross Section: Unifying Cross-Platform Assessment of Separation Efficiency in Ion Mobility Spectrometry. *Anal Chem* **89**, 12176-12184, doi:10.1021/acs.analchem.7b02827 (2017).
215 Giles, K., Williams, J. P. & Campuzano, I. Enhancements in travelling wave ion mobility resolution. *Rapid Commun Mass Sp* **25**, 1559-1566, doi:10.1002/rcm.5013 (2011).
216 Stiving, A. Q., Jones, B. J., Ujma, J., Giles, K. & Wysocki, V. H. Collision Cross Sections of Charge-Reduced Proteins and Protein Complexes: A Database for Collision Cross Section Calibration. *Anal Chem* **92**, 4475-4483, doi:10.1021/acs.analchem.9b05519 (2020).
217 Allison, T. M., Landreh, M., Benesch, J. L. P. & Robinson, C. V. Low Charge and Reduced Mobility of Membrane Protein Complexes Has Implications for Calibration of Collision Cross Section Measurements. *Anal Chem*, doi:10.1021/acs.analchem.6b00691 (2016).
218 Allison, T. M. *et al.* Computational Strategies and Challenges for Using Native Ion Mobility Mass Spectrometry in Biophysics and Structural Biology. *Anal Chem* **92**, 10872-10880, doi:10.1021/acs.analchem.9b05791 (2020).
219 Mack, E. Average cross-sectional areas of molecules by gaseous diffusion methods. *J Am Chem Soc* **47**, 2468-2482, doi:DOI 10.1021/ja01687a007 (1925).
220 Benesch, J. L. & Ruotolo, B. T. Mass spectrometry: come of age for structural and dynamical biology. *Curr Opin Struct Biol* **21**, 641-649, doi:10.1016/j.sbi.2011.08.002 (2011).
221 Shvartsburg, A. A. & Smith, R. D. Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem* **80**, 9689-9699, doi:10.1021/ac8016295 (2008).
222 Koeniger, S. L., Merenbloom, S. I., Sevugarajan, S. & Clemmer, D. E. Transfer of structural elements from compact to extended states in unsolvated ubiquitin. *J Am Chem Soc* **128**, 11713-11719, doi:10.1021/ja062137g (2006).
223 Kemper, P. R., Dupuis, N. F. & Bowers, M. T. A new, higher resolution, ion mobility mass spectrometer. *International Journal of Mass Spectrometry* **287**, 46-57, doi:10.1016/j.ijms.2009.01.012 (2009).
224 Eldrid, C. *et al.* Gas Phase Stability of Protein Ions in a Cyclic Ion Mobility Spectrometry Traveling Wave Device. *Anal Chem* **91**, 7554-7561, doi:10.1021/acs.analchem.8b05641 (2019).
225 Ruotolo, B. T. *et al.* Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes. *Angew Chem Int Edit* **46**, 8001-8004, doi:10.1002/anie.200702161 (2007).
226 Wanasundara, S. N. & Thachuk, M. Theoretical investigations of the dissociation of charged protein complexes in the gas phase. *J Am Soc Mass Spectr* **18**, 2242-2253, doi:10.1016/j.jasms.2007.09.022 (2007).

227 Benesch, J. L. P., Aquilina, J. A., Ruotolo, B. T., Sobott, F. & Robinson, C. V. Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chem Biol* **13**, 597-605, doi:10.1016/j.chembiol.2006.04.006 (2006).
228 Thachuk, M., Fegan, S. K. & Raheem, N. Description and control of dissociation channels in gas-phase protein complexes. *J Chem Phys* **145**, doi:10.1063/1.4960615 (2016).
229 Fegan, S. K. & Thachuk, M. A Charge Moving Algorithm for Molecular Dynamics Simulations of Gas-Phase Proteins. *J Chem Theory Comput* **9**, 2531-2539, doi:10.1021/ct300906a (2013).
230 Popa, V., Trecroce, D. A., McAllister, R. G. & Konermann, L. Collision-Induced Dissociation of Electrosprayed Protein Complexes: An All-Atom Molecular Dynamics Model with Mobile Protons. *J Phys Chem B* **120**, 5114-5124, doi:10.1021/acs.jpcb.6b03035 (2016).
231 Zhong, Y., Han, L. & Ruotolo, B. T. Collisional and Coulombic unfolding of gas-phase proteins: high correlation to their domain structures in solution. *Angew Chem Int Ed Engl* **53**, 9209-9212, doi:10.1002/anie.201403784 (2014).
232 Eschweiler, J. D., Martini, R. M. & Ruotolo, B. T. Chemical Probes and Engineered Constructs Reveal a Detailed Unfolding Mechanism for a Solvent-Free Multidomain Protein. *J Am Chem Soc* **139**, 534-540, doi:10.1021/jacs.6b11678 (2017).
233 Allison, T. M. *et al.* Quantifying the stabilizing effects of protein-ligand interactions in the gas phase. *Nat Commun* **6**, doi:10.1038/ncomms9551 (2015).
234 Landreh, M., Costeira-Paulo, J., Gault, J., Marklund, E. G. & Robinson, C. V. Effects of Detergent Micelles on Lipid Binding to Proteins in Electrospray Ionization Mass Spectrometry. *Anal Chem* **89**, 7425-7430, doi:10.1021/acs.analchem.7b00922 (2017).
235 Aquilina, J. A. The major toxin from the Australian Common Brown Snake is a hexamer with unusual gas-phase dissociation properties. *Proteins* **75**, 478-485, doi:10.1002/prot.22259 (2009).
236 van den Heuvel, R. H. H. *et al.* Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* **78**, 7473-7483, doi:10.1021/ac061039a (2006).
237 Loo, R. R. & Loo, J. A. Salt Bridge Rearrangement (SaBRe) Explains the Dissociation Behavior of Noncovalent Complexes. *J Am Soc Mass Spectrom* **27**, 975-990, doi:10.1007/s13361-016-1375-3 (2016).
238 Mistarz, U. H., Chandler, S. A., Brown, J. M., Benesch, J. L. P. & Rand, K. D. Probing the Dissociation of Protein Complexes by Means of Gas-Phase H/D Exchange Mass Spectrometry. *J Am Soc Mass Spectrom*, doi:10.1007/s13361-018-2064-1 (2018).
239 Cooks, R. G., Terwilliger, D. T., Ast, T., Beynon, J. H. & Keough, T. Surface Modified Mass-Spectrometry. *J Am Chem Soc* **97**, 1583-1585, doi:DOI 10.1021/ja00839a056 (1975).
240 Mabud, M. D. A., Dekrey, M. J. & Cooks, R. G. Surface-Induced Dissociation of Molecular-Ions. *Int J Mass Spectrom* **67**, 285-294, doi:Doi 10.1016/0168-1176(85)83024-X (1985).
241 Snyder, D. T., Panczyk, E. M., Somogyi, A., Kaplan, D. A. & Wysocki, V. Simple and Minimally Invasive SID Devices for Native Mass Spectrometry. *Anal Chem* **92**, 11195-11203, doi:10.1021/acs.analchem.0c01657 (2020).
242 Wysocki, V. H., Joyce, K. E., Jones, C. M. & Beardsley, R. L. Surface-induced dissociation of small molecules, peptides,and non-covalent protein complexes. *J Am Soc Mass Spectr* **19**, 190-208, doi:10.1016/j.jasms.2007.11.005 (2008).
243 Nuzzo, R. G. & Allara, D. L. Adsorption of Bifunctional Organic Disulfides on Gold Surfaces. *J Am Chem Soc* **105**, 4481-4483, doi:DOI 10.1021/ja00351a063 (1983).
244 Porter, M. D., Bright, T. B., Allara, D. L. & Chidsey, C. E. D. Spontaneously Organized Molecular Assemblies .4. Structural Characterization of Normal-Alkyl Thiol Monolayers on Gold by Optical Ellipsometry, Infrared-Spectroscopy, and Electrochemistry. *J Am Chem Soc* **109**, 3559-3568, doi:DOI 10.1021/ja00246a011 (1987).

245 Somogyi, A., Kane, T. E., Ding, J. M. & Wysocki, V. H. Reactive Collisions of C6h6.+ and C6d6.+ at Self-Assembled Monolayer Films Prepared on Gold from N-Alkanethiols and a Fluorinated Alkanethiol - the Influence of Chain-Length on the Reactivity of the Films and the Neutralization of the Projectile. *J Am Chem Soc* **115**, 5275-5283, doi:DOI 10.1021/ja00065a044 (1993).
246 Morris, M. R. *et al.* Ion Surface Collisions at Functionalized Self-Assembled Monolayer Surfaces. *Int J Mass Spectrom* **122**, 181-217, doi:Doi 10.1016/0168-1176(92)87016-8 (1992).
247 Grill, V., Shen, J., Evans, C. & Cooks, R. G. Collisions of ions with surfaces at chemically relevant energies: Instrumentation and phenomena. *Rev Sci Instrum* **72**, 3149-3179, doi:Doi 10.1063/1.1382641 (2001).
248 Jones, C. M. *et al.* Symmetrical gas-phase dissociation of noncovalent protein complexes via surface collisions. *J Am Chem Soc* **128**, 15044-15045, doi:10.1021/ja064586m (2006).
249 Beardsley, R. L., Jones, C. M., Galhena, A. S. & Wysocki, V. H. Noncovalent Protein Tetramers and Pentamers with "n" Charges Yield Monomers with n/4 and n/5 Charges. *Anal Chem* **81**, 1347-1356, doi:10.1021/ac801883k (2009).
250 Galhena, A. S., Dagan, S., Jones, C. M., Beardsley, R. L. & Wysocki, V. N. Surface-induced dissociation of peptides and protein complexes in a quadrupole/Time-of-Flight mass spectrometer. *Anal Chem* **80**, 1425-1436, doi:DOI 10.1021/ac701782q (2008).
251 Zhou, M. W., Dagan, S. & Wysocki, V. H. Protein Subunits Released by Surface Collisions of Noncovalent Complexes: Nativelike Compact Structures Revealed by Ion Mobility Mass Spectrometry. *Angew Chem Int Edit* **51**, 4336-4339, doi:10.1002/anie.201108700 (2012).
252 Levy, E. D., Erba, E. B., Robinson, C. V. & Teichmann, S. A. Assembly reflects evolution of protein complexes. *Nature* **453**, 1262-U1266, doi:10.1038/nature06942 (2008).
253 Villar, G. *et al.* Self-Assembly and Evolution of Homomeric Protein Complexes. *Phys Rev Lett* **102**, doi:10.1103/PhysRevLett.102.118106 (2009).
254 Quintyn, R. S., Yan, J. & Wysocki, V. H. Surface-Induced Dissociation of Homotetramers with D-2 Symmetry Yields their Assembly Pathways and Characterizes the Effect of Ligand Binding. *Chem Biol* **22**, 583-592, doi:10.1016/j.chembiol.2015.03.019 (2015).
255 Harvey, S. R. *et al.* Relative interfacial cleavage energetics of protein complexes revealed by surface collisions. *Proc Natl Acad Sci U S A* **116**, 8143-8148, doi:10.1073/pnas.1817632116 (2019).
256 Seffernick, J. T., Harvey, S. R., Wysocki, V. H. & Lindert, S. Predicting Protein Complex Structure from Surface-Induced Dissociation Mass Spectrometry Data. *ACS Cent Sci* **5**, 1330-1341, doi:10.1021/acscentsci.8b00912 (2019).
257 Song, Y., Nelp, M. T., Bandarian, V. & Wysocki, V. H. Refining the Structural Model of a Heterohexameric Protein Complex: Surface Induced Dissociation and Ion Mobility Provide Key Connectivity and Topology Information. *ACS Cent Sci* **1**, 477-487, doi:10.1021/acscentsci.5b00251 (2015).
258 Zhou, M. *et al.* Surface Induced Dissociation Coupled with High Resolution Mass Spectrometry Unveils Heterogeneity of a 211 kDa Multicopper Oxidase Protein Complex. *J Am Soc Mass Spectrom* **29**, 723-733, doi:10.1007/s13361-017-1882-x (2018).
259 Sahasrabuddhe, A. *et al.* Confirmation of intersubunit connectivity and topology of designed protein complexes by native MS. *Proc Natl Acad Sci U S A* **115**, 1268-1273, doi:10.1073/pnas.1713646115 (2018).
260 Zhou, M. W., Dagan, S. & Wysocki, V. H. Impact of charge state on gas-phase behaviors of noncovalent protein complexes in collision induced dissociation and surface induced dissociation. *Analyst* **138**, 1353-1362, doi:10.1039/c2an36525a (2013).
261 Jacobs, A. D. *et al.* Resolution of Stepwise Cooperativities of Copper Binding by the Homotetrameric Copper-Sensitive Operon Repressor (CsoR): Impact on Structure and Stability. *Angew Chem Int Ed Engl* **54**, 12795-12799, doi:10.1002/anie.201506349 (2015).

262 Busch, F. *et al.* Localization of Protein Complex Bound Ligands by Surface-Induced Dissociation High-Resolution Mass Spectrometry. *Anal Chem* **90**, 12796-12801, doi:10.1021/acs.analchem.8b03263 (2018).
263 Harvey, S. R., Liu, Y., Liu, W., Wysocki, V. H. & Laganowsky, A. Surface induced dissociation as a tool to study membrane protein complexes. *Chem Commun (Camb)*, doi:10.1039/c6cc09606a (2017).
264 Stiving, A. Q., Gilbert, J. D., Jones, B. J. & Wysocki, V. H. A Tilted Surface and Ion Carpet Array for SID. *J Am Soc Mass Spectrom* **31**, 458-462, doi:10.1021/jasms.9b00009 (2020).
265 VanAernum, Z. L. *et al.* Surface-Induced Dissociation of Noncovalent Protein Complexes in an Extended Mass Range Orbitrap Mass Spectrometer. *Anal Chem* **91**, 3611-3618, doi:10.1021/acs.analchem.8b05605 (2019).
266 Yan, J. *et al.* Surface-Induced Dissociation of Protein Complexes in a Hybrid Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Anal Chem* **89**, 895-901, doi:10.1021/acs.analchem.6b03986 (2017).

2 Application of Surface-Induced Dissociation to Detect Interfacial Lipids for Membrane Protein Complexes by Native Mass Spectrometry

Abstract

Membrane proteins (MPs) are important drug targets that are very challenging to characterise using standard biophysical techniques. Native mass spectrometry (MS) has emerged as a complementary method, providing important information on MP stoichiometry and lipid binding. However, by being an inherently low structural resolution method, native MS is often unable to provide positional or functional information about the bound lipids.

In this chapter I develop a novel method which allows distinguishing between the subunit-stabilising and interface-stabilising lipids. To achieve the former, I employ collision-induced unfolding (CIU), which is a well-established method of investigating MP-lipid interactions. To discover the interface-stabilising lipids, I use surface-induced dissociation (SID), which has been extensively characterised for soluble proteins but so far only had limited examples of MP-related applications.

Here, I demonstrate for the first time the ability of SID to distinguish between the lipid classes and successfully identify interfacial species. Moreover, I show that by combining novel instrumentation, careful selection of conditions and mathematical and statistical analysis, it is possible to observe very small differences in interfacial stabilities of MPs induced by closely related lipid species. I describe a protocol for a combined CIU-SID experiment, which can be used to investigate specific lipids

and to determine their effect on MP structure. Finally, I outline a few potential sources of error, which can be extremely challenging to spot but which could drastically change data interpretation if they are present.

2.1 Introduction

Membrane proteins (MPs) perform many different functions, from transporting small molecules in and out of cells to being involved in complex signalling pathways[1]. They are embedded in lipid bilayers, where the proteins are surrounded by a variety of different lipid species. MPs interact with lipids selectively, often requiring phospholipids (PLs) containing specific head groups[2-4] or hydrophobic tails[5-10] for correct structure and function (see **Sections 1.1-1.3** for further discussion).

Despite the great progress in studying MP-lipid interactions, there are still many aspects that elude our understanding. Research is hindered by the difficulty in expressing and purifying MPs, the need to use membrane mimetics for *in vitro* investigations and the great diversity of lipid species, many of which are hard to differentiate[11]. Native mass spectrometry (MS) has made significant contributions to this field, identifying and investigating many MP-lipid interactions[1] (**Section 1.6**). Native MS is an inherently low structural resolution technique, so in order to obtain information on the nature of the observed binding events it has to be combined with complementary methods. Ion mobility (IM) spectrometry can be naturally coupled with native MS and allows measuring the physical dimensions, in the form of collisional cross section (CCS)[12] (**Section 1.13**). IM can be employed to investigate the ligand-induced stabilisation in a collision-induced unfolding (CIU) experiment, which involves gas phase activation of the protein and the analysis of

the resulting unfolding profiles (**Section 1.14**). CIU is a well-established method for investigating specific binding of lipids to MPs, with multiple examples of structurally important PLs discovered this way[4,13] (**Section 1.6**).

One of the main limits of CIU is that it only probes the MP structure with respect to unfolding and does not provide information on the oligomeric state stabilisation. While native MS was used to investigate such cases as well, it was mostly limited to indirect methods, such as observing the changes in abundances of different oligomeric species upon lipid addition or removal[14-17]. Although these methods are informative, various challenges exist for data interpretation, including different probabilities of ionisation, transmission and detection of species with different masses, and possible losses of bound lipids in the gas phase[18] (**Section 1.5**). Some of these factors are incredibly difficult to control, because even minor variations to the instrumental set-up (such as varying the distance between the needle and the spectrometer inlet) can dramatically change the energy of the protein ions[19].

Surface-induced dissociation (SID) is a method that can be used in combination with quadrupole selection to probe a single protein species, eliminating most of the complications described above. SID specifically breaks weaker subunit interfaces and has been successfully employed to identify interface-stabilising ligands for soluble proteins[20] (**Section 1.15**). Moreover, SID energy was shown to be directly related to the interfacial cleavage energetics, allowing for semi-quantitative experiments to be performed[21]. Overall, SID clearly has the potential to identify lipids, which are important for MP oligomerisation; however, only limited success has been achieved so far when looking at this type of interactions, with no significant stabilisation observed for a few initial case studies[22]. The main aim of

this project is to develop a method to reliably employ SID for identification of interfacial lipids that stabilise MP complexes. In this chapter, I demonstrate the ability of SID to distinguish between interfacial and non-interfacial lipids and even to differentiate between several structurally related interfacial species based on small differences in the resulting dissociation energies.

2.2 Experimental Outline

As was described above, the mechanisms of CIU and SID are the opposites of each other: 'idealised' CIU is unfolding without dissociation, while 'idealised' SID is dissociation without unfolding (**Figure 2.1**). It should be noted that such an idealised scenario rarely occurs in practice, with some degree of unfolding occurring due to the surface collision[23,24] and lipids falling off the protein during CIU[13]. However, the former can be minimised by using lower charge states[23,25], while the latter only becomes significant when looking at multiple binding events of weakly bound lipids[13]. This means that with careful selection of experimental conditions the two techniques can be employed to obtain complementary information about the lipid binding: CIU can identify the lipids that stabilise individual monomers, whereas SID can identify the lipids that strengthen the subunit interface. Here, I demonstrate this behaviour by looking at two different proteins: ammonium transporter (AmtB) from *Escherichia coli* (*E.coli*) and SemiSWEET (SS) from *Vibrio sp*. N418.

AmtB is a trimeric bacterial protein that has been extensively studied by native MS due to high expression yields and good stability in solution and gas-phase[4,6,26,27]. One of the key findings was the discovery of a highly stabilising interaction between AmtB and phosphatidylglycerol (PG) lipid (**Figure 2.2a**), which resulted in a crystal

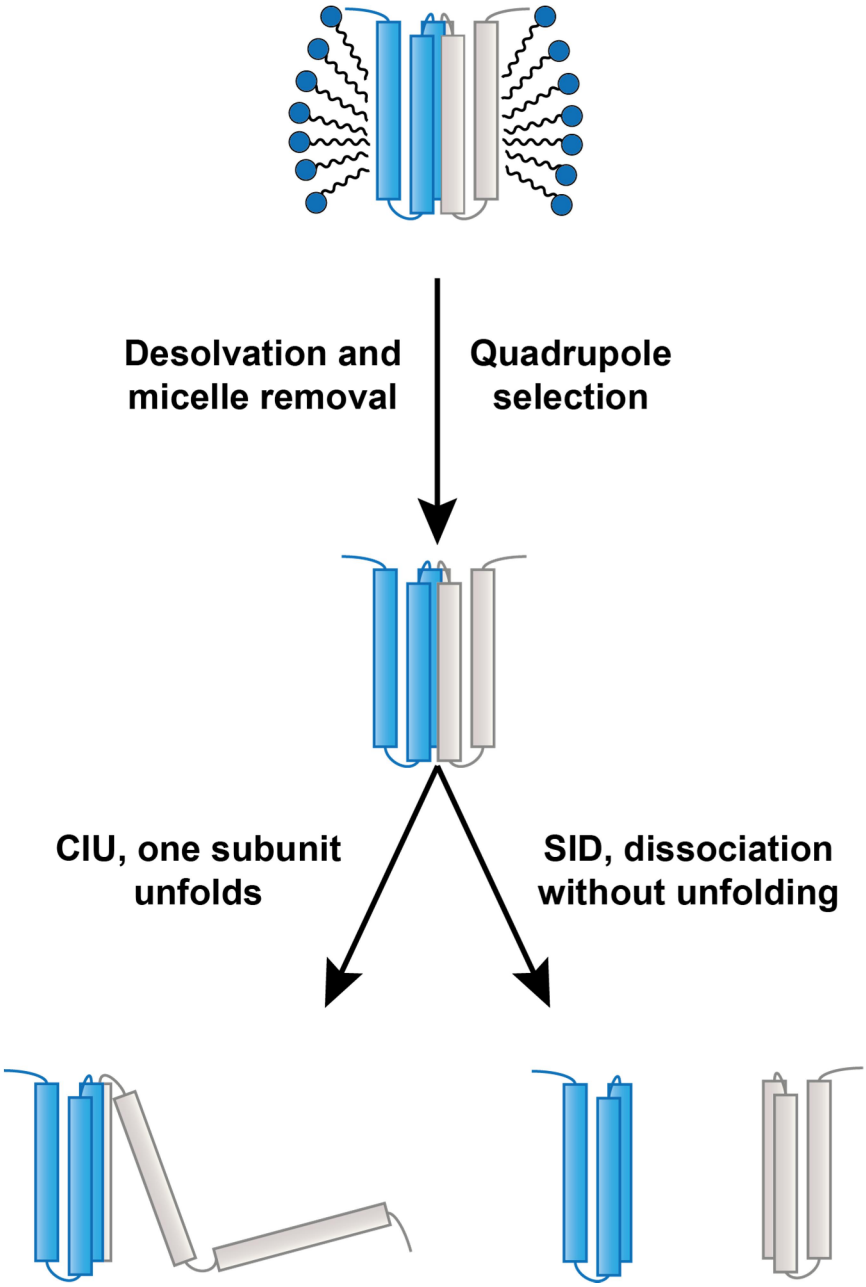


Figure 2.1 Collision-induced unfolding and surface-induced dissociation. A schematic, representing a dimeric protein (such as semiSWEET) subjected to either CIU or SID. CIU results in unfolding of a single subunit, but preserves the dimer interface. SID disrupts the interface, but subunits remain folded.

structure with a previously unobserved conformation of AmtB in the presence of this lipid species (**Figure 2.2b**)[4]. Importantly, both the CIU result and the location of the specific binding pocket in the crystal structure reveal that PG primarily stabilises the individual monomeric subunits of AmtB. Therefore, PG binding would

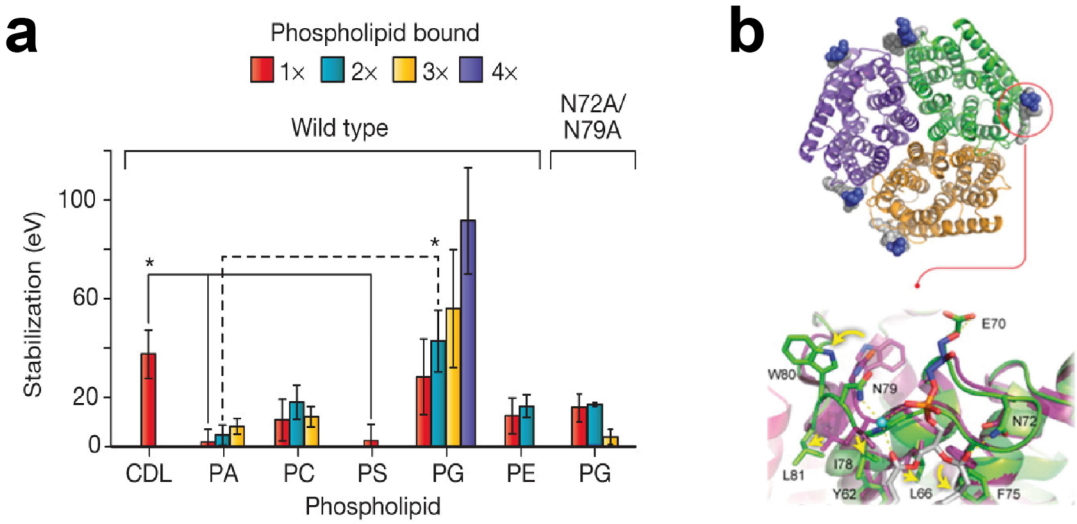


Figure 2.2 AmtB. a, Collision-induced unfolding experiments of AmtB bound to various lipids revealed the extra stabilisation by PG lipids that was abolished upon mutation. **b,** Crystal structure of AmtB bound to PG revealed a specific binding pocket within each of the monomeric subunits (top), resulting in a change of conformation compared to the *apo* protein (bottom). Figure adapted from ref(4), with permission from Springer Nature.

be expected to have no significant oligomeric stabilisation upon the AmtB trimer, which should be reflected by the results of an SID experiment according to the mechanism in **Figure 2.1**. Thus, AmtB was selected as a negative control for the SID, with expected outcome of no extra stabilisation of PG-bound species compared to the *apo* trimer. (*Note: the AmtB data were obtained prior to the publication of similar data by Harvey et al*[22]*, with the experiments designed and performed independently from them*).

SemiSWEETs are a family of bacterial dimeric sugar transporters, with each monomer containing three transmembrane domains (**Figure 2.3a**)[28,29]. Dimer is functionally important for transport, as is reflected by the SWEET proteins – SS homologues from plant, in which the two 'subunits' are genetically fused together by an extra transmembrane domain. Interestingly, SS dimers have an 'intermediate' subunit interface strength (see **Section 1.6**), meaning that while they

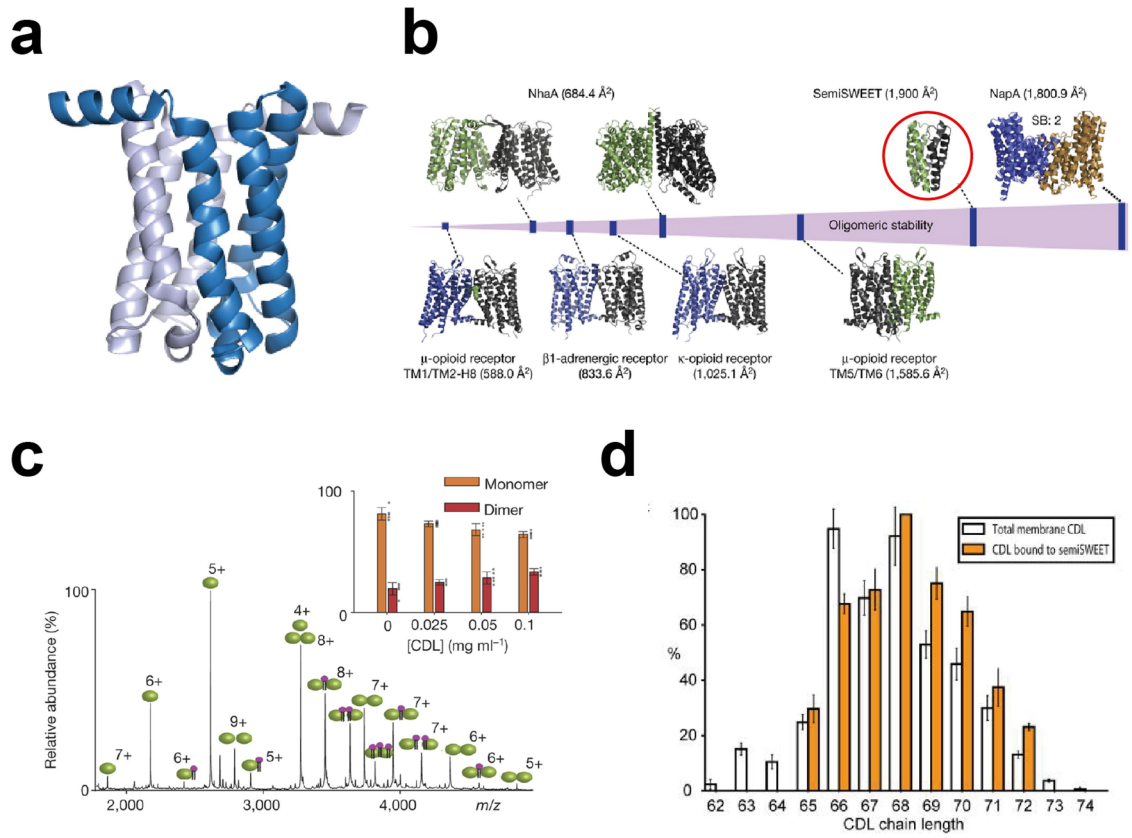


Figure 2.3 SemiSWEET. a, A crystal structure of SemiSWEET from *Vibrio sp.* (PDB 4QND). **b,** Oligomeric interface sizes for various proteins – SemiSWEET falls between weak interface proteins that require lipids for oligomerisation, and strong interface proteins that oligomerise on their own. **c,** Native mass spectrometry data revealing increased dimerisation of SemiSWEET in the presence of increasing concentrations of cardiolipin. **d,** Chain lengths of cardiolipin species present in the membrane (white) and those observed bound to SemiSWEET (orange) – a clear preference for binding of longer lipids is observed. Panels **b** and **c** adapted from ref(17), panel **d** adapted from ref(30), all with permission from Springer Nature.

do not require lipids for formation, dimerisation was shown to be influenced by cardiolipin (CDL) (**Figure 2.3 b,c**)[15]. In addition, the distribution of endogenous lipids was assigned by high resolution native MS and revealed a higher percentage of CDL molecules with long chain lengths bound to SS compared with all the lipids in the cell extract (**Figure 2.3d**)[30]. Unfortunately, this method does not allow determination of whether this distributional shift arises from the increased binding affinity of SS towards longer species or if this represents greater propensity of larger molecules to remain bound during gas phase activation. Here, I use SS to

determine whether SID can be employed to reveal stabilisation induced by CDL on the dimer interface. I then compare the results obtained upon SS binding to lipids with different hydrophobic tails to see whether SID has sufficient sensitivity and resolution to observe very subtle differences in induced stabilisation.

2.3 Surface-Induced Dissociation of AmtB

Initial SID experiments were performed on a modified Synapt G2 instrument (Waters Corporation) which was described in the introductory **Chapter 1 (Figure 1.6a)**. Briefly, the instrument was modified for high mass transmission by using a low-frequency quadrupole and enabling pressure control in the source region as previously described[31]. Another modification was the replacement of the Trap travelling wave ion guide (TWIG) with a truncated version, with an SID device consisting of 10 individually controlled electrodes installed between the Trap exit and the Helium cell entrance[32]. I have chosen the SID-IM configuration, rather than an alternative IM-SID configuration[33], in order to monitor the amount of unfolding of any dissociation products, as such unfolding represents an alternative energy pathway to the desired interfacial cleavage.

AmtB was selected as the model protein for the initial SID experiments. The first consideration for any MP is the choice of an appropriate membrane mimetic (**Section 1.6**). For AmtB, tetraethylene glycol monooctyl ether (C8E4) detergent was shown to be compatible with native MS, resulting in protein ions which retain their native-like fold after the micelle removal step[4]. C8E4 is also naturally charge reducing, which is beneficial for SID experiments as the low charge species were shown to produce more native-like dissociation products[23,25]. C8E4 was therefore an appropriate choice of detergent for this study.

Interpretation of SID data is greatly improved by the tandem mass spectrometry (MS/MS) approach. If dissociation is performed without prior quadrupole selection, a variety of different sub-species will be produced arising from all of the different charge states. This final distribution of products can be extremely difficult to analyse. Selection of an individual peak with a certain mass-to charge ratio (*m/z*) before SID alleviates this problem, as dissociation occurs only for this well-defined species. In order to perform quadrupole selection, any unwanted adducts, including detergent molecules, need to be removed from the protein in order to avoid major signal losses due to the bound protein species being outside of the quadrupole *m/z* selection window. Synapt instruments allow for a maximum application of 200 V of potential on the sample cone in the source region of the spectrometer, which proved to be insufficient for complete micelle removal from AmtB. I have taken advantage of the previous 'high-energy source' modification of the Synapt G2 that I was using, which increased the maximum potential that can be applied to the extraction cone from 10 V to 200 V[17]. Adding 50V on the extraction cone, together with 200 V on the sample cone, allowed sufficient micelle removal to perform peak selection (**Figure 2.4a**) (an alternative strategy for removing detergent micelle from AmtB was demonstrated, where instead of increasing the extraction cone potential a temperature increase was employed for a similar outcome[13]).

Three factors were considered to determine which charge state would be used for the SID experiment. Firstly, the ion should retain a compacted, folded state prior to activation in order to achieve the most 'idealised' SID dissociation (**Figure 2.1**). Secondly, the peaks must have sufficient initial intensity in order to obtain high quality data, as significant intensity drops are expected during both quadrupole selection and SID steps[34]. Thirdly, because the peaks are separated by *m/z* rather

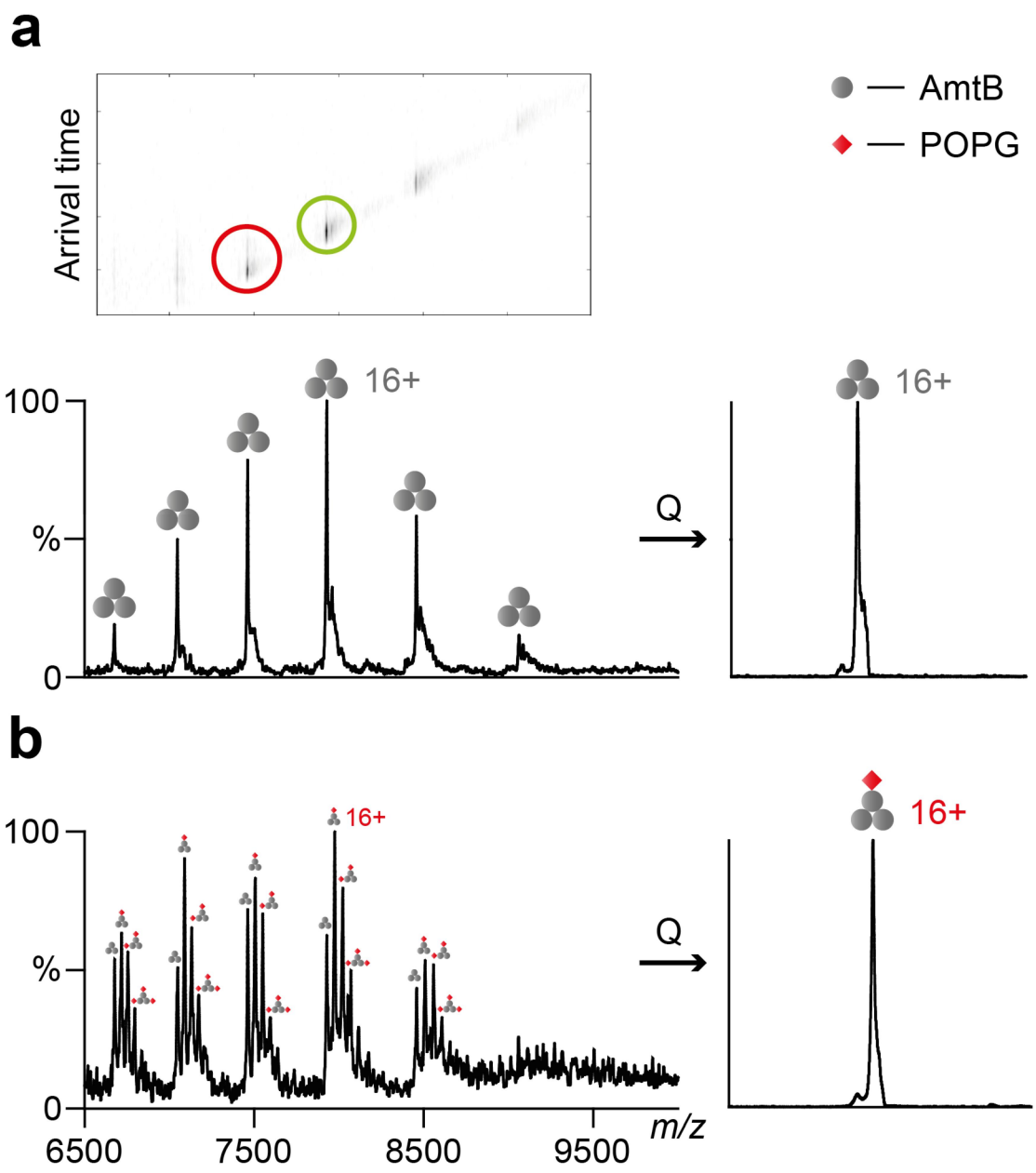


Figure 2.4 Mass spectra of AmtB. a, Charge state distribution of AmtB, with arrival time distribution obtained by ion mobility (top panel). 16+ and 17+ are the two states which met the selection criteria for dissociation (see text). Ion mobility data shows that 16+ (green circle) has a smaller arrival time distribution than 17+ (red circle, larger), indicating a more native-like conformation. Quadrupole (Q) selection of the 16+ charge state greatly reduces initial complexity. **b,** Mass spectrum of AmtB upon addition of POPG, with up to three bound lipids visible for every charge state. The 16+ charge state with a single bound POPG is selected.

than by mass, any charge states divisible by the number of monomeric subunits in a protein complex are undesirable due to a potential overlap with the resulting product species. For example, if the 18+ charge state of trimeric AmtB was

dissociated by SID, monomers with the 6+ charge and dimers with the 12+ charge would all have the same *m/z* ratio as the precursor, resulting in an overlap of all of these species in the mass spectrum. This problem is especially relevant for SID due to its propensity to create dissociation products with symmetrical charge distributions; for collision-induced dissociation (CID), formation of such species is highly unlikely in most cases[35]. With these factors in mind, 16+ and 17+ charge states were selected as the possible candidates, with 16+ chosen due to having a narrower arrival time distribution (ATD) than the 17+, indicating a more native-like state **(Figure 2.4a, top panel**). Unfolded species are associated with broader ATDs due to a variety of possible conformations that they can occupy. The exact CCS values were not calculated due to the challenging nature of MP calibration (discussed further in **Section 1.13**), however the most compact state of AmtB was shown to be correlated to the native-like conformation by previous studies performed on a drift tube instrument[4].

For the lipid binding studies, 1-Palmitoyl-2-Oleoyl PG (POPG) was exogenously added to AmtB until several lipid-bound peaks could be observed in the spectra (**Figure 2.4b**). A major advantage of native MS is the ability to see individual binding events, for example, here up to three molecules of POPG bound to each of the charge states of AmtB can be clearly resolved. I have selected the same charge state (16+) as for the *apo* AmtB protein to enable direct comparison between the two cases, because different charge states receive different amount of activation at the same potential and are also likely to have a different initial protein conformation. For the initial SID experiment, I have chosen to focus on the protein bound to one lipid molecule due to the highest abundance of such species. In addition, using a singly-bound protein allows inferring some information on the

nature of the MP-lipid interaction from the dissociation pattern. If a lipid is stabilising a subunit interface, it will be stronger than the other two and therefore will not be cleaved by SID. In this case, only the dimeric fragments will retain the lipid molecule, while all the monomers can be expected to be *apo*. Such a pattern would not be produced if there was more than one lipid present.

Ligand or lipid induced stabilisation is typically assessed by SID in the form of an energy resolved mass spectrometry (ERMS) plot. In order to perform an ERMS experiment, the SID energy is increased in a stepwise manner and the intensities of the precursor and all of the dissociation products are obtained by measuring the areas under the peaks. The relative amount of the precursor, *P*, is then calculated by the equation:

$$P = \frac{\sum I_P}{\sum I_t}, \tag{2.1}$$

where I_p is the intensity of each precursor peak and I_t is the intensity of every peak in the spectrum. While in an MS/MS experiment only one precursor peak is selected, sometimes additional peaks arise due to charge stripping – a loss of a single charge by a protein due to a loss of a proton (or possibly due to a gain of an electron) after the surface collision (discussed in the later parts of this chapter). As the energy increases, a greater proportion of the precursor is dissociated (**Figure 2.5**).

The top panel of **Figure 2.6** shows ATDs for the SID products of the *apo* AmtB. It can be seen that the vast majority of dissociation products retain a narrow ATD, indicative of them maintaining the folded state. A few unfolded species are also present, representing an alternative energy pathway. It should be noted that while

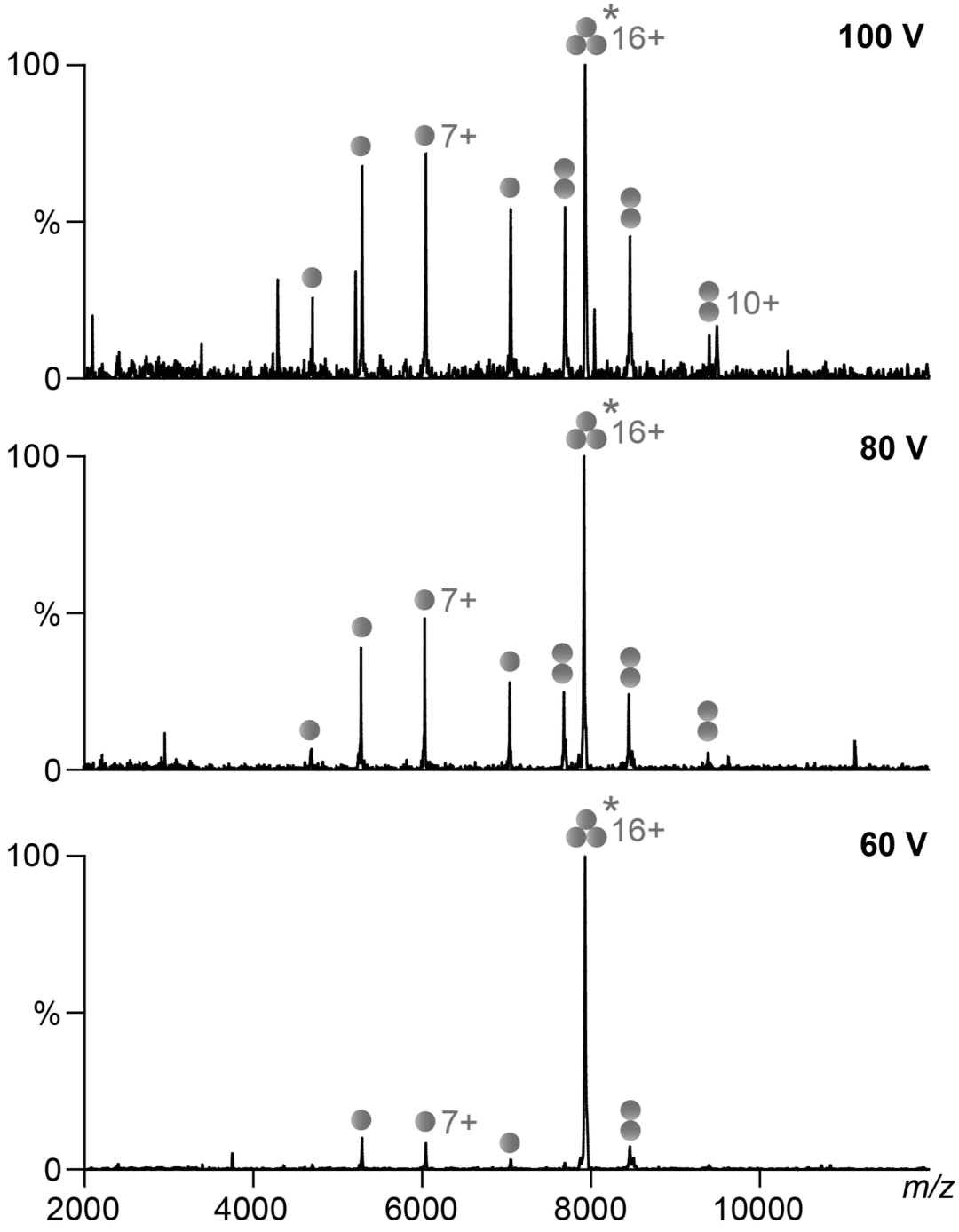


Figure 2.5 SID of AmtB. The quadrupole selected precursor peak (*) is subjected to increasing SID energy (bottom to top). The relative amount of the trimer goes down at the higher dissociation energy, while more monomers and dimers are produced. The unassigned peaks are due to detector noise.

most of the dissociation products retain their compact conformations, the surviving precursor clearly undergoes some degree of unfolding. Similar results were

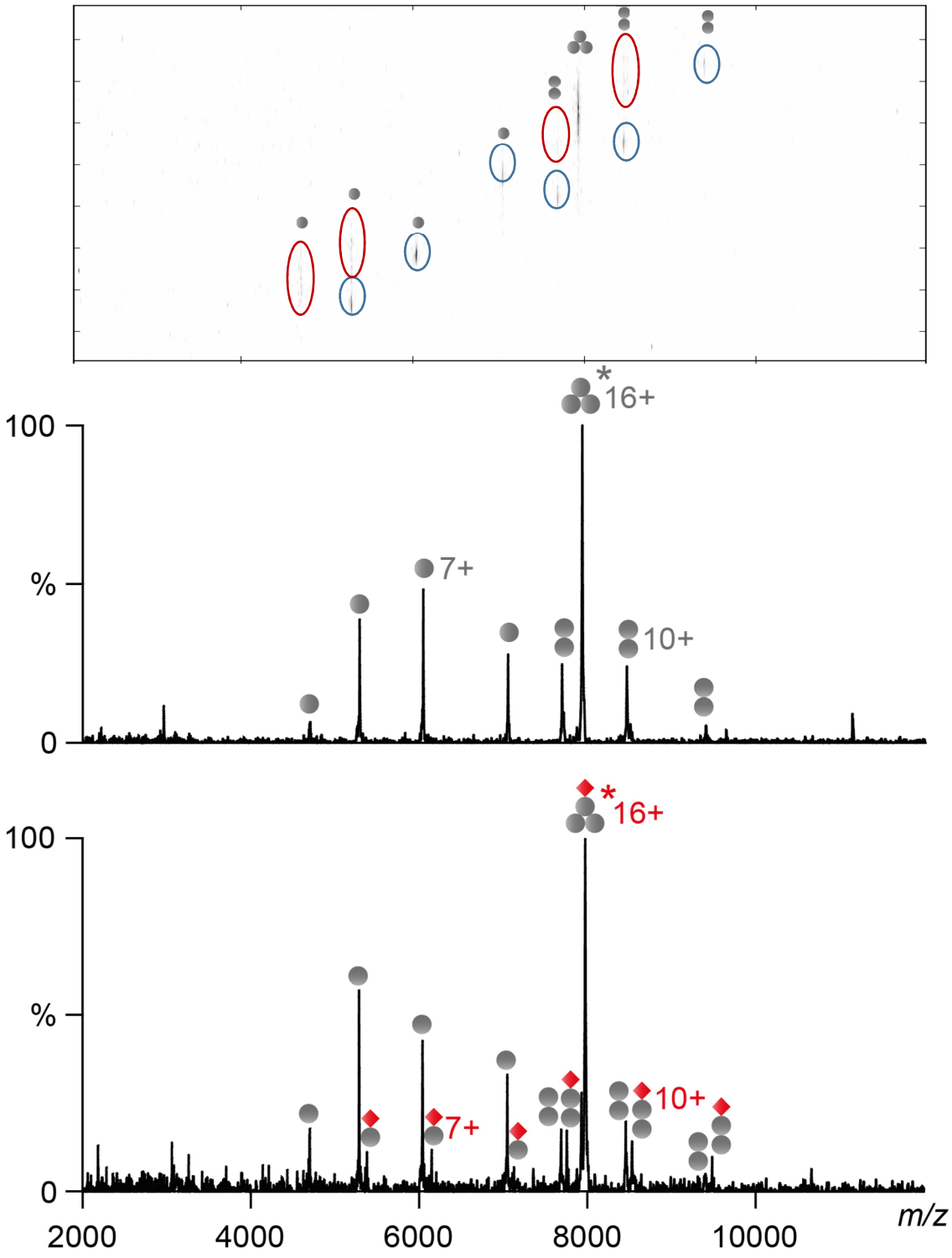


Figure 2.6 SID of AmtB bound to POPG. SID of AmtB at 80 V with (bottom) and without (middle) POPG (red). Amount of the surviving precursor is visually similar for the two cases. Both monomers and dimers retain POPG binding upon dissociation. Top panel shows arrival time distribution of SID products for *apo* AmtB – most species remain folded (blue circles), although a small amount of unfolding also occurs (red circles). The unassigned peaks in the spectra are due to detector noise.

previously observed for SID of other proteins[25]. This behaviour is discussed further in the later parts of this chapter in conjunction with SS data.

The middle and the bottom panels of **Figure 2.6** show mass spectra for the *apo* and the POPG-bound AmtB at the same SID voltage (80 V). The relative amount of the surviving precursor is similar in both cases, at least visually. In addition, both POPG-bound dimers and monomers are observed upon dissociation of the lipidated trimer. By the argument outlined above, this implies that POPG is not significantly stabilising the subunit interface. While the fraction of dimer retaining lipids is greater than the fraction of corresponding monomers, this can be expected purely on statistical grounds as dimers have twice as many subunits. In the case of oligomer-stabilising lipid, the population of lipid-bound monomers should be virtually non-existent.

In order to further verify the non-interfacial nature of POPG, the SID-ERMS plot was produced (**Figure 2.7**). No significant differences are observed between the undissociated precursors for *apo* and the POPG-bound AmtB across the whole range of voltages. This is in agreement with expected behaviour due to PG being located away from the subunit interfaces (**Figure 2.2**). This observation is in stark contrast with the CIU data, where a large stabilisation of AmtB by PG lipids is observed. Therefore, the data presented here demonstrates the proposed orthogonal nature of CIU and SID (**Figure 2.1**), so the two methods can indeed be used in combination for localisation of lipid binding. It should be noted that the independent study by Harvey *et al*. reached very similar conclusions, while starting from a 17+ charge state of AmtB, instead of the 16+ investigated here. This

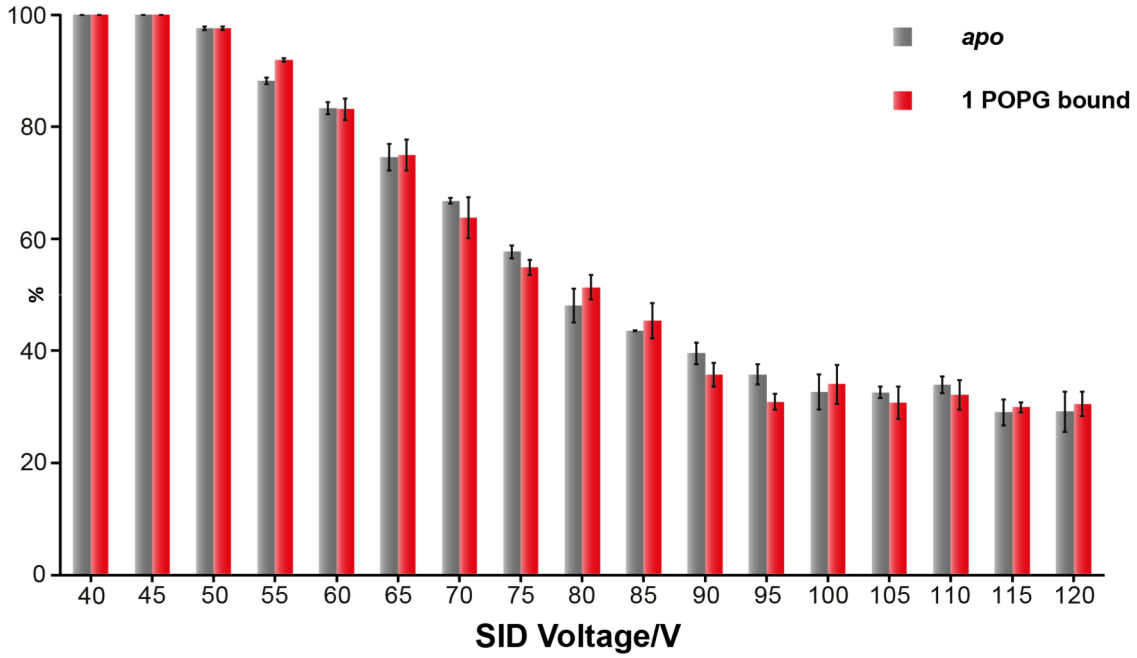


Figure 2.7 SID-ERMS plot of AmtB. The amount of the surviving precursor does not differ significantly between the *apo* and the POPG-bound AmtB (16+), indicating no stabilising effect of the lipid on the subunit interface. Data plotted as a bar chart to aid visualisation. Values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation.

confirms that SID results are not exclusive to one particular charge state and are representative of the system as a whole.

2.4 Surface-Induced Dissociation of SemiSWEET on Synapt G2

Having obtained the initial results for applicability of SID method to a model membrane protein, I have moved to a more challenging system, namely SS. As was mentioned in **Section 2.2**, stabilisation of SS dimer by CDL is expected from previous studies (**Figure 2.3**). Moreover, since SS is capable of forming dimers in the absence of lipids, a relatively small effect of CDL on the interfacial dissociation energy of the protein can be expected (although this is not certain, as factors such as large-scale structural rearrangements induced by lipid binding can play a role).

This implies that any method for measuring this energy will need to be highly reproducible, otherwise the effect can be completely obscured by the large error bars. Unfortunately, performing the SID experiments on the Synapt G2 resulted in exactly this outcome, where no significant stabilisation could be extracted from the data due to large variations between repeats. As a consequence, I switched to a more sensitive, improved version of a Synapt (**Figure 1.6b**), where I managed to obtain data of much higher quality, which I present in **Section 2.6**. For that reason, in this section I will only show the final results in terms of the ERMS plots and not any mass spectra or ion mobility measurements. I will show these data in **Section 2.6** instead, in order to avoid presenting copies of mostly identical spectra, with the only major difference being the signal-to-noise ratio (S/N). Also, some of the details briefly mentioned here will be discussed in more detail in that section as well.

The 7+ charge state was selected for SS by a similar algorithm to the one described for AmtB. After identifying suitable voltages to maintain the native fold while stripping off the C8E4 detergent (see **Section 2.6** for details), the 7+ charge state of SS was subjected to SID. For the binding partner, CDL 18:1 was chosen (all four acyl chains have 18 carbons in them, with 1 double bond on each). Selection of this particular species of CDL was inspired by **Figure 2.3d**, which showed preference of SS for similar endogenous lipids (CDL 18:1 has 72 acyl carbons in total).

The efforts to observe any interfacial stabilisation were hindered by low S/N of the data, which resulted in large variations between repeats. **Figure 2.8** shows an example of that, with 3 repeats of SID performed on the *apo* protein and two repeats of SID performed on the CDL-bound SS are plotted as individual data points. Very large differences can be clearly observed between repeats, especially

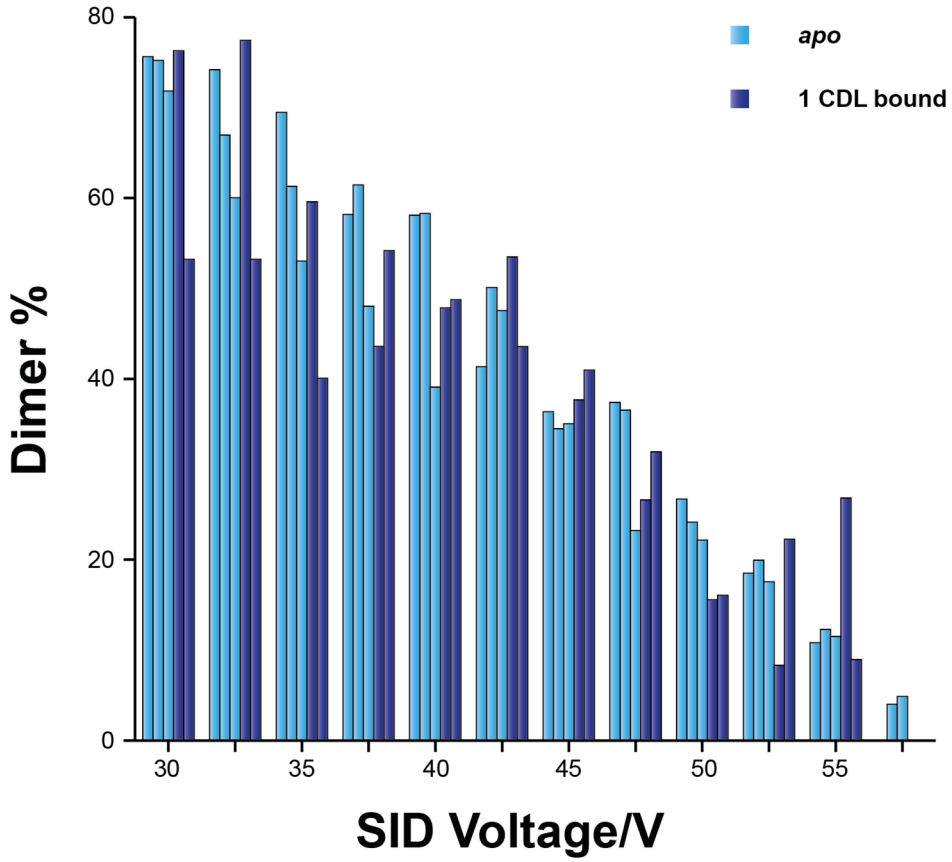


Figure 2.8 SID-ERMS of SS on Synapt G2. Three individual repeats of the *apo* and two individual repeats of the CDL-bound SS (7+) are shown. A large variation is observed between repeats, especially for the lipid-bound species. Data plotted as a bar chart to aid visualisation

the two CDL runs, which have the amount of surviving dimer vary by as much as 20% at the same voltage. Some variation is expected during native MS experiments, which is partly attributed to the needles used to initiate electrospray ionisation. In order to check for any needle effects, I repeated the experiment in **Figure 2.8**, but with the individual repeats obtained sequentially from the same needle. A similar spread of values was observed (data not shown).

I have employed multiple strategies in order to attempt increasing S/N, including increasing protein concentration, testing various settings both in the spectrometer and in the SID cell and changing the charge state to the 6+. However, obtaining data of sufficient quality was not possible. The problem was only alleviated by

switching to a different instrument (see **Section 2.6**). These results indicate that the method described in this chapter may not be suitable for the older generations of instruments, at least in the most challenging cases.

In order to better understand the limits of Synapt G2, I have looked for a system where a relatively large difference was expected. A mutation of the interfacial amino acids can have such an effect, significantly disrupting the subunit interface. Y62W mutant of SS was designed by Daniel Quetschlich using Swissprot, with expected weakening of the interface. I have obtained SID data on this mutant and compared the results against the wild type (WT) SS (**Figure 2.9**). The difference in interfacial stability was indeed sufficiently large to be clearly observed, despite the large error bars. Surprisingly, the Y62W mutant dimer was shown to be more stable than the WT SS. While this effect is very intriguing, understanding its nature is beyond the scope of current work.

In conclusion, the results discussed in this section indicate that the newest mass spectrometers are required to discern subtle stabilisation effects; however, when relatively large differences are expected, the instruments below the current state-of-the-art might well be sufficient, as was demonstrated by the data in **Figure 2.9**.

2.5 Collision-Induced Unfolding of SemiSWEET

Unlike AmtB which had been extensively characterised by CIU, SS was not investigated by this method. Therefore, I have used CIU to investigate the interaction of SS with lipids. I have chosen CDL 18:1 and CDL 14:0 to determine whether the preference for the long-chain CDLs (**Figure 2.3d**) can be observed by this method. In addition, I have selected POPG and POPE lipids as controls, because of the presence of those lipids in the native environment of *Vibrio*

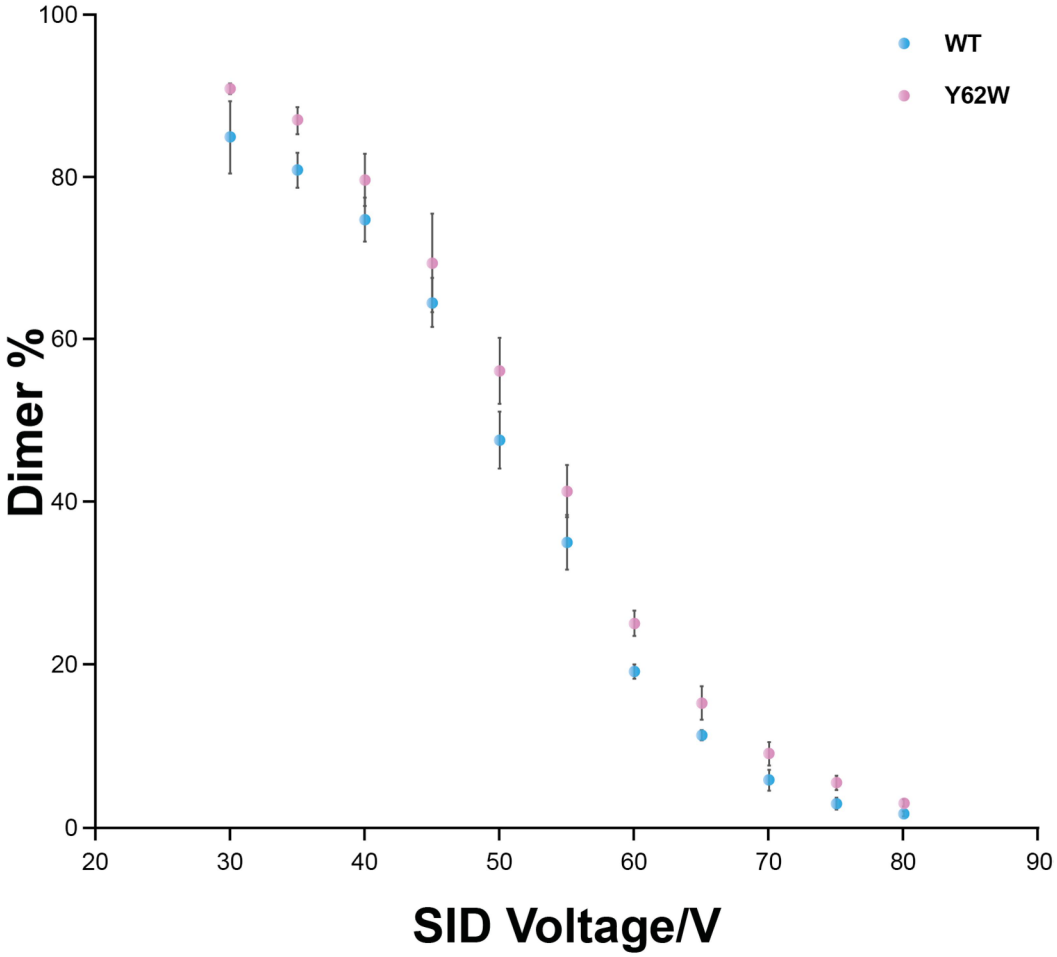


Figure 2.9 SID-ERMS plot of wild type and mutant SS on Synapt G2. A clear difference is observed between the wild type (WT) and the Y62W mutant SS dimers. Surprisingly, the mutant dimer is more stable, despite being specifically designed to have a weaker subunit interface. The range of voltages is different to the previous figure due to a change of instrumental parameters. Values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation.

bacteria[36]. The CIU experiments were performed on the 7+ charge state of SS, due to its high abundance and the native-like conformation prior to activation, and to enable direct comparison between the SID and the CIU data. The experiments were also performed on the Synapt G2; however, due to no quadrupole selection requirement and no signal losses due to imperfect refocusing of ions after the surface collision the CIU data was of a considerably higher quality and sufficiently reproducible for the purposes of this experiment.

Upon the gas-phase activation, the 7+ charge state of SS shows a single unfolding event (**Figure 2.10a**). The voltage at which this unfolding event occurs can increase upon lipid binding, indicating its stabilising nature. The stabilisation value can be calculated with special software (such as PULSAR used here), by fitting the raw data to a mathematical model[37] (**Figure 2.10a**). The stabilisation voltage can be converted into the laboratory frame collision energy (E_{lab}) by simply multiplying it by the charge of the ion (in this case, 7). The resulting plot is shown in **Figure 2.10b**. All of the lipid species result in some induced stability compared to the *apo* SS; this is a commonly observed outcome for CIU experiments[4]. Because of this fact, comparing the stabilisation values between the different lipid species is required. For SS, PE and PG show similar stabilisation values; the two CDL species are more stabilising with respect to unfolding than the smaller PLs, but are virtually identical to each other. It should be noted that the data for the two CDLs was acquired with both lipids in the same spectrum; this method minimises the 'between-samples' variations and, therefore, should be more able to detect any differences. One major distinction between SID and CIU is the fact that CIU measures the voltage while SID measures the ion intensity. This means that CIU is inherently unable to detect stability differences below the size of the 'step' of the voltage increase (for example, 2 V increments were used to record the data shown in **Figure 2.10**). In practice, this means that in order to increase the resolution, the number of experiments performed also has to be increased, significantly lengthening the process. In addition, reducing the voltage steps below a certain value (roughly 1 V) is impractical due to precision level of various electronics in the spectrometer. SID, on the other hand, measures ion intensity, so the resolution is limited only by the sensitivity of the detector and the inherent reproducibility of

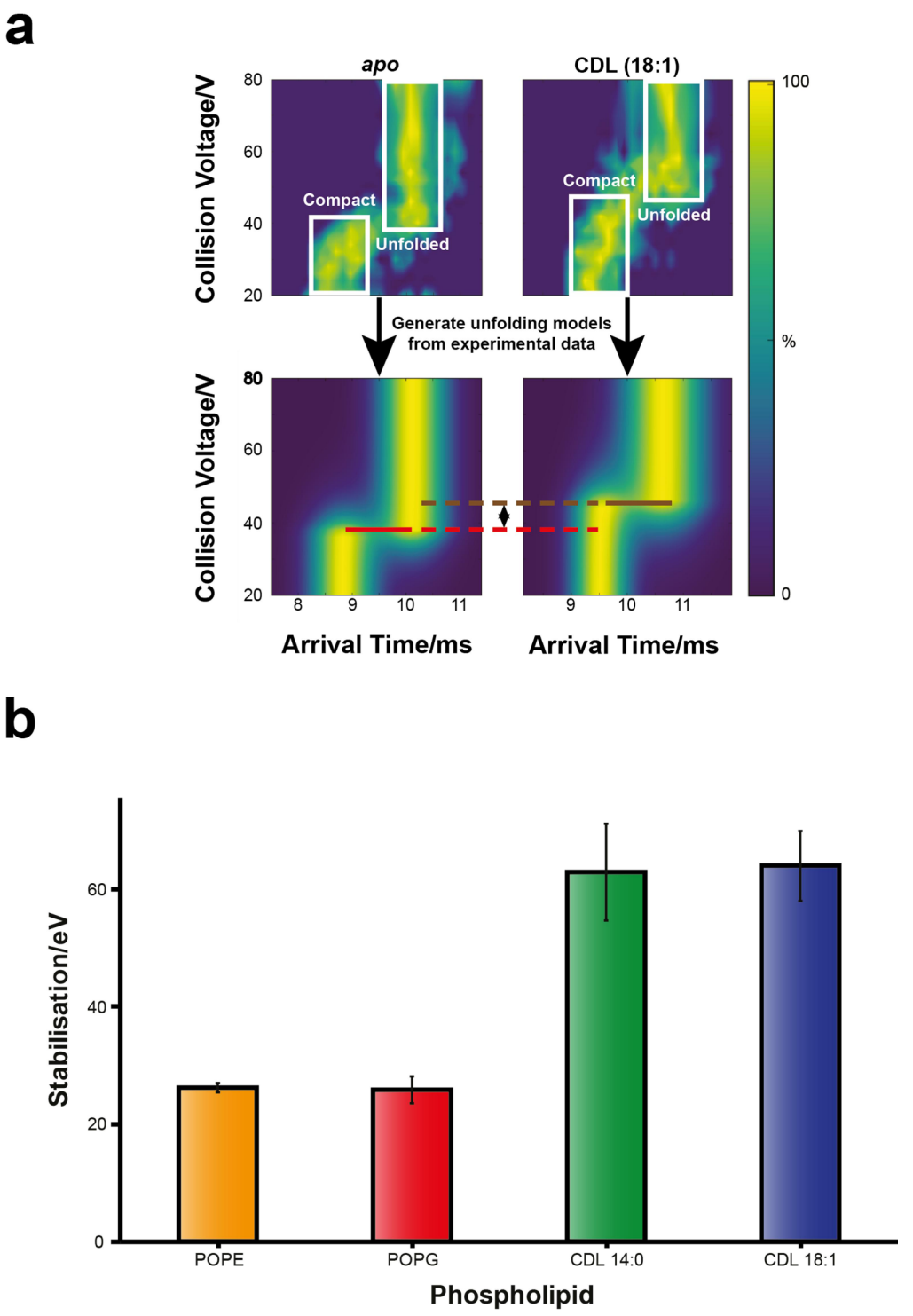


Figure 2.10 CIU of SemiSWEET. a, Activation of the 7+ charge state of SS results in a single unfolding transition. SS is stabilised by lipid binding, resulting in increased energy at which the unfolding is observed. Stabilisation can be calculated by fitting the raw data to a mathematical unfolding model. **b,** Stabilisation induced by various lipids. Collision voltage was changed in 2 V increments. Values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation.

native MS experiments. For that reason, SID is more likely to pick up very subtle differences between the samples.

CDL is expected to stabilise the subunit interface of SS based on our current understanding (**Figure 2.3**), so it is a little surprising to see a significant stabilising effect observed by CIU, especially in light of the 'orthogonal' nature of CIU and SID demonstrated in **Section 2.3** for AmtB. There are several interpretations of this outcome. One possible explanation relies on the fact that CDL is practically a 'double' lipid, with twice as many headgroups, acyl chains and roughly twice the mass of the other PLs. Considering the fact that even non-specific binding usually results in some degree of unfolding stabilisation[38], it is not improbable that doubling the mass of the lipid can roughly double the CIU effect (which is observed here). The data I present in **Chapter 5** corroborates this explanation, as CDL is not expected to be present in the native environment of the MP investigated there, but still has a large stabilisation observed by CIU. However, some counterexamples also exist: for example, mechanoselective channel protein MscL shows no particular preference for CDL compared with the other PLs[4]. An alternative interpretation is the possibility of CDL to stabilise both the dimer interface and individual subunits of SS. Considering that SS is a relatively small protein, with CDL having more than 10% of mass of SS monomer, it is not hard to imagine this lipid affecting the protein structure in several different ways at once. Regardless of whether the stabilisation observed by CIU is biologically relevant or a gas-phase artefact, it clearly does not probe the effect of CDL on the dimer stability and is unable to distinguish between the different CDL species. These issues can be addressed by SID, as is demonstrated in the next section.

2.6 Surface-Induced Dissociation of SemiSWEET on the 'High Resolution' Synapt

As was mentioned in **Section 2.4**, the sensitivity of Synapt G2 proved insufficient to perform the desired SID experiments to probe lipid-facilitated dimerisation of SS. Kevin Giles and Jakub Ujma from Waters Corporation (Wilmslow, UK) have kindly offered me to use a development version of a modified Synapt with improved sensitivity and resolution[39] (which I refer to as 'high-res' Synapt throughout this thesis). The modifications have been described in **Chapter 1 (Figure 1.6b)**. Briefly, a Stepwave ion guide in the source region and a novel segmented quadrupole preceding the extended time-of-flight (TOF) are the main improvements compared to the Synapt G2. Similarly to the other instrument, an SID cell is installed between the Trap ion guide and the IM cell.

A spectrum of SS (solubilised in C8E4 detergent) obtained on the 'high-res' Synapt is shown in **Figure 2.11**. Both monomers and dimers are observed, in agreement with previous studies[15]. Several lower intensity peaks are present around the main species – these are attributed to incomplete cleavage of the N-terminal methionine, as well as non-specific cleavage of one or two C-terminal amino acids by the protease during the His-tag cleavage step (see **Materials and Methods** for purification details). These species only differ by a small amount from the desired protein in *m/z* space. However, removing them during selection is beneficial for the later interpretation of the SID spectra. The improved sensitivity of the 'high res' Synapt enables using a very tight selection window, while still maintaining a high intensity of the precursor ion required for the SID experiment. **Figure 2.11** shows

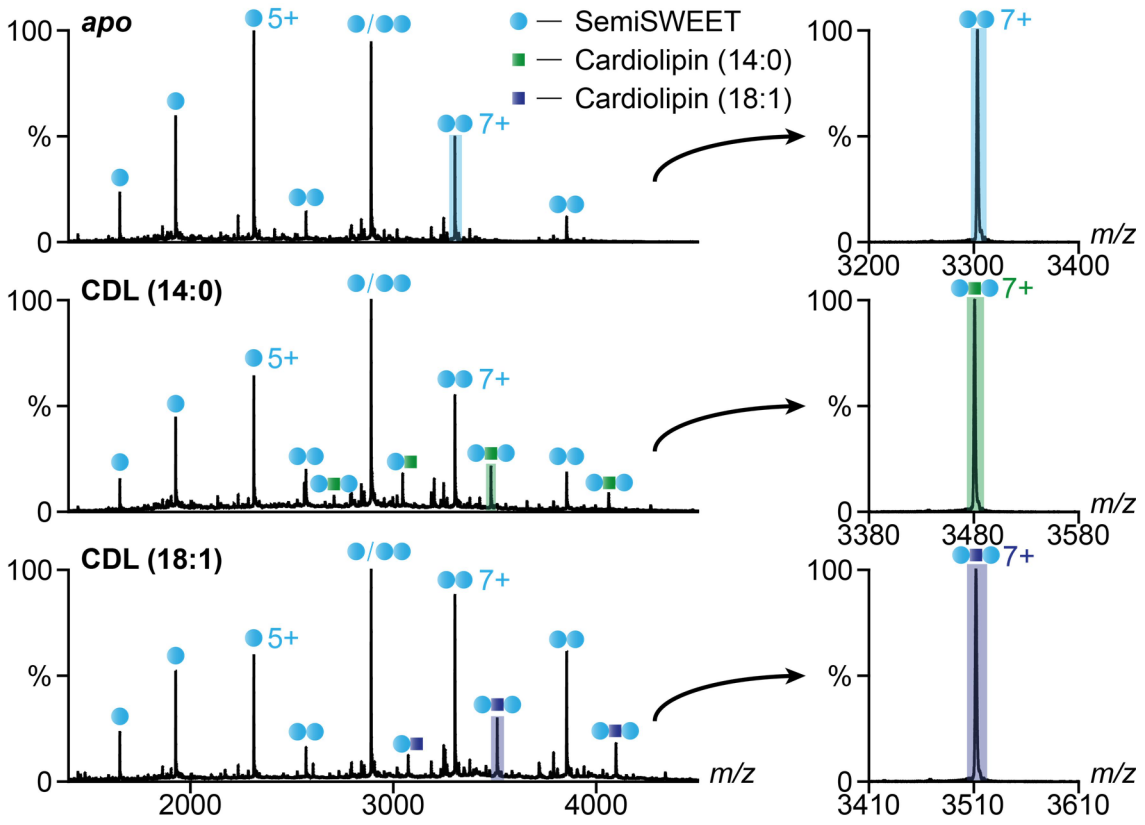


Figure 2.11 Spectra of SemiSWEET obtained on the 'high-res' Synapt. The 7+ charge state of SS was quadrupole selected for the *apo* dimer, as well as for the CDL-bound complexes. The high sensitivity of the instrument enables the use of a very tight selection window, which is beneficial for SID experiments.

a highly specific selection of a single species for the *apo* 7+ SS dimer, as well as for the two different CDL-bound species.

As was already mentioned above, maintaining the precursor in its folded state prior to SID activation is highly desirable. In order to select the optimal conditions, a cone voltage ramp was performed on the 7+ charge state of the dimer **(Figure 2.12**) (it matches the other two criteria of high abundance and not being divisible by the number of subunits outlined in **Section 2.3**). The protein maintains the most compact state until 130 V, followed by a slight increase in ATD that indicates a presence of more extended conformations, followed by an unfolding event at 160 V and a different unfolding event at 190 V. This analysis points towards the use of cone voltages of 130 V or below. However, when working with membrane proteins

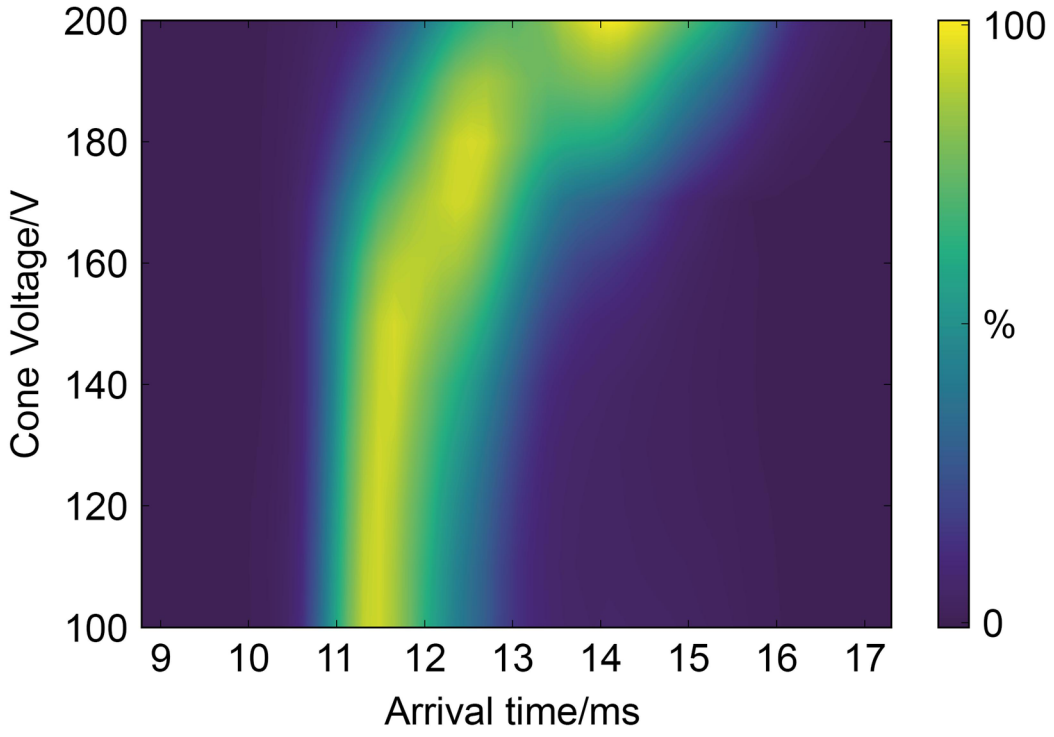


Figure 2.12 In-source unfolding of SemiSWEET. The arrival time of the 7+ charge state of SS was monitored at different cone voltages applied in the source region of the spectrometer. Changes in arrival time indicate protein unfolding. The voltage was changed in 10 V increments.

another factor has to be taken into account, namely the removal of the detergent micelle from the protein before it reaches the quadrupole. Incomplete desolvation decreases data quality, because a protein bound to even one detergent molecule differs in mass from the completely naked protein, and, therefore, is not selected by the mass filter, reducing precursor intensity for the following analysis. In the case of lipid binding experiments another complication can arise, where different combinations of protein, lipid and detergent can accidentally overlap in *m/z* space within the size of a quadrupole selection window – this can result in co-selection of unwanted species that produce their own dissociation products and make accurate data analysis extremely difficult or even impossible. This problem is discussed further in **Section 2.8**, with examples being reported for SS.

Finding the balance between maximising desolvation and minimising unfolding, the experiments were performed using a cone voltage of 150 V. It should be noted, that once the initial conditions have been selected, they must be used for all of the experimental results that will be compared to each other, including those with lipids bound. SID is extremely sensitive towards the initial conformation of the protein, so a small change in the initial conformation of precursor results in a noticeable difference in both the distribution of the dissociation products, and, importantly, the amount of energy required for dissociation[40]. Interestingly, pre-activing the protein with energetic collisions into gas molecules shifts dissociation energy not to lower, but to higher values. To our current understanding, this effect is attributed to the increased number of internal rotational modes upon unfolding, which can be used to dissipate some of the energy[24,40]. A similar effect was observed for SS, with up to 5% more dimer surviving at a given SID energy when the cone voltage of 150 V was used compared to the cone voltage of 130 V (data not shown).

The first goal of the SID experiment was to determine whether a difference in the interfacial cleavage energy can be observed upon CDL binding. **Figure 2.13a** shows an SID-ERMS plot for the 7+ dimer of apo SS, and SS bound to one CDL 14:0. (*Note: while ERMS plots usually have E_{lab} rather than voltage as one of the axes, in this case the two values can be interconverted by simply multiplying the voltage by a constant (7) as only a single charge state is investigated, meaning that no fundamental differences in terms of statistical or other analysis exist between the two plots*). There are two notable regions on that graph where the lipid-bound dimer percentage is greater than the *apo* at the same SID voltage: the first one happens around 20-30 V with a relatively large difference between the two species; the second one is more subtle and occurs around 50 V. It should be noted

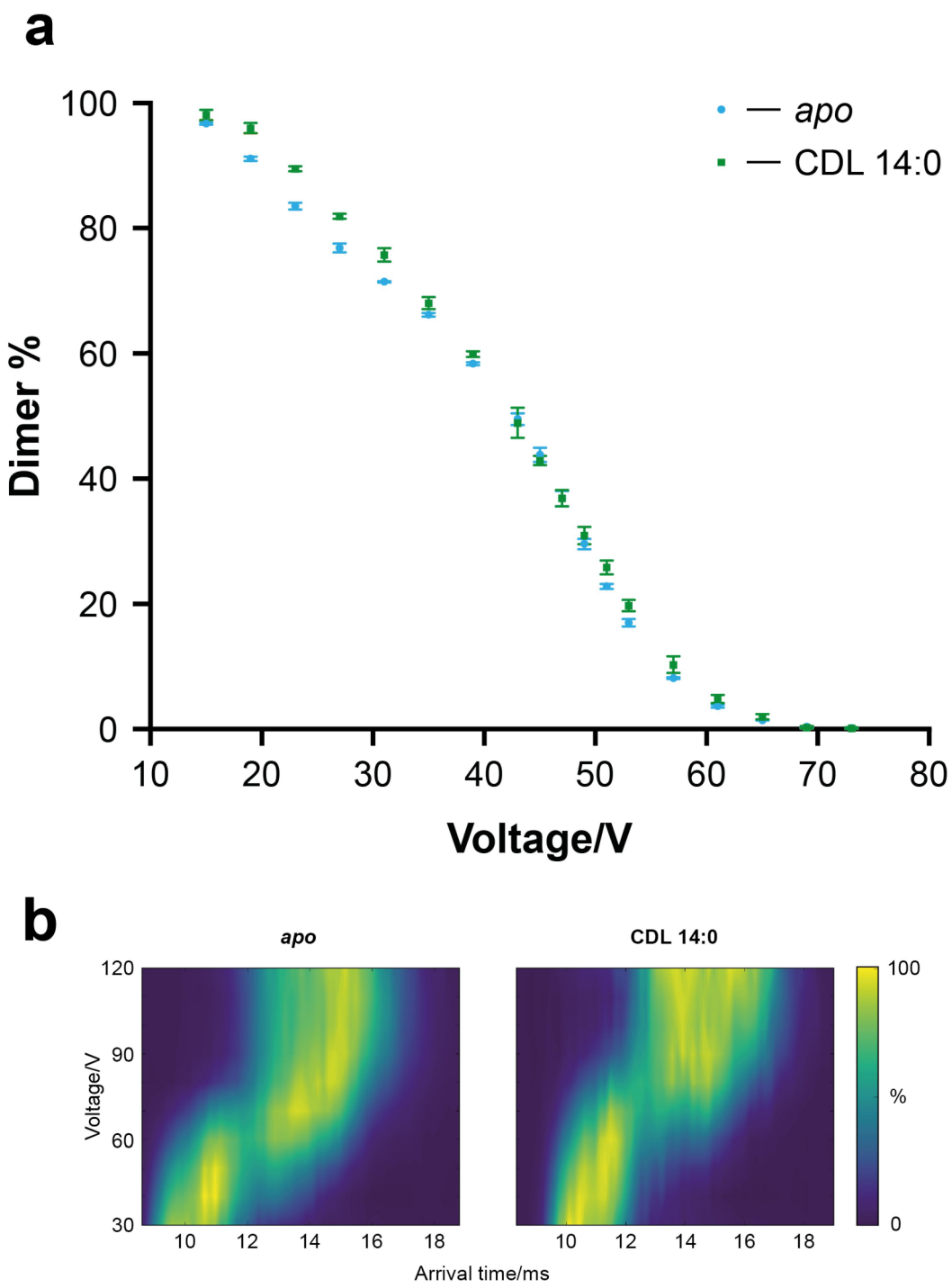


Figure 2.13 SID of SemiSWEET on the 'high-res' Synapt. a, ERMS plot of the 7+ SS dimer, either *apo* or bound to CDL 14:0. Values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation. **b,** Unfolding of the 7+ SS dimer due to increased IM Bias voltage in fly-through mode. A corresponding SID voltage (in SID mode) can be calculated by subtracting 20 V from the values on the y-axis. The voltage was changed in 10 V increments.

that the 'sigmoidal' shape of the CDL 14:0 curve is the expected outcome of an SID experiment, based on previous reports[20,22,25]; the *apo* SS deviated from it significantly in the first region, before returning to a more standard path later. This pattern is very reproducible, both within the three technical repeats used to produce **Figure 2.13a**, but also across two different versions of the 'high-res' instrument (with a linear IM cell and a cyclic IM cell), and another set of repeats on the 'cyclic' instrument at a different time period (data not shown). One plausible explanation is that the increase in the IM Bias voltage parameter (which increases the potential drop before the SID cell and is required to 'set' the SID energy) causes some activation of the *apo* SS dimer, effectively artificially enhancing the amount of dissociation observed. While such CIU events are undesirable and are minimised by reducing the gas pressure in the Trap region of the instrument, some degree of activation is often inevitable, especially for a protein which is already on the brink of unfolding due to desolvation requirements. This pre-activation is not observed for the CDL-bound species at the same energy due to induced stabilisation with respect to unfolding (**Figure 2.10**). However, this reasoning is contradicted by the previous observations, indicating that an unfolded protein would appear more, not less stable (as was explained in the previous paragraph). To settle this matter, I have obtained an IM Bias voltage increase unfolding plot on both precursors without performing the SID activation step (in fly-through mode), to see whether there would be any unfolding features matching the anomalous behaviour observed for the SS dimer (**Figure 2.13b**). While an unfolding event is indeed observed at a lower voltage for the *apo* protein than for the CDL-bound species, the energy at which it occurs does not match the unusual 'first region 'on the plot (*note: an IM bias voltage of 60 V corresponds to roughly 40 V in SID energy*).

However, it is not impossible that the opposite is true and the unfolding is instead responsible for the observed deviation from the 'first' curve to a seemingly more stable one. Regardless of whether or not this is actually the case, the fact that even during its 'second' stage the *apo* protein is observably less stable than the lipid-bound one clearly indicates that CDL is stabilising the dimer interface of SS and that SID has sufficient resolution to report this effect.

In order to show that the observed effect does not occur simply due to a presence of extra mass (and, consequently, a few extra degrees of freedom available for energy dissipation) I have selected two control lipids, POPE and POPG. However, there is a complication with the fact that, while CDL remains bound throughout an SID experiment, PE and PG start falling off the protein even at relatively low activation energies, as indicated by a high amount of *apo* dimer produced (**Figure 2.14a**). This fact significantly hinders direct comparison between CDL and these species. If the *apo* dimer is included in the calculation, PE- and PG-bound SS appear to be even more stabilised than the CDL-bound protein (**Figure 2.14b**) – this is unlikely to be realistic, considering that they do not even stay bound throughout the experiment. The increased stabilisation probably arises from the fact that the energy being used to promote the lipid departure does not contribute towards the SS dimer dissociation – similar effects have been observed by SID[20] and CIU[13] before. Alternatively, if the *apo* dimer is completely excluded from the calculation, PE and PG appear to be destabilising with respect to unbound SS – again, this is unrealistic, given that the produced *apo* dimer remains intact. One way around the problem is to consider only the lower-end of the SID energies, where less than 10% of PE and PG (**Figure 2.14d**) is dissociated and to include the *apo* dimer in the plot (**Figure 2.14e**). This way, the interface stabilisation

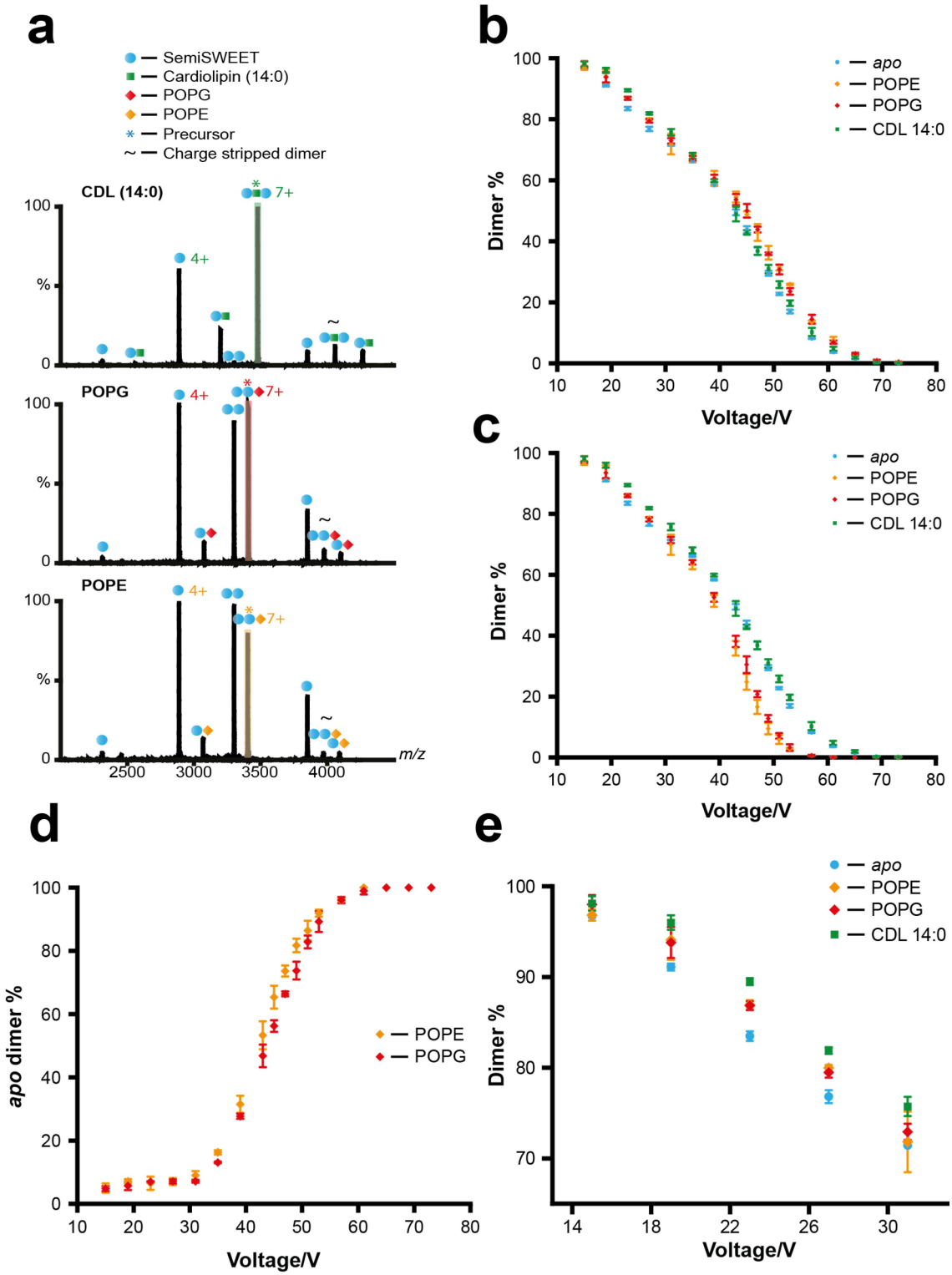


Figure 2.14 SID of SemiSWEET in the presence of various lipids. a, SID (43 V) of the highlighted precursor (7+ dimer bound to one lipid molecule). Only minimal amounts of *apo* dimer are produced from the CDL-bound dimer, but large scale falling off is observed for PE and PG. **b,** ERMS plot, *apo* dimer contributes towards dimer intensity. **c,** ERMS plot, *apo* dimer excluded from the plot. **d,** Percentage of *apo* dimer out of the total (*apo* and lipid-bound) dimer. **e,** ERMS plot, only the first five data points considered. For the panels **b, c, d** and **e** values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation.

induced by PG and PE is clearly lower than that induced by CDL, despite PG and PE curves being slightly enhanced by the 'falling-off effect'. This finding confirms that the dimer stabilisation by CDL is not an artefact due to a simple increase in mass. It should also be noted, that previous findings confirm this conclusion, with no stabilisation observed by SID due to binding of non-interfacial lipids[20].

Having observed the difference in SS-binding modes between the PL classes, the next step was to probe whether the effects of CDLs with different hydrophobic tails on the SS dimerisation can be observed by SID. The tails can vary in two main ways: the length (the number of carbons in each acyl chain) and the number of double bonds. In order to investigate these two parameters, I have selected the following species of CDL: 14:0, 14:1, 16:0 and 18:1. This set of lipids allows to probe the effect of the double bond for the lipids of the same length (14:0 and 14:1) as well as length for the lipids with the same number of double bonds (14:0 and 16:0; and 14:1 and 18:1) with a minimal number of experiments required. As mentioned above, CDL 18:1 is expected to have the most influence on the dimerisation of SS by previous data[30] (**Figure 2.3d**, see the 72C case). It would be interesting to also have a longer species than CDL 18:1 in the set, but unfortunately they are not commercially available. Considering the distribution of native lipids in **Figure 2.3d**, CDL 18:1 appears to be close to the limit of size for naturally occurring CDL, which is probably the reason why it is not mass produced.

Figure 2.15a shows the SID-ERMS plot of the 7+ charge state of SS bound to each of the lipids in 'the set', as well as the *apo* dimer. From this graph, the *apo* SS can be seen to be less stable than the lipid-bound species, as was already described above for the individual case of CDL 14:0 (**Figure 2.13a**). Also, CDL 18:1-bound dimer visually appears to be slightly more stable than the other

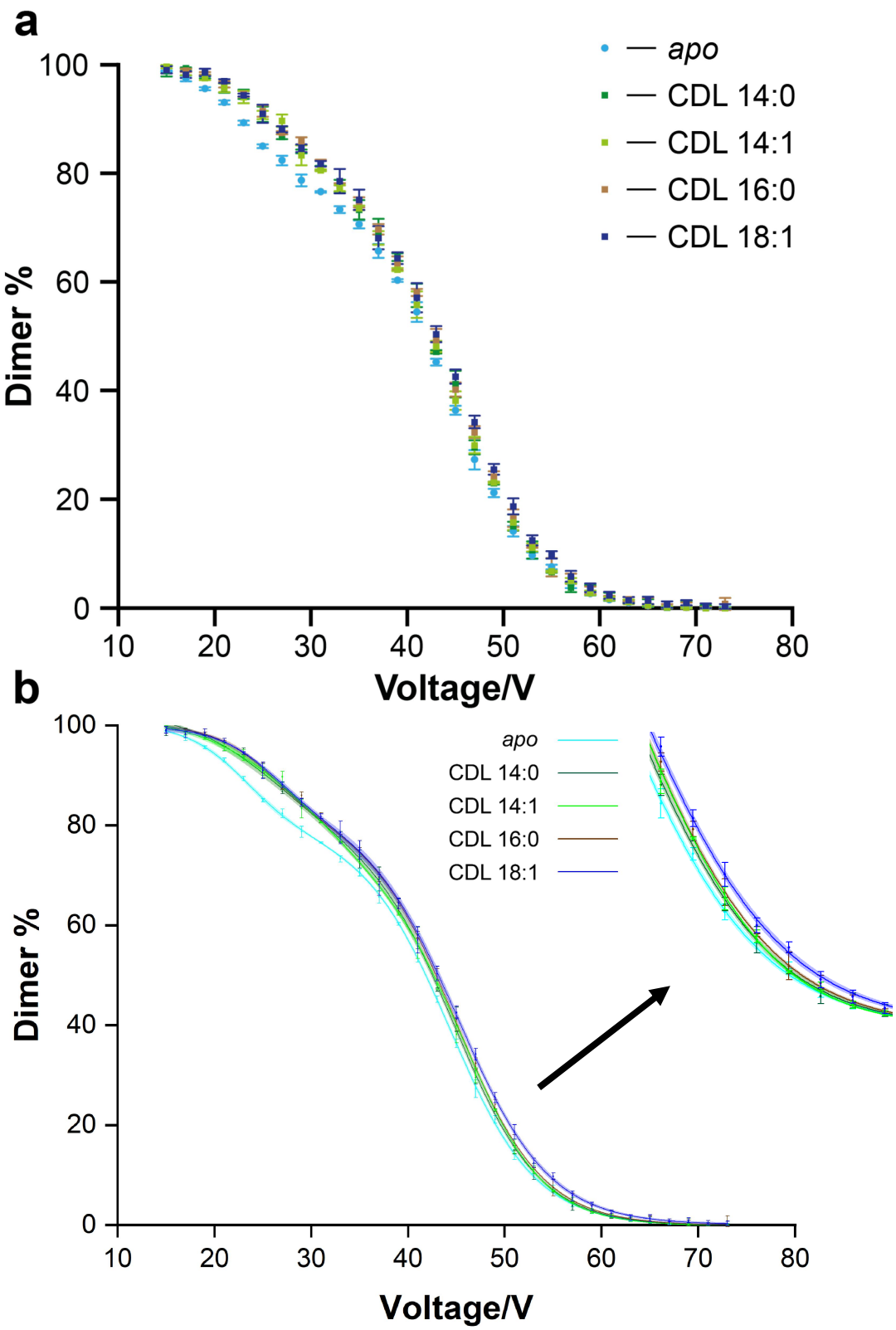


Figure 2.15 SID of SemiSWEET in the presence of cardiolipin. a, SID-ERMS plot of SS bound to various CDL species. **b,** The same plot as above, but fitted with a sigmoidal curve. Inset: a zoomed in region of the curve around the 50 V mark. Values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation.

lipidated species, especially in the region roughly between 40-60 V. However, it is extremely difficult to come to any certain conclusions from these data, because unlike some of the previously reported cases[20], these differences are not clear cut. Currently, there is no great way to extract information from the SID-ERMS plots, apart from qualitative comparison. When a single energy value is required, an arbitrary threshold of 10% dissociation has been used as appearance energy[21,41] (alternatively an inflection point of the curve[42] or the 50% dissociation point[43] can be used as the energy value). While this method works well for large deviations in appearance energies between different proteins, it is not applicable to the data here as the main differences occur during the 'steep' part of the sigmoidal curve. Here, I propose a new way of data analysis, designed specifically to identify very subtle differences between the species.

The first step towards better data interpretation includes fitting a mathematical curve to the individual data points. As the plots show roughly sigmoidal behaviour, I have selected one of sigmoidal functions in the Origin Pro software, choosing the one that produced the best overall fit (**Figure 2.15b**) (see **Materials and Methods** for more detail). This picture allows for the better identification of the voltage region, where significant differences can be expected to be observed, namely between 43 V and 55 V (inclusive). This is extremely beneficial for a statistical analysis step, as I will discuss below. I would like to stress that, while this particular sigmoidal function is not necessarily the best one to choose and may not be relevant to the SID process in terms of physics, the only role of this fitting is to identify the area on the graph where the largest changes can be observed. The following statistical analysis is performed directly on the data and does not involve the mathematical fitting in any way, apart from selecting a particular subset of data to focus on.

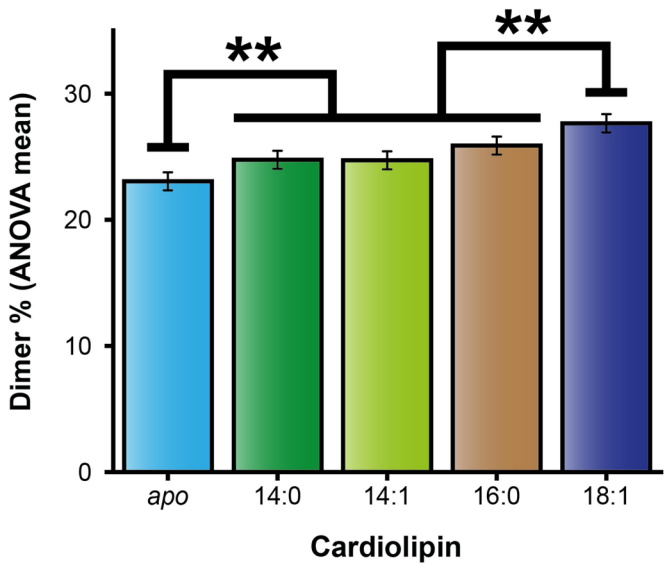
Since the data points have overlapping error-bars, a statistical analysis is necessary to determine whether the observed differences are significant or not. Since there is a categorical dependent variable (the type of lipid or the SID voltage used) and a continuous independent variable (the percentage of intact dimer remaining) analysis of variance (ANOVA) test is a natural choice for these data. As ANOVA is only used here for practical implementation, the full mathematical ramifications of this method are beyond the scope of this work. Therefore, I will only give a brief outline of this method sufficient for practical use (in this section and in slightly more detail in **Materials and Methods**); a more comprehensive description can be found in specialised literature[44]. ANOVA is a statistical technique used to test whether the population means for 3 or more categorical groups are all identical. The null hypothesis is that the means are all identical – the probability of the null hypothesis below a certain value (usually 0.05) implies that at least one of the population means is sufficiently different from the others. ANOVA by itself only shows whether or not all of the means are equal – but it does not indicate where the differences lie exactly. If the means are shown to be statistically different by ANOVA, a post hoc test needs to be run to identify which means in particular are different.

If there is only one type of independent variable (e.g. different lipids interacting with the same protein), one way ANOVA is used. For the SID curves here, one way ANOVA can be applied by treating each voltage separately; then the lipid identity is used as the only dependent variable. One way ANOVA revealed statistically significant differences at multiple different SID voltages, so a post-hoc Tukey test was employed to identify the pairwise differences of the means (shown for the 43V-55V 'subset' of voltages in **Figure 2.16a**). These data are inconsistent from voltage

a

Voltage/V	43	45	47	49	51	53	55
	16:0>*apo*	14:0>*apo*	16:0>*apo*	14:1>*apo*	18:1>*apo*	18:1>*apo*	18:1>14:0
	18:1>*apo*	18:1>*apo*	18:1>*apo*	16:0>*apo*	18:1>14:0		18:1>14:1
		18:1>14:1	18:1>14:0	18:1>*apo*			18:1>16:0
			18:1>14:1	18:1>14:0			
				18:1>14:1			

b



c

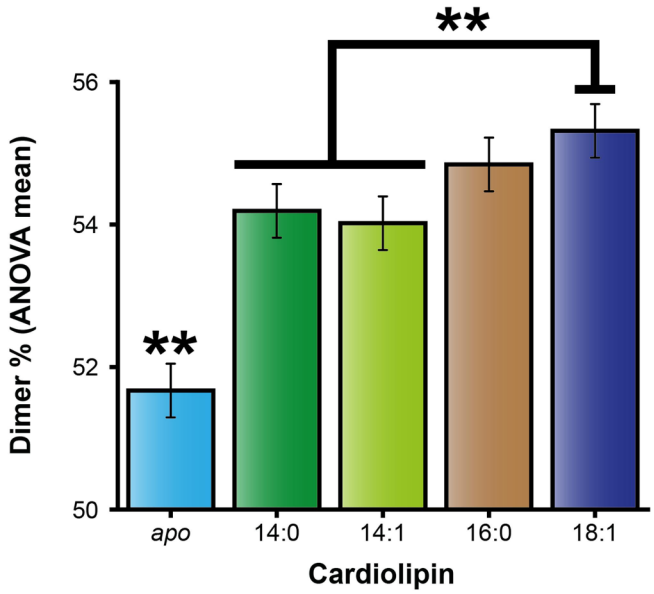


Figure 2.16 Statistical analysis of the SID-ERMS data of CDL-bound SS. a, One-way ANOVA test results. A>B should be read as dimer A is more stable than dimer B, according to the Tukey post-hoc analysis at 0.05 confidence level. The values in red represent voltages for which the homogeneity of variances test was not passed. **b,** Two-way mixed ANOVA performed on the data in the range from 43 V to 55 V. **c,** Two-way mixed ANOVA performed on the data in the range from 15 V to 63 V. For panels **b** and **c** double star (**) implies statistical significance (Bonferroni test, 0.05 confidence level).

to voltage regarding which differences are statistically significant, and which are not. In addition, the data points at 49 V and 55 V did not pass the homogeneity of variances test (see **Materials and Methods**), so technically the Tukey test is not appropriate to use there (although there is no reason to assume that the data at these voltages are fundamentally different to the other ones). The outcome of the one way ANOVA indicates that the little variations for each data point in either mean or standard deviation have a major effect on this test. A better statistical analysis would look at a region of several subsequent voltage values as a whole.

A two-way mixed ANOVA was then applied to the data, with lipids used as non-repeating variable and voltages as the repeating variable (in practice, it means that each repeat run for each lipid bound is expected to be completely independent from the others, whereas the voltages that are part of the same run are treated as not entirely independent, as they might share some common biases due to a state of the instrument at a particular time. This setup allows to ignore such systematic variations in a particular data set, and instead looks at the change from one energy to another). Also, including the SID voltage as a second independent variable enables the ANOVA test to directly compare the data for the different lipid species at the same energy, as desired (for example, comparing the amount of the *apo* SS at 47 V, but at 55 V for the CDL 14:0-bound dimer is not informative of any fundamental differences due to the presence of the lipid). The outcome of this ANOVA experiment (with a post hoc Bonferoni test) is presented in **Figure 2.16b**. It is important to note that the bars do not represent the means of the data, but the 'ANOVA means' with a corresponding error. If these means are sufficiently different (meaning that those error bars are not overlapping), it implies that the data in the two categories are statistically significantly different (at the 0.05 confidence level).

This graph shows that in the voltage region between 43 V and 55 V all the CDL species are significantly stabilising compared to the *apo* SS dimer and that, excitingly, CDL 18:1 is significantly more stabilising than the other three lipids. The means of CDLs 14:0, 14:1 and 16:0 are not significantly different from each other. This fact further validates the proposed relevance of these results: if the differences observed were simply an effect of the increased mass (either directly due to SID or indirectly due to some effect on unfolding observed in **Figure 2.13b**), the change between 14:X CDLs to 16:0 compared to that between 16:0 to 18:1 would be expected to be either the same (as the absolute value of mass increase is the same in both cases) or lower for the latter (because the relative increase in mass is lower percentage-wise). Instead, the increase in stability caused by CDL 18:1 is the greatest, in accordance with prediction outlined above based on previously reported data. It is also important to note that, while it might be tempting to speculate that the length of the chain plays an important role for the SS dimer stability, whereas the double bond is not relevant based on the fact that the ANOVA mean of CDL 16:0 is slightly larger than the virtually identical means of the two 14-carbon species, the fact that these differences were not identified as statistically significant implies that this observation might have simply occurred by chance.

Figure 2.16c provides some evidence on why doing the mathematical fit prior to the ANOVA step is so important. When the two way mixed ANOVA is run on the more complete data set instead (15 V to 63 V), its findings are very similar to the above, with the exception of CDL 16:0, which is now not significantly higher than the two 14:Xs, nor significantly lower than CDL 18:1. This observation can be explained by the presence of data points at which the differences between the lipids are minimal or virtually non-existent; so when such a 'diluted' data set is

analysed with ANOVA, the test can no longer confidently confirm that the overall means are in fact different. Therefore, limiting the analysis exclusively to a region of interest is beneficial. In addition, using a high sampling interval during the experiment (2 V increments of SID voltage) results in more data points being available for statistical analysis in the region of interest.

As was discussed above, the increase in the number of degrees of freedom (DOF) upon lipid binding is unlikely to be the main cause of the observed differences in dimer stabilities between the different CDL species. It is also possible to account for the different number of DOF between the complexes in a more quantitative way. A simple linear DOF correction can be applied, that was previously successfully employed to investigate SID fragmentation efficiencies of peptides with varying numbers of atoms[42], by using the equation:

$$E_{DOF} = \frac{DOF_{apo}}{DOF_{com}} Vz, \tag{2.2}$$

where E_{DOF} is the DOF corrected dissociation energy, DOF_{apo} is the number of DOF for the *apo* protein, DOF_{com} is the number of DOF for the protein-lipid complex, V is the SID voltage and z is the protein charge. After applying this correction on the data, the ANOVA test can be run to investigate whether the differences observed in **Figure 2.16b** remain statistically significant when taking the number of DOF into account. Since the vibrational modes are the most efficient in distributing energy, only the vibrational DOF are usually taken into account, which can be calculated from the total number of atoms, n, by the equation $DOF_{vib}=3n-6$ [45] (alternatively, the 3 rotational degrees of freedom can also be included; however, since n is very large for the complexes investigated here,

increasing or decreasing the values by 3 will have no observable difference on the outcome of the calculation).

There are two important points that need to be taken into consideration regarding equation (2.2). Firstly, while the total number of DOF provides a good indication of the energy distribution during SID fragmentation of small peptides, the situation is more complicated for a comparatively large protein complex: for example, as was mentioned earlier in this chapter, the conformation of a protein can have a large impact on the observed dissociation energy and, therefore, on the internal energy distribution, despite no change in the total number of DOF[24,40]. Consequently, the location of the lipid binding site and any allosterically induced structural changes to the protein complex upon binding would need to be taken into account in order to completely understand the effect of the additional DOF on the dissociation energy. However, assuming that all of the CDL species bind roughly in the same location, the linear DOF correction can be used to compare the dissociation energies between them.

Secondly, the DOF correction described by equation (2.2) applies to dissociation energy, while the outcome of the data analysis presented in **Figure 2.16b** is the fraction of the undissociated dimer at a given energy. Therefore, the data need to be processed differently and the dissociation energies need to be extracted before the DOF correction can be applied. As was discussed earlier, one way to achieve this is by taking the threshold energy required to dissociate a certain fraction of the precursor. Given that the two way ANOVA applied to several values produced a more interpretable outcome than the one way ANOVA applied to individual data points for the original data analysis (**Figure 2.16 a,b**), I have decided to also use several threshold values for this analysis. I have selected the dissociation

percentages corresponding to the 43V-55V region used previously (**Figure 2.15b**), so the voltages required to dissociate 50% of the precursor, 55% of the precursor… – all the way down to 90% of the precursor in 5V steps were chosen (note that 90% dissociation of the precursor is equivalent to 10% intact precursor remaining). In order to extract these values, a similar sigmoidal curve to that shown in **Figure 2.15b** was fitted to each of the individual (not averaged) data sets; and the extracted voltage values for each lipid species were analysed by the two way mixed ANOVA (**Figure 2.17a**) (similarly to the data in **Figure 2.16b**). However, by this method, the only significantly different pairs were CDL 16:0 and *apo* SS, and CDL 18:1 and *apo* SS (**Figure 2.17b**). Unfortunately, no significant differences in voltages between the CDL species were observed even without the DOF correction, meaning that the proposed experiment of applying the correction and re-running the ANOVA test to see if the statistically significant differences remain cannot be carried out.

While it might initially seem surprising that the same statistically significant differences are not observed by the ANOVA analysis shown in **Figure 2.16b** (comparing the percentage of dissociated precursor at a given voltage between the different CDL-SS complexes) and **Figure 2.17b** (comparing the voltages required to achieve a certain percentage of dissociation of the precursor between the different CDL-SS complexes) as they are based on the same raw but differently processed data, this outcome can be explained by the nature of the experimental data acquisition. The voltage is set experimentally, meaning that the ANOVA analysis on the intact precursor percentage (**Figure 2.16b**) can be run on the actual data; whereas fitting and interpolation is required in order to extract the threshold dissociation voltages (**Figure 2.17a**), increasing the uncertainty.

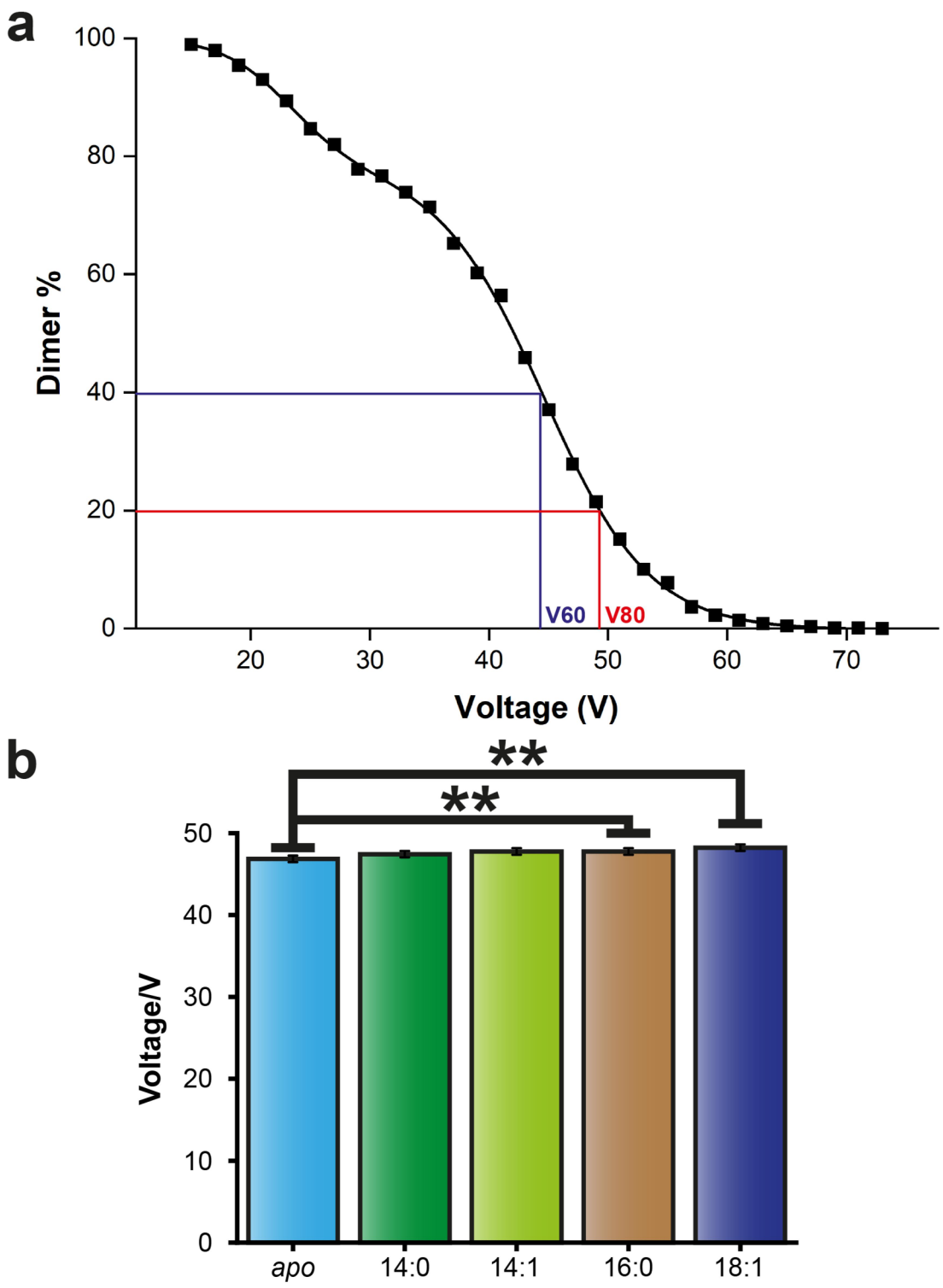


Figure 2.17 Statistical analysis of the SID threshold energies of CDL-bound SS. a,. SID-ERMS plot of *apo* SS (individual run). A sigmoidal curve is used to extract threshold dissociation voltages, highlighted for 60% precursor dissociation voltage (V60, blue) and 80% precursor dissociation voltage (V80, red). **b,** Two-way mixed ANOVA performed on the threshold voltages between 50% and 90% precursor dissociation (5V steps). For panel **b** double star (**) implies statistical significance (Bonferroni test, 0.05 confidence level).

Although it is not possible to apply the actual correction described by equation (2.2) to the SS dissociation voltages, the DOF correction factor relative to the *apo* protein (DOF_{apo}/DOF_{com}) can still be calculated for the SS-CDL complexes (**Table 2.1**).

	SS	SS+CDL	SS+CDL	SS+CDL	SS+CDL
DOF	9987	10617	10593	10689	10737
DOF_{apo}/DOF_{com}	1.000	0.941	0.943	0.934	0.930

Table 2.1 Degrees of freedom correction factors.

These correction factors have very close values for the four SS-CDL complexes and, therefore, are unlikely to account for a particularly large stabilisation of the dimer by CDL 18:1.

2.7 Molecular Dynamics Simulations of SemiSWEET

Applying several biological methods can provide more insight into a biological system than any of those methods could do on their own. In accordance with this principle, I have decided to use the assistance of MD simulations in order to get more detail about the nature of SS-CDL binding (*all of the MD experiments were performed by, and the data analysed by, Daniel Quetschlich*). Before MD simulations can be performed, a starting structure of SS needs to be selected. In a native membrane, SS can interconvert between three different conformations: inward-open, outward-open and occluded (closed on both sides) (**Figure 2.18a**)[28,29]. Importantly, the conformational changes are not induced by the substrate and can occur freely[28,46]. Therefore, it is difficult to predict which structure SS would have in the gas phase, or even in solution. I have recorded IM spectra on a drift-tube (DT) instrument (described in **Chapter 1**, **Figure 1.6c**) in order to measure the CCS value. The CCS can then be compared to a theoretical value calculated from the crystal structure of each conformation, providing an estimate

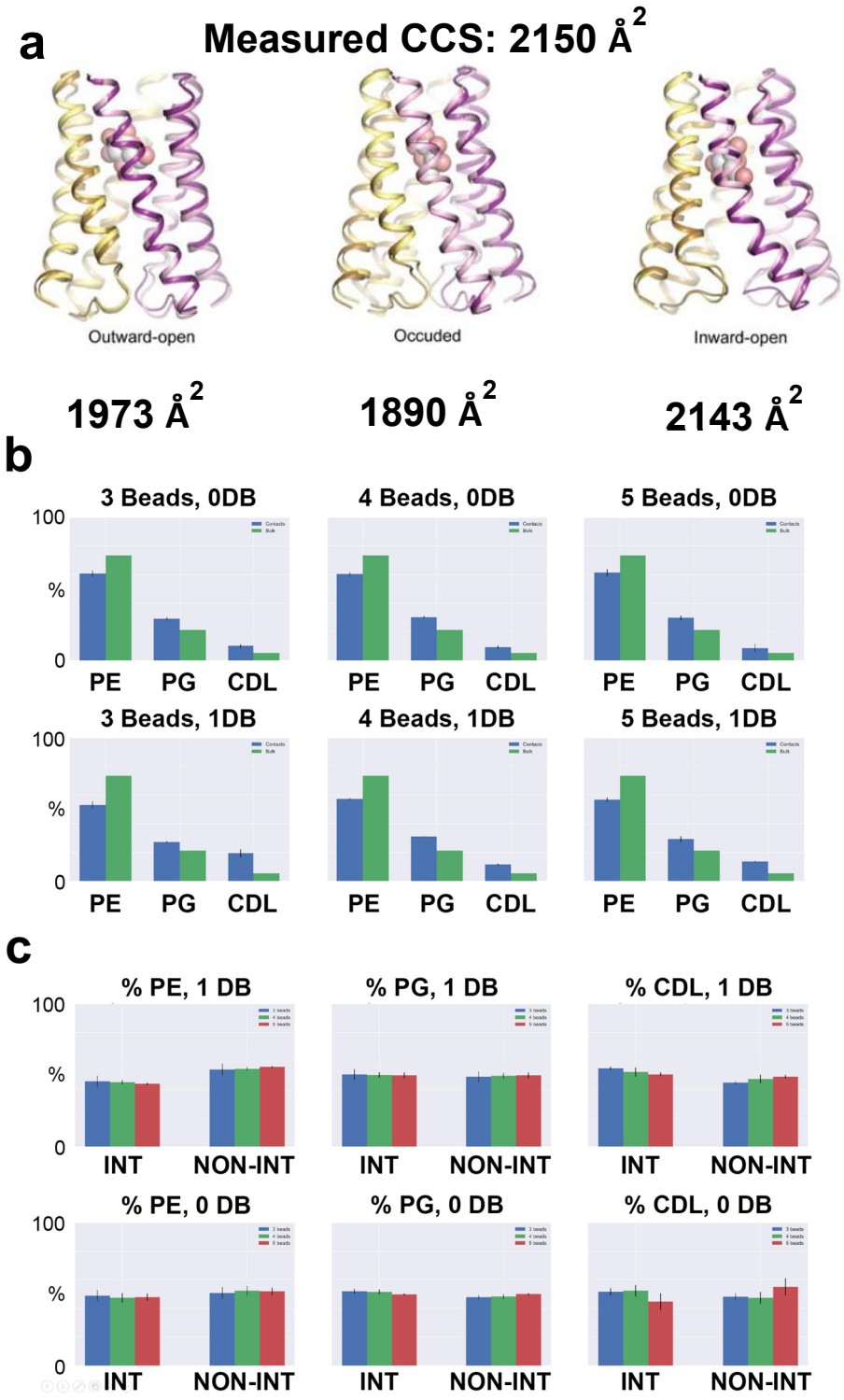


Figure 2.18 Molecular dynamics simulations of CDL-bound SS. a, Experimental and theoretical CCS values of SS dimer. Figure adapted from ref(46). **b,** MD simulation of lipids in contact with SS (blue) or in bulk (green). DB – double bonds. Number of beads represents chain length of CDL (3=14C, 4=16C, 5=18C) **c,** Interfacial (INT) and non-interfacial (NON-INT) contacts for different lipids. 3 beads (blue), 4 beads (green) and 5 beads (red) and the number of DBs refer to the CDL species present in the simulation, and not to PE or PG. Panels **b** and **c** adapted from the figure made by Daniel Quetschlich.

of the gas-phase conformation of SS. It should be noted, that it is highly possible that none of the crystal structures actually exist outside the bilayer – so the conformation observed in the gas phase might represent some combination of the three arrangements in **Figure 2.18a**, or even a completely separate conformation. The CCS analysis is only used to obtain an estimate for a reasonable starting point – more advanced computational techniques, such as structure relaxation[46], would be more suitable for advanced and comprehensive MD simulations which are beyond the scope of current work.

A further complication arises from the fact that, while crystal structures for all of the three conformations are available, they were not simultaneously shown for a single SS homologue (for example, for Vs SS used in this work only the structure in the outward-open conformation has been obtained)[29]. However, it has been shown for SS homologues that the structures they form are very closely aligned, despite modest sequence homologies[28,29]. Therefore, I have used all of the currently known crystal structures to roughly estimate the expected CCS value for each conformation of SS (see **Materials and Methods** for a detailed description). DT IM data was acquired using gentle activating conditions, designed to preserve the most-compact conformation of SS. The experimental CCS value was determined to match the inward-open conformation, so it was selected for the MD simulations.

MD simulations were performed using a coarse-grained approach. This means that the CDLs with different tails could not be reconstructed exactly, but instead were represented by a different number of beads: 3 for 14-C chains, 4 – for 16-C and 5 – for 18-C. In total, 6 different CDL species were used, one for each combination of a particular chain length and either zero or one double bond. Two modes of analysis were employed.

First, the number of lipids in the simulated bilayer which were in direct contact with the protein and in 'bulk' was determined (**Figure 2.18b**). Here, any direct contact is counted, not just the interfacial region. It can be seen that SS prefers binding to the negatively charged PG and CDL lipids than to zwitterionic PE, despite the prevalence of the latter in the simulated bilayer. Interestingly, the native SID data revealed a greater amount of PE falling off compared to PG, seemingly reflecting the preference observed by the simulation (**Figure 2.14d**).

Second, the contacts for the different CDL species were separated into interfacial and non-interfacial (**Figure 2.18c**). Unfortunately, no significant differences were observed as the small variations were well within the experimental error. This could imply that the resolution of the coarse-grained simulation is insufficient for this investigation; computationally expensive atomistic simulations might be more appropriate in this case.

2.8 Potential Issues for SID-ERMS of Membrane Proteins

I would like to conclude this chapter with a brief description of potential pitfalls that could be encountered while performing the SID-ERMS experiments. Here, I focus on the subtle effects that are not obvious, but that can significantly affect data interpretation and potentially lead to erroneous conclusions.

As was mentioned in **Section 2.6**, the biggest problem of this type is an accidental quadrupole co-selection of an undesired species together with the peak of interest. There are three common sources of this error:

1. Incomplete desolvation of the MP prior to selection. The remaining detergent molecules bound to various protein-lipid complexes present in the spectrum can

cause an overlap in the *m/z* space that is extremely difficult to predict. For example, when CDL 14:0 is selected at 130 V cone energy and then gently activated in the trap a pattern of several peaks is produced instead of the anticipated single peak (**Figure 2.19a**). This arises due to the fact that a part of the intensity does not belong to the 7+ SS dimer bound to one CDL 14:0 lipid, but instead represents the 7+ SS dimer bound to 4 copies of C8E4 detergent (a difference of 1-2 units in *m/z* space, which is too small to practically separate using the quadrupole). If such an ion was activated by SID, a large amount of *apo* dimer would be quickly produced due to the relatively low energies required to strip off the detergent. More importantly, calculating the amount of the undissociated lipid-bound dimer becomes impossible, because the peak will contain a varying amount of detergent-bound dimer at each SID voltage. Considering the fact that the space of potential protein-lipid-detergent complexes is vast, an overlap between some species is almost inevitable. As I have mentioned in **Section 2.6**, a practical solution is to increase the cone energy to promote complete desolvation, even at the expense of causing a minor amount of unfolding.

2. Incomplete detergent exchange. The effect is similar to that described above, with a caveat that it often cannot be resolved by a simple increase of the in-source activation. This is because the exchange usually occurs from a very ‘sticky’ and mild detergent, such as *n*-dodecyl-β-d-maltopyranoside (DDM), into an MS-compatible harsher detergent, such as C8E4. The implication is that the energy required to dissociate any remaining DDM molecules from the protein is much higher than for C8E4, resulting in a large degree of protein unfolding, or even complex dissociation, occurring before the complete detergent removal. In the case of 7+ SS dimer bound to CDL 18:1, it happened to overlap with a 4+ SS monomer

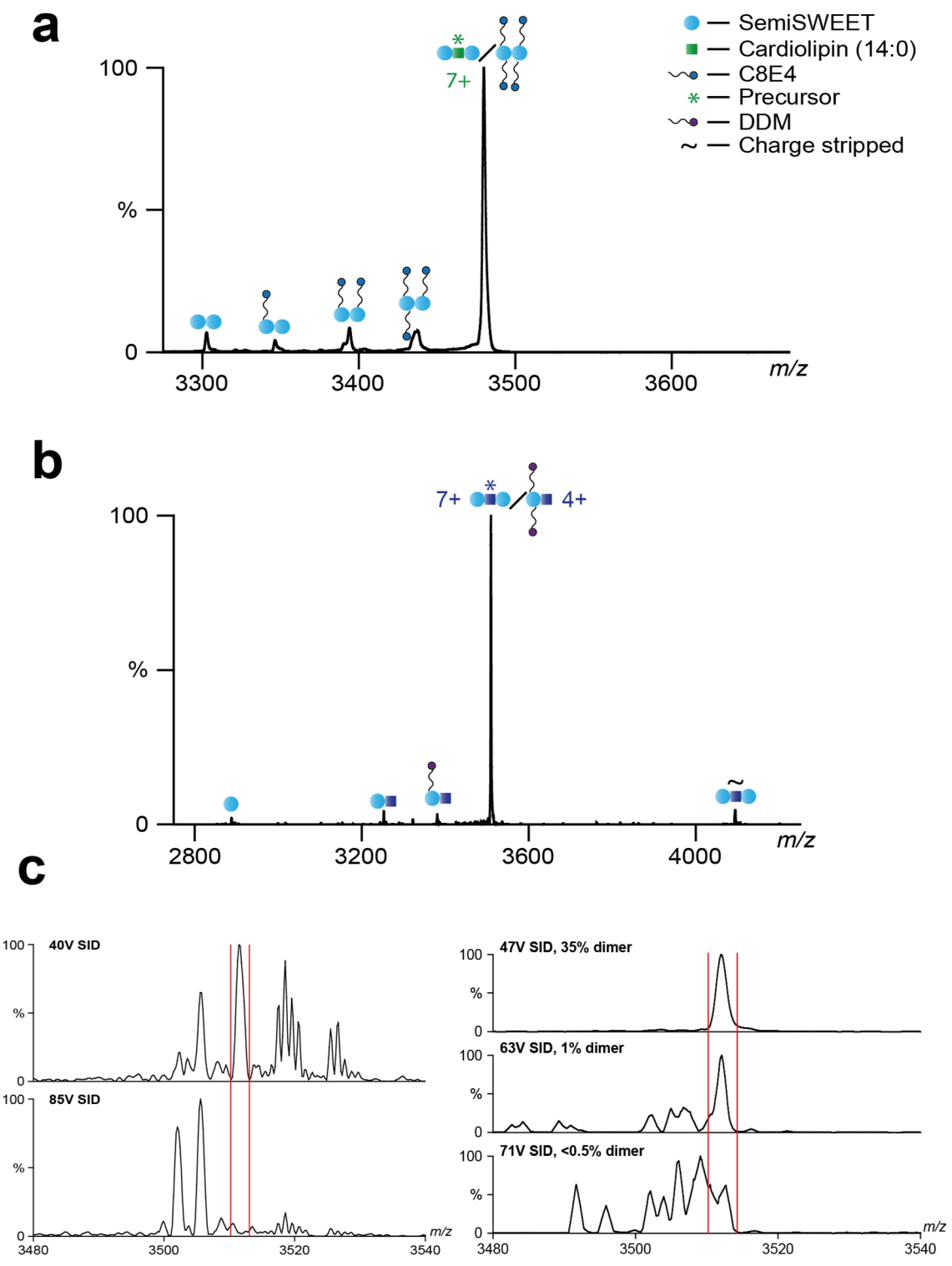


Figure 2.19 Effect of impurities on SS spectra. a, Selection of SS-CDL 14:0 at 130 V cone overlaps with SS+4xC8E4 as revealed by gentle trap activation. **b,**. Selection of SS-CDL 18:1 overlaps with SS monomer+CDL+2xDDM, revealed at 17V SID. **c,** A contaminant present on SS-CDL 18:1 selection as seen on QE (left) and Synapt (Right). Red lines roughly indicate the region used for the peak area calculation. Data for the left panel of **c** provided by Joseph Gault.

bound to one molecule of CDL 18:1 and two molecules of DDM. Here, it is important to note that, unlike the case above, where the presence of an unwanted co-selection was relatively obvious due to the formation of unexpected dissociation products, in this case a lipid-bound 4+ SS monomer is expected to form from the 7+ dimer. Therefore, the only unusual species is the intermediate 4+ SS monomer bound to CDL 18:1 and one DDM – however, it only occurs at a few low SID voltages and could be miscounted as noise unless careful attention is paid (**Figure 2.19b**). If this co-selection would go unnoticed, the amount of monomer produced by SID from the dimer would be enhanced by the selected monomer, stripped from its detergent. As a result, CDL 18:1-bound SS dimer would appear much less stabilised than it is in reality, potentially completely reversing the final conclusion. The only way around this problem is by improving purification conditions, to achieve complete detergent exchange. For example, this can be done by running two rounds of size-exclusion chromatography (SEC) detergent exchange, instead of just one; or, as was done in this case, by discarding a part of the protein yield by not collecting any shoulder peaks on a SEC chromatogram.

3. Presence of impurities. Unlike the two cases above, this is a problem that is much less likely to occur. It arises when an impurity remains in the sample that happens to be really close in mass to the species of interest. Due to the extensive purification protocol used to prepare the protein samples for native MS, a number and intensity of any impurities in the spectrum is small, making a chance of observing an overlap very low. Nevertheless, a co-selection of an impurity was detected for the 7+ SS-CDL 18:1 dimer – this is especially clear on Q-Exactive SID data (*obtained by Joseph Gault*) (**Figure 2.19c, left**). It is possible that more than one impurity is present, as the ratio of the peaks not assigned to be SS changes

at different collision energies. Importantly, there is a significant difference in *m/z* between the impurities and the sample, so the presence of the contaminants can be minimised by employing a tight quadrupole selection window, as is shown for the 'high-res' Synapt data (**Figure 2.19c, right**). From these data, it is clear that the contaminant is only significantly present at very high SID voltages, where the overall amount of dimer is at or below 1%. For the vast majority of data point, the impurity is extremely low in intensity and is also not included in the amount of dimer calculation due to a sufficiently different mass. Therefore, in this particular case the presence of this contaminant did not affect the overall results or conclusions. However, if the difference in *m/z* between the sample and the contaminant was low enough to prevent base-line resolution, the only way around the problem would be to completely eliminate the contaminant during purification stage. This can potentially result in a tedious and time-consuming purification protocol optimisation being required.

Another potential source of peak overlap arises not from an undesired co-selection, but as a direct consequence of SID. During a surface collision, some amount of charge stripping can occur, resulting in a precursor with less charge – in some cases, it can then overlap with one of the monomeric dissociation products[25,40]. This is especially true if the precursor is dimeric (such as SS), as either the selected or the charge-stripped precursor must be divisible by two and, therefore, occupies the same *m/z* space as a monomer with half the charge. Fortunately, the dimers and the monomers are clearly separated in the IM space (**Figure 2.20a**). As a side note, this figure shows that the dimeric SS precursor shows extended conformation (as was also observed for AmtB in **Section 2.3**). However, the charge-stripped dimer is visibly less unfolded, even though it is formed directly from the precursor.

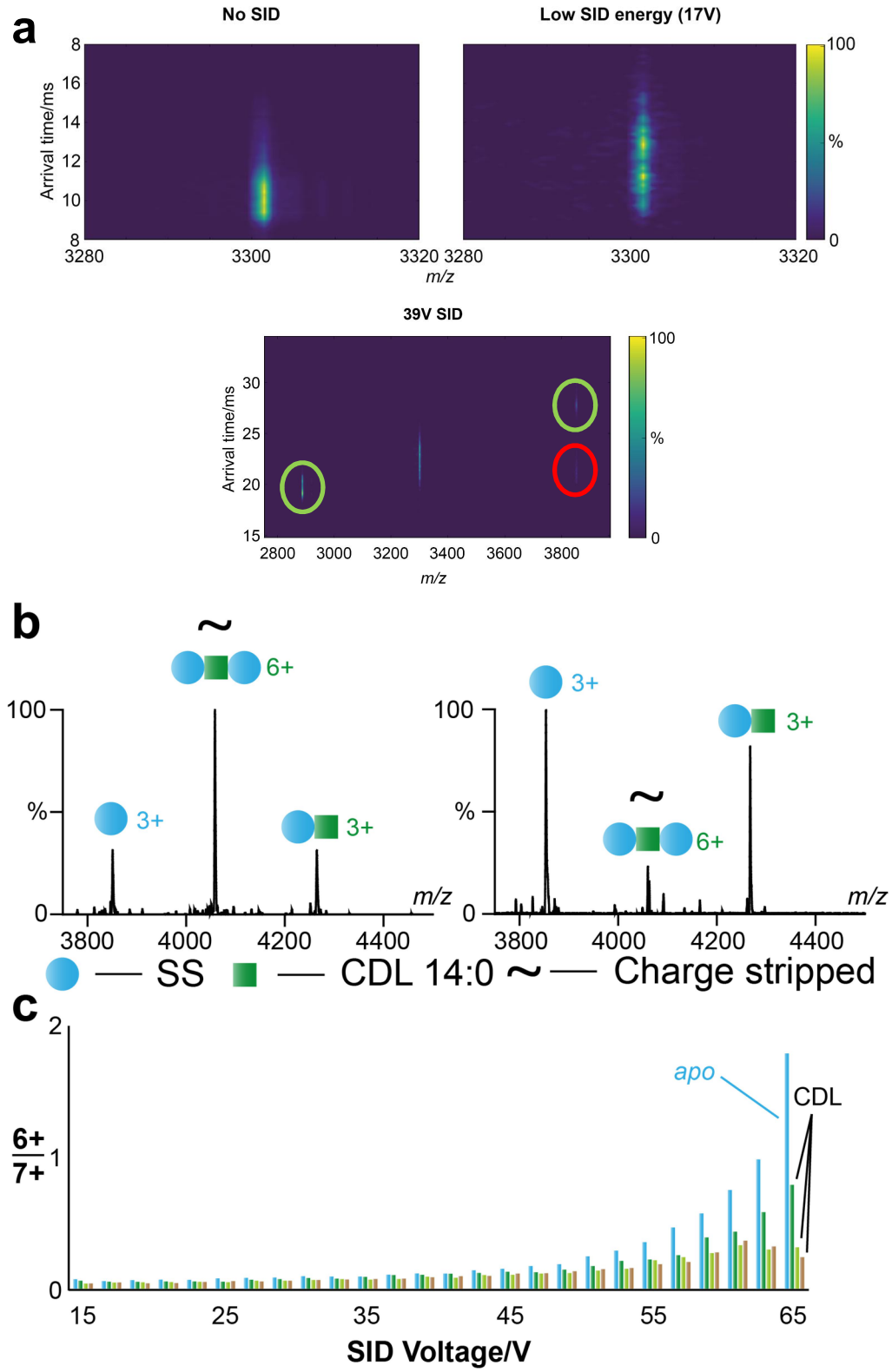


Figure 2.20 Charge stripping considerations. a, IM data for SS 7+ dimer. Lower panel: green – monomer, red – charge stripped dimer. **b,** The amount of charge stripped dimer (relative to the 3+ monomer) produced when using an old (left) or a fresh (right) surface at SID voltage of 31V. **c,** 6+/7+ ratios are similar for lipidated and non-lipidated (calculated using IM data) SS; the total amount of dimer decreases with voltage, so the absolute difference remains low, despite the increasing relative difference. Values shown are the average of three repeats performed from different needles; standard deviations were low and are not shown.

This is similar to the formation of folded dissociation products from a seemingly unfolded precursor, discussed for AmtB earlier. Two potential explanations exist for this behaviour.

The first explanation is that the extended conformation of the precursor is an alternative form of energy dissipation to dissociation or neutralisation. This means that if the precursor dissociates after the surface collision, then that is where its internal energy goes, so the 'fragments' remain folded; while if the dissociation does not occur, the internal energy has to go into vibrational modes, promoting unfolding. The data obtained on the charge reduced dimer (**Figure 2.20a**) supports this conclusion, because the energy is deposited into the surface to promote charge reduction – leaving less energy to cause protein unfolding. This reasoning explains the puzzling formation of folded dissociation products from a seemingly unfolded precursor[25] – as the precursor does not in fact unfold, unless it also does not dissociate.

The second explanation relies on the fact that the charge-stripped intact oligomer and the dissociation products have less charge than the original precursor. The amount of charge is an important determinant of the protein conformation in the gas phase, and a significant loss of charge can cause even solution-denatured proteins to appear indistinguishable from the native ones by CCS[47] (although such an extreme scenario is not observed during an SID experiment). Therefore, by this explanation the precursor unfolding happens first, while the lower charge products have sufficient time to collapse into more folded conformations on a ms time scale of an IM experiment. Personally, I prefer the first explanation, because the second explanation fails to account for the SID products looking very different to those obtained by CID[34].

Returning to the issue of overlap between the charge-stripped precursor and the associated problems, it could be more pressing when performed on a non-IM compatible instrument, such as FT-ICR or Orbitrap based mass spectrometers. It was demonstrated that the high resolution provided by these machines can be exploited to separate such overlapping species by determining the charge from the isotopic pattern[48,49]. Unfortunately, resolution has both practical and theoretical limits[50], so such a solution is not always available. An alternative is to only analyse the 'distinct' peaks, for which no overlap is theoretically possible[51] (such as the odd charge states of SS dimer, for example). This method has a disadvantage of discarding good data, potentially including the most intense peaks in a spectrum. Here, I present two alternative solutions to this problem.

1. Use freshly prepared surfaces. To investigate whether the 'age' of the surface has any effect on the induced charge-stripping, data was collected on the same instrument, using a 'several-months-in-use' surface (**Figure 2.20b, left**) or a 'freshly' re-coated surface (**Figure 2.20b, right**). The intensity of the charge-stripped peak is significantly higher for an old surface than it is for a fresh one. While data is shown for a single point, the same pattern was observed for all experiments performed in those two sessions. Therefore, while using a 'new' surface does not completely eliminate charge stripping, it can significantly minimise its effect on data interpretation, especially in the cases where non-resolvable peak overlaps are likely. After continuous use, a surface can be covered with stuck ions or potentially even have defects where patches of 'naked' metal are exposed (although the latter is less likely), which can facilitate charge transfer by either proton or electron transfer (a mass difference of 1 Da for these two cases cannot be confidently identified for SS even on a 'high-res' Synapt, due to a relatively low

intensity of the charge stripped peak; previous soft landing experiments performed on peptides identified protons as charge carriers in that case[52]). It should be noted that this method applies to the 'fluorinated-SAM on gold' surfaces only; the extent of charge stripping of native-like proteins for the recently introduced stainless steel surfaces[53,54] requires further investigation.

2. Estimate charge stripping by using a singly-bound precursor. As can be clearly seen in **Figure 2.20b**, adding a single adduct to the protein removes any possible overlap (because dividing the mass of the bound dimer by two would overlap with a monomer with 'half a lipid' attached). If charge stripping occurs solely due to the surface properties, this clearly-resolved case can be used to estimate the proportion of the *apo* precursor which is expected to be charge-stripped. This calculated value can then be subtracted from the overlapping peak, to estimate the amount of monomer and dimer in it. I have compared the 6+/7+ ratio for the *apo* SS calculated directly from the IM data to 6+CDL/7+CDL ratios for the lipid-bound species calculated from MS data (**Figure 2.20c**). The values observed are indeed very similar, especially at lower SID voltages. At high end the variation in fractions are higher, but the overall abundances of both the 6+ and the 7+ are very low at this point, so the absolute error remains small. While this method provides only an estimate, it might well be the best alternative in some challenging cases. One limitation of this technique is the requirement for a lipid or ligand that can remain bound throughout the entire experiment, which might not always exist.

2.9 Summary and Conclusions

In this chapter, I demonstrated the utility of a combined CIU-SID experiment to investigate the role of lipids bound to MP complexes by native MS. Firstly, I showed

that the binding of a PG lipid does not induce any stabilisation of AmtB trimer as indicated by SID, despite a large stabilisation observed by CIU – this is attributed to a non-interfacial nature of the specific lipid-binding pocket[4]. Secondly, I showed that SID has sufficient resolution and sensitivity to distinguish between the interface stabilising (CDL) and non-stabilising (PE, PG) PL classes for SS. Thirdly, I demonstrated the ability of SID to observe very subtle changes to the surface cleavage energy and developed a novel method to analyse these small differences. Finally, I outlined a practical way to perform these experiments, including potential sources of error. In summary, the following steps are proposed:

1 Sample Preparation. The protein complex of interest has to be extensively purified, so that no or minimal impurities can be observed in the spectra. A complete detergent exchange has to be achieved into an MS-compatible detergent, such as C8E4. All endogenous lipids have to be removed. For the best results, high concentration of the protein has to be used (10-20 μM). Add synthetic homogeneous lipid to the sample right before the experiment, one species at the time.

2 Native MS. Desolvate the protein completely in the source region, so that a clean peak can be quadrupole selected for MS/MS experiments. If an IM cell is present, perform an unfolding ramp to find a balance between insufficient desolvation and induced unfolding. It is important to check the initial conditions for all of the lipid-bound states which will be investigated by SID, as changing them between species will compromise comparison.

3 CIU. Perform CIU experiments (if IM is available) on the proteins bound to the lipids of interest; make sure that at least one control lipid is included.

4 SID. Select a sampling interval (SID voltage increment); a quick preliminary experiment can be run to identify the region of interest and increase the sampling interval accordingly. Aim for a minimum of 5 to 6 data points if possible. Select the range to start at 100% precursor and finish at 0% precursor (or until a 'plateau' is reached).

If non-resolvable peak overlaps occur due to charge reduction, use a freshly coated SAM surface.

Create an SID-ERMS plot, with the relative intensity of the undissociated dimer plotted against the SID voltage. If the lipids are falling off, only go up to the voltage at which significant falling off (roughly above 10%) is observed. Detect large differences visually.

5 Statistical analysis. Only perform this step if looking for very subtle differences. Ideally, a 'state-of-the-art' mass spectrometer needs to be used to collect the data.

Using a mathematical fit (for example, a sigmoidal curve), or otherwise, identify the region of the largest difference.

Perform a two-way mixed ANOVA, with SID energy as the independent repeating variable, species identity as the independent non-repeating variable, and percentage of the undissociated precursor as the dependent variable. If ANOVA test detects a statistically significant difference at 0.05 confidence level, use a post-hoc Bonferroni test (or another appropriate post-hoc test) to identify pairwise differences.

By utilising this workflow, important positional information about MP-lipid interactions can be obtained, with SID identifying specific interfacial lipids and CIU

identifying specific structural (subunit-stabilising) lipids. Potentially, this approach can be extended to MP interactions with other ligands, such as the novel 'matchmaker' drugs that work by creating a complex between two entities, so that the 'effector' can act on the 'target'[55]. With the recent technological advancements of both mass spectrometers in general[39,56] and the SID technology in particular[53,54] it is possible to envision the high-resolution ERMS-SID experiments described here to be more widely applicable in the near future.

Materials and Methods

Membrane protein expression and purification. Vs SS plasmid was transformed into BL21 λDE3 *E. coli* (Novagen) and expressed and purified as reported previously[15]. Briefly, colonies were used to inoculate 4 100ml flasks containing LB media and kanamycin antibiotic at the concentration of 50mg/ml and grown at 37°C overnight. 15 ml of this primary culture was used to inoculate each of the 24 flasks containing 1L of the same media, which were then grown at 37°C until the optical density reached 0.8. After that, expression was induced with 1ml of 0.2M isopropyl β- d-1-thiogalactopyranoside (IPTG) and the cells were grown overnight at 22°C.

After centrifugal harvesting at 5000g at 4°C, the cells were supplemented with protease inhibitor tablets and lysed using M-110 PS Microfluidiser (Microfluidics). After pelleting the cell debris at 20000g for 20 min, membranes were collected by ultracentrifugation at 100000g for 2h. After homogenising, powdered DDM was added at 2% w/v concentration, and the protein was extracted with gentle agitation at 4°C overnight. His-tagged SS was purified by immobilised metal affinity chromatography (IMAC), dialysed overnight with HRV 3C Protease to cleave the His-tag, and then purified with reverse IMAC. The protein was concentrated using

a 50K molecular weight cut-off (MWCO) concentrator to a final volume of 250 μl, exchanged into a buffer containing C8E4 detergent (at 2xCMC concentration) over a 24-ml SEC column and re-concentrated using a 10K MWCO concentrator. Flash-frozen samples were stored at -80 ^{o}C.

Y62W SS mutation was performed with the use of geneBlock (Integrated DNA Technologies), then expressed and purified in the same manner.

Ec AmtB was expressed and purified as reported previously[4].

Sample preparation for native MS. Proteins were exchanged into 200 mM AA (Sigma Aldrich) solution containing 2xCMC of C8E4 detergent using a Biospin column (Micro Bio-Spin 6, Bio-Rad), to a final concentration of 10-15 μM (measured using DS 11+ Spectrophotometer (DeNovix)). Phospholipids were purchased from Avanti. Aqueous solutions of PLs were prepared in the form of 10 mg/ml stocks as described elsewhere[57]. Each PL was diluted by 200 mM AA/2xCMC C8E4 solution and mixed with the protein at 4:1 Protein-lipid v/v ratio directly before native MS analysis. PL concentrations were adjusted until high intensity of the lipid-bound protein was achieved in mass spectra.

Native MS. Nanoelectrospray ionisation was initiated from borosilicate needles, pulled and gold-coated in house as previously described[58]. The mass spectrometers used and their related modifications are described in **Chapter 1 (Figure 1.6).** For the SID experiments, 2-(perfluorodecyl)ethanethiol was used to prepare the SAM surfaces, as reported previously[59].

Synapt G2 data. For SID of AmtB, the following parameters were used: Capillary voltage: 1.4 kV, Sampling cone: 200 V, Extraction cone: 50 V, LM resolution and

HM resolution: 10, Aperture 1: 10 (increasing from 0 significantly improved transmission of high *m/z* ions), Trap collision energy (CE): 10 V, Transfer CE: 5 V, Trap gas flow (Argon): 6 ml/min, IMS gas flow: 40 ml/min, Helium cell gas flow: 120 ml/min, Trap wave velocity (WV): 800, Trap wave height (WH): 10.0, IM WV: 700, IM WH: 40.0, Transfer WV: 85, Transfer WH: 3.0, Backing Pressure: 8.5-9.0 mbar. For 50 V SID, Trap Bias of 120 V was used; to change SID energy, Trap Bias, as well as Entrance 1 and Front Bottom SID parameters were changed by the same amount.

For SID of SS, the following parameters were used: Capillary voltage: 1.2 kV, Sampling cone: 150 V, Extraction cone: 4 V, LM resolution: 10, HM resolution: 12, Trap CE: 4 V, Transfer CE: 2 V, Trap gas flow (Argon): 2 ml/min, IMS gas flow: 40 ml/min, Helium cell gas flow: 120 ml/min, Trap WV: 250, Trap WH: 3.0, IM WV: 950, IM WH: 40, Transfer WV: 100, Transfer WH: 1.0, Backing Pressure: 5 mbar. For 50 V SID, Trap Bias of 110 V was used (corresponding to the injection voltage of 50 V); to change SID energy, Trap Bias, as well as Entrance 1 and Front Bottom SID parameters were changed by the same amount.

For CIU of SS, the following parameters were used: Capillary voltage: 1.2 kV, Sampling cone:130 V, Extraction cone:4 V, Transfer CE: 2 V, Trap gas flow (Argon): 4 ml/min, IMS gas flow: 40 ml/min, Helium cell gas flow: 120 ml/min, Trap WV: 250, Trap WH: 3.0, IM WV: 950, IM WH: 40, Transfer WV: 100, Transfer WH: 1.0, Backing Pressure: 5 mbar.

Synapt G1 data. To calculate the CCS of SS, drift voltages from 60 V to 95 V were used. The following parameters were used: Capillary voltage: 1.4 kV, Sampling cone: 200 V, Extraction cone: 1 V, Trap CE: 4 V, Trap gas flow (Nitrogen): 2 ml/min,

IMS gas flow (Helium): 60 ml/min, Trap WV: 100, Trap WH: 1.0, Transfer WV: 100, Transfer WH: 10.0, Backing Pressure: 4 mbar.

Cyclic prototype Synapt data. For SID of SS, the following parameters were used: Capillary voltage: 1.2 kV, Sampling cone: 150 V, Stepwave 1,2 velocity: 300, Stepwave 1,2 Height: 5, LM resolution: 5, HM resolution: 5, Trap CE: 5 V, Transfer CE: 0 V, Trap gas flow (Argon): 2 ml/min, IMS cell pressure: 2.2 mbar, Trap WV: 650, Trap WH: 0.5, Cyclic WV: 375, Cyclic WH: 40. For 38V SID, IMS Bias (equivalent to Trap Bias on the older Synapts) of 50 V was used, (injection voltage: 15 V); to change SID energy, IMS Bias, as well as Entrance 1 and Front Bottom SID parameters were changed by the same amount. SID parameters used were: Entrance 1: -130 V, Entrance 2: -135 V, Front Top: -209 V, Front Bottom: -136 V, Surface: -160 V, Middle Bottom: -219 V, Back Top: -291 V, Back Bottom: -151 V, Exit 1: -180 V, Exit 2: -170 V.

Data analysis. CIU experiments. Data were processed with PULSAR software as previously described[37]. The final conversion of data into a model was performed using an in-house python script (provided by Tim Allison).

SID experiments. Peak fitting and peak area calculation were performed manually in UniDEC software[60]. Each peak was manually defined as a 1+ charge species (so that just the peak areas are calculated, not adjusted for charge by the software) (**Figure 2.21a**). AmtB: Data were processed in MassLynx software (Waters Corporation) (Savitzky Golay, smooth window: 80, number of smooths: 3, background subtraction polynomial order: 2). UniDEC parameters: Gaussian smoothing: 100, Bin every: 1, Sample mass every: 100, Peak shape function: Gaussian, Peak FWHM: Automatic (calculated for every spectrum).

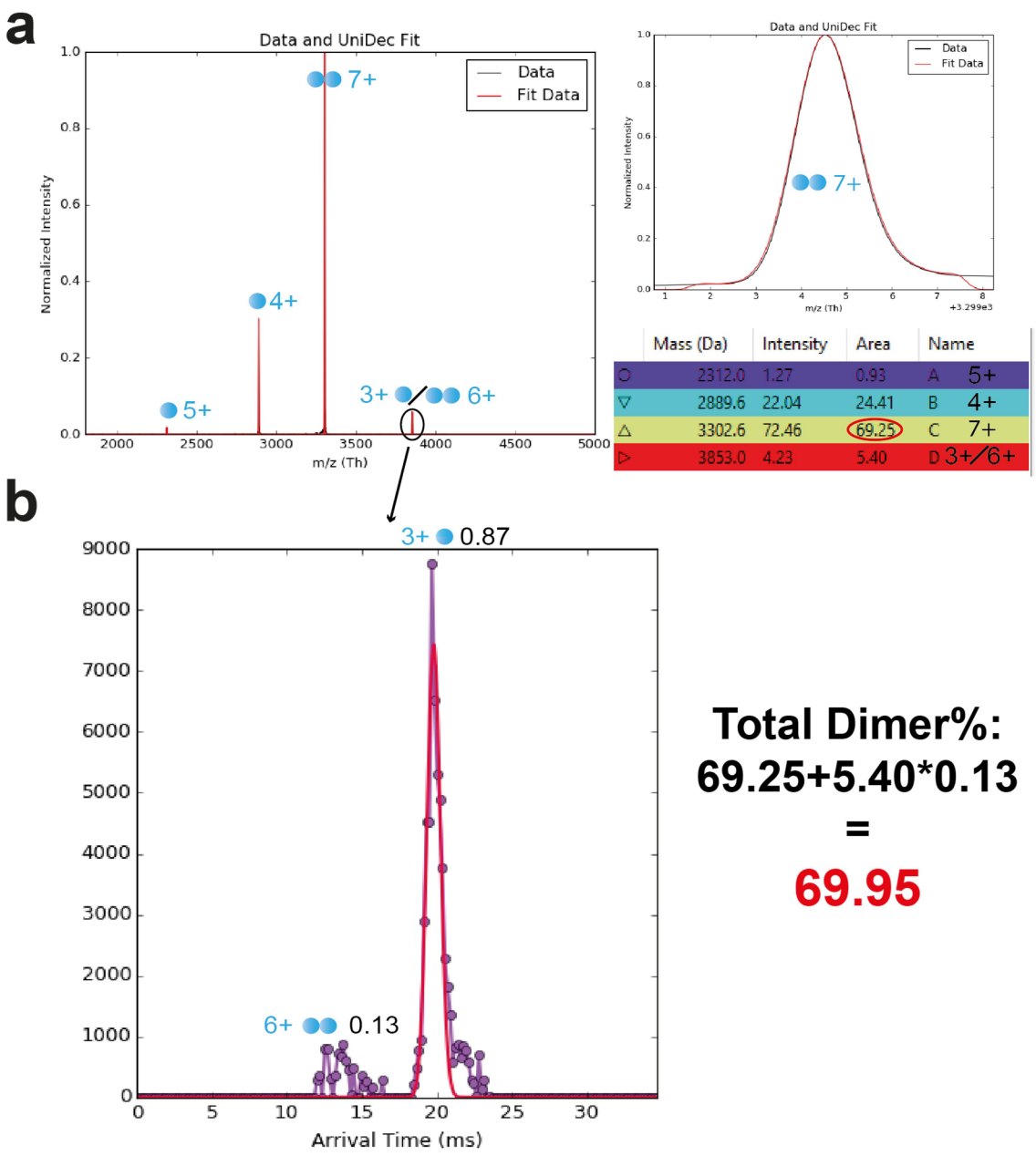


Figure 2.21 Dimer percentage calculations. An example of conversion of MS data into dimer percentage values for *apo* SS at 35V SID. **a,** Smoothed MS data are fitted in UniDEC software (a zoomed in view of the fitting for the 7+ precursor peak is shown on the right) and relative peak areas are calculated. **b,** For the overlapping monomer/dimer peak, IM data is used to calculate the relative fraction of monomer and dimer in the peak using PULSAR software. The relative intensities of the 7+ precursor peak and the 6+ striped dimer peak are summed to produce a single dimer percentage value. An average of 3 such values from separate repeat experiments at the same SID voltage are then used to produce ERMS plots, such as the one shown in **Figure 2.15a**.

SS: Data were processed in MassLynx Software (Mean, smooth window: 30, number of smooths: 2, no background subtraction). UniDEC parameters: Gaussian smoothing: 0, Bin every: 0, Sample mass every: 0.2, Peak shape function: Gaussian, Peak FWHM: 0.4, Intensity threshold: 0.005.

To calculate the amount of charge reduction, IM data for the 6+ dimer / 3+ monomer peak were extracted from PULSAR software into excel file, with contribution of each species calculated as the relative intensity of the corresponding peak (**Figure 2.21b**).

CCS calculation. SS CCS was calculated from the drift-tube IM data using PULSAR software. To estimate theoretical CCS values for the different conformations of SS, the following PDB files were used: Inward-open: 4X5M; Occluded: 4QNC, 4RNG; Outward Open: 4QND, 4X5N, 5UHS. (*Note: another structure in the inward-open conformation exists (5UHQ), but the PDB file failed to produce a reasonably looking dimer*). IMPACT software[61] was used to calculate CCS values from crystal structures using projection approximation (PA), which were then corrected using equation (1.6):

$$CCS_{calc} = 1.14 CCS_{PA} \left(\frac{M_{exp}}{M_{PDB}}\right)^{\frac{2}{3}},$$

Where M_{exp} is the molecular weight of the protein used in the experiment and M_{PDB} is the molecular weight of the protein in the crystal structure. Where more than one structure was present for a conformation, the values were averaged to produce the final estimate.

Curve fitting. All of the sigmoidal curve models available in Origin Pro software were checked, to see which one gives the best initial fit to the data (visually).

BiDoseRep function was chosen, with $1/s^2$ used for weighting (where *s* is the standard deviation). For the data points where the standard deviation was equal to 0 (only occurred when the dimer intensity was equal to 0 for all three repeats), the values were changed to 0.1.The *prob*>|*t*| test was used to access the relevance of all function parameters by setting them to 0 and checking if that affects the fit. This test was passed for all of the parameters at 0.05 level, apart from the A1 parameter which is supposed to be 0 as it represents the point of complete dimer dissociation.

ANOVA testing. ANOVA tests were performed in Origin Pro software. For the one-way ANOVA assumption of normality was verified by Shapiro-Wilks test (at 0.05 significance level); assumption for homogeneity of variance was verified by Levene's test (at 0.05 significance level). For the data points where Levene's test was not passed, Kruskal-Wallis H test was used to verify ANOVA results (at 0.05 significance level; note: for the value at 55V SID in **Figure 2.16a** the outcome narrowly missed the target, with a value of 0.051). In the cases where ANOVA confirmed statistically significant differences between means, post hoc Tukey HSD test was used (at 0.05 significance level) to identify pairwise differences.

For the two-way mixed ANOVA, normality check is replaced by Maulchy's sphericity test (at 0.05 significance level). It is passed for the smaller voltage set data (**Figure 2.16b**), but is violated for the larger set (**Figure 2.16c**) – a correction was applied. Bonferroni post hoc test is used instead of Tukey.

MD simulations. The simulations of SS have been carried out in GROMACS[62] using the Martini force field version 2.2 [63] which merges around four heavy atoms into a bead. The atomistic structures of both *Vibrio sp.* SS (PDB: 4QND[29]) and the

homology model have been coarse-grained and inserted into lipid bilayers using the MemProtMD pipeline[64].

The lipid bilayers for the simulations were composed of 70 % POPG 20% POPE and 10% cardiolipin (CDL) of different chainlength. For the unsaturated, the unsaturated hydrogen bond was kept central. The simulations were carried out for five microseconds and five replica for each lipid composition. Contacts of residues was defined by a distance of less than six angstrom between a residue bead and a lipid headgroup bead. The interfacial residues have been defined visually.

Acknowledgments

First and foremost, I would like to thank Joseph Gault for his continuous support, useful discussions and contributions to experimental design throughout this project (and for providing Ec AmtB plasmid and Q Exactive SID data). I thank Daniel Quetschlich for performing molecular dynamics simulations, as well as designing and co-purifying the Y62W SemiSWEET mutant, with assistance from Jani Bolla. I would like to thank Todd Mize and Andy Dolan for their help with setting and maintaining the SID device on the Synapt G2, and Sophie Harvey and Vicki Wysocki for their help with SID methodology and instrumentation and for providing the SID cell and the surface coating material. I thank people at Waters Corporation, especially Kevin Giles and Jakub Ujma, for allowing me to use their prototype Synapt instrument – without them a major part of this chapter would be impossible. Jakub also helped me with the instrumental set-up. I would like to thank Tim Allison for his help with PULSAR software. I thank Andy Baldwin for writing a custom piece of code for me to speed up the SID data analysis (even though I ended up performing manual analysis on all of the final datasets for improved control over

accuracy). Finally, I would also like to thank Stephen Ambrose for providing Vs SS cell stocks and plasmid.

References

1 Bolla, J. R., Agasid, M. T., Mehmood, S. & Robinson, C. V. Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. *Annu Rev Biochem* **88**, 85-111, doi:10.1146/annurev-biochem-013118-111508 (2019).

2 Liko, I. *et al.* Lipid binding attenuates channel closure of the outer membrane protein OmpF. *Proc Natl Acad Sci U S A* **115**, 6691-6696, doi:10.1073/pnas.1721152115 (2018).

3 Contreras, F. X., Ernst, A. M., Wieland, F. & Brugger, B. Specificity of intramembrane protein-lipid interactions. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a004705 (2011).

4 Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172-+, doi:10.1038/nature13419 (2014).

5 Rohacs, T., Chen, J., Prestwich, G. D. & Logothetis, D. E. Distinct specificities of inwardly rectifying K(+) channels for phosphoinositides. *J Biol Chem* **274**, 36065-36072, doi:10.1074/jbc.274.51.36065 (1999).

6 Cong, X., Liu, Y., Liu, W., Liang, X. & Laganowsky, A. Allosteric modulation of protein-protein interactions by individual lipid binding events. *Nat Commun* **8**, 2203, doi:10.1038/s41467-017-02397-0 (2017).

7 Perini, D. A., Alcaraz, A. & Queralt-Martin, M. Lipid Headgroup Charge and Acyl Chain Composition Modulate Closure of Bacterial beta-Barrel Channels. *Int J Mol Sci* **20**, doi:10.3390/ijms20030674 (2019).

8 Lee, A. G. Lipid-protein interactions. *Biochem Soc Trans* **39**, 761-766, doi:10.1042/BST0390761 (2011).

9 van Gestel, R. A. *et al.* The influence of the acyl chain composition of cardiolipin on the stability of mitochondrial complexes; an unexpected effect of cardiolipin in alpha-ketoglutarate dehydrogenase and prohibitin complexes. *J Proteomics* **73**, 806-814, doi:10.1016/j.jprot.2009.11.009 (2010).

10 Contreras, F. X. *et al.* Molecular recognition of a single sphingolipid species by a protein's transmembrane domain. *Nature* **481**, 525-529, doi:10.1038/nature10742 (2012).

11 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, doi:10.1038/nrm4080 (2015).

12 Goth, M. & Pagel, K. Ion mobility-mass spectrometry as a tool to investigate protein-ligand interactions. *Anal Bioanal Chem*, doi:10.1007/s00216-017-0384-9 (2017).

13 Liu, Y., Cong, X., Liu, W. & Laganowsky, A. Characterization of Membrane Protein-Lipid Interactions by Mass Spectrometry Ion Mobility Mass Spectrometry. *J Am Soc Mass Spectrom* **28**, 579-586, doi:10.1007/s13361-016-1555-1 (2017).

14 Landreh, M. *et al.* Integrating mass spectrometry with MD simulations reveals the role of lipids in Na+/H+ antiporters. *Nat Commun* **8**, 13993, doi:10.1038/ncomms13993 (2017).

15 Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **541**, 421-424, doi:10.1038/nature20820 (2017).

16 Chorev, D. S. *et al.* Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science* **362**, 829-834, doi:10.1126/science.aau0976 (2018).

17 Gupta, K. *et al.* Identifying key membrane protein lipid interactions using mass spectrometry. *Nat Protoc* **13**, 1106-1120, doi:10.1038/nprot.2018.014 (2018).

18 Boeri Erba, E., Barylyuk, K., Yang, Y. & Zenobi, R. Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal Chem* **83**, 9251-9259, doi:10.1021/ac201576e (2011).

19 Xia, Z. & Williams, E. R. Effect of droplet lifetime on where ions are formed in electrospray ionization. *Analyst*, doi:10.1039/c8an01824c (2018).

20 Quintyn, R. S., Yan, J. & Wysocki, V. H. Surface-Induced Dissociation of Homotetramers with D-2 Symmetry Yields their Assembly Pathways and Characterizes the Effect of Ligand Binding. *Chem Biol* **22**, 583-592, doi:10.1016/j.chembiol.2015.03.019 (2015).

21 Harvey, S. R. *et al.* Relative interfacial cleavage energetics of protein complexes revealed by surface collisions. *Proc Natl Acad Sci U S A* **116**, 8143-8148, doi:10.1073/pnas.1817632116 (2019).

22 Harvey, S. R., Liu, Y., Liu, W., Wysocki, V. H. & Laganowsky, A. Surface induced dissociation as a tool to study membrane protein complexes. *Chem Commun (Camb)*, doi:10.1039/c6cc09606a (2017).

23 Zhou, M. W., Dagan, S. & Wysocki, V. H. Impact of charge state on gas-phase behaviors of noncovalent protein complexes in collision induced dissociation and surface induced dissociation. *Analyst* **138**, 1353-1362, doi:10.1039/c2an36525a (2013).

24 Donor, M. T., Mroz, A. M. & Prell, J. S. Experimental and theoretical investigation of overall energy deposition in surface-induced unfolding of protein ions. *Chem Sci* **10**, 4097-4106, doi:10.1039/c9sc00644c (2019).

25 Ma, X., Zhou, M. W. & Wysocki, V. H. Surface Induced Dissociation Yields Quaternary Substructure of Refractory Noncovalent Phosphorylase B and Glutamate Dehydrogenase Complexes. *J Am Soc Mass Spectr* **25**, 368-379, doi:10.1007/s13361-013-0790-y (2014).

26 Sipe, S. N., Patrick, J. W., Laganowsky, A. & Brodbelt, J. S. Enhanced Characterization of Membrane Protein Complexes by Ultraviolet Photodissociation Mass Spectrometry. *Anal Chem* **92**, 899-907, doi:10.1021/acs.analchem.9b03689 (2020).

27 Patrick, J. W. *et al.* Allostery revealed within lipid binding events to membrane proteins. *Proc Natl Acad Sci U S A* **115**, 2976-2981, doi:10.1073/pnas.1719813115 (2018).

28 Lee, Y., Nishizawa, T., Yamashita, K., Ishitani, R. & Nureki, O. Structural basis for the facilitative diffusion mechanism by SemiSWEET transporter. *Nat Commun* **6**, 6112, doi:10.1038/ncomms7112 (2015).

29 Xu, Y. *et al.* Structures of bacterial homologues of SWEET transporters in two distinct conformations. *Nature* **515**, 448-452, doi:10.1038/nature13670 (2014).

30 Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat Methods* **13**, 333-336, doi:10.1038/nmeth.3771 (2016).

31 Sobott, F., Hernandez, 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, doi:10.1021/ac0110552 (2002).

32 Zhou, M. W., Dagan, S. & Wysocki, V. H. Protein Subunits Released by Surface Collisions of Noncovalent Complexes: Nativelike Compact Structures Revealed by Ion Mobility Mass Spectrometry. *Angew Chem Int Edit* **51**, 4336-4339, doi:10.1002/anie.201108700 (2012).

33 Zhou, M. W., Huang, C. S. & Wysocki, V. H. Surface-Induced Dissociation of Ion Mobility-Separated Noncovalent Complexes in a Quadrupole/Time-of-Flight Mass Spectrometer. *Anal Chem* **84**, 6016-6023, doi:10.1021/ac300810u (2012).

34 Zhou, M. W. & Wysocki, V. H. Surface Induced Dissociation: Dissecting Noncovalent Protein Complexes in the Gas phase. *Accounts Chem Res* **47**, 1010-1018, doi:10.1021/ar400223t (2014).

35 Benesch, J. L. P. Collisional Activation of Protein Complexes: Picking Up the Pieces. *J Am Soc Mass Spectr* **20**, 341-348, doi:10.1016/j.jasms.2008.11.014 (2009).

36 Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol Rev* **40**, 133-159, doi:10.1093/femsre/fuv008 (2016).

37 Allison, T. M. *et al.* Quantifying the stabilizing effects of protein-ligand interactions in the gas phase. *Nat Commun* **6**, doi:10.1038/ncomms9551 (2015).

38 Landreh, M., Costeira-Paulo, J., Gault, J., Marklund, E. G. & Robinson, C. V. Effects of Detergent Micelles on Lipid Binding to Proteins in Electrospray Ionization Mass Spectrometry. *Anal Chem* **89**, 7425-7430, doi:10.1021/acs.analchem.7b00922 (2017).

39 Giles, K. *et al.* A Cyclic Ion Mobility-Mass Spectrometry System. *Anal Chem* **91**, 8564-8573, doi:10.1021/acs.analchem.9b01838 (2019).

40 Quintyn, R. S., Zhou, M. W., Yan, J. & Wysocki, V. H. Surface-Induced Dissociation Mass Spectra as a Tool for Distinguishing Different Structural Forms of Gas-Phase Multimeric Protein Complexes. *Anal Chem* **87**, 11879-11886, doi:10.1021/acs.analchem.5b03441 (2015).

41 Seffernick, J. T., Harvey, S. R., Wysocki, V. H. & Lindert, S. Predicting Protein Complex Structure from Surface-Induced Dissociation Mass Spectrometry Data. *ACS Cent Sci* **5**, 1330-1341, doi:10.1021/acscentsci.8b00912 (2019).

42 Dongre, A. R., Jones, J. L., Somogyi, A. & Wysocki, V. H. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. *J Am Chem Soc* **118**, 8365-8374, doi:DOI 10.1021/ja9542193 (1996).

43 Jellen, E. E., Chappell, A. M. & Ryzhov, V. Effects of size of noncovalent complexes on their stability during collision-induced dissociation. *Rapid Commun Mass Sp* **16**, 1799-1804, doi:10.1002/rcm.791 (2002).

44 Field, A. *Discovering Statistics using IBM SPSS Statistics*. (Sage Publications Ltd., 2013).

45 Crowe, M. C. & Brodbelt, J. S. Evaluation of noncovalent interactions between peptides and polyether compounds via energy-variable collisionally activated dissociation. *J Am Soc Mass Spectr* **14**, 1148-1157, doi:10.1016/S1044-0305(03)00452-5 (2003).

46 Latorraca, N. R. *et al.* Mechanism of Substrate Translocation in an Alternating Access Transporter. *Cell* **169**, 96-107 e112, doi:10.1016/j.cell.2017.03.010 (2017).

47 Laszlo, K. J., Buckner, J. H., Munger, E. B. & Bush, M. F. Native-Like and Denatured Cytochrome c Ions Yield Cation-to-Anion Proton Transfer Reaction Products with Similar Collision Cross-Sections. *J Am Soc Mass Spectrom*, doi:10.1007/s13361-017-1620-4 (2017).

48 Yan, J. *et al.* Surface-Induced Dissociation of Protein Complexes in a Hybrid Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Anal Chem* **89**, 895-901, doi:10.1021/acs.analchem.6b03986 (2017).

49 Zhou, M. *et al.* Surface Induced Dissociation Coupled with High Resolution Mass Spectrometry Unveils Heterogeneity of a 211 kDa Multicopper Oxidase Protein Complex. *J Am Soc Mass Spectrom*, doi:10.1007/s13361-017-1882-x (2018).

50 Lossl, P., Snijder, J. & Heck, A. J. R. Boundaries of Mass Resolution in Native Mass Spectrometry. *J Am Soc Mass Spectr* **25**, 906-917, doi:10.1007/s13361-014-0874-3 (2014).

51 Busch, F. *et al.* Localization of Protein Complex Bound Ligands by Surface-Induced Dissociation High-Resolution Mass Spectrometry. *Anal Chem* **90**, 12796-12801, doi:10.1021/acs.analchem.8b03263 (2018).

52 Alvarez, J., Futrell, J. H. & Laskin, J. Soft-landing of peptides onto self-assembled monolayer surfaces. *J Phys Chem A* **110**, 1678-1687, doi:10.1021/jp0555044 (2006).

53 Stiving, A. Q., Gilbert, J. D., Jones, B. J. & Wysocki, V. H. A Tilted Surface and Ion Carpet Array for SID. *J Am Soc Mass Spectrom* **31**, 458-462, doi:10.1021/jasms.9b00009 (2020).

54 Snyder, D. T., Panczyk, E. M., Somogyi, A., Kaplan, D. A. & Wysocki, V. Simple and Minimally Invasive SID Devices for Native Mass Spectrometry. *Anal Chem* **92**, 11195-11203, doi:10.1021/acs.analchem.0c01657 (2020).

55 Deshaies, R. J. Multispecific drugs herald a new era of biopharmaceutical innovation. *Nature* **580**, 329-338, doi:10.1038/s41586-020-2168-1 (2020).

56 Fort, K. L. *et al.* Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. *Analyst* **143**, 100-105, doi:10.1039/c7an01629h (2017).

57 Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat Protoc* **8**, 639-651, doi:10.1038/nprot.2013.024 (2013).

58 Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* **2**, 715-726, doi:10.1038/nprot.2007.73 (2007).

59 Galhena, A. S., Dagan, S., Jones, C. M., Beardsley, R. L. & Wysocki, V. N. Surface-induced dissociation of peptides and protein complexes in a quadrupole/Time-of-Flight mass spectrometer. *Anal Chem* **80**, 1425-1436, doi:DOI 10.1021/ac701782q (2008).

60 Marty, M. T. *et al.* Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal Chem* **87**, 4370-4376, doi:10.1021/acs.analchem.5b00140 (2015).

61 Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J. & Benesch, J. L. Collision cross sections for structural proteomics. *Structure* **23**, 791-799, doi:10.1016/j.str.2015.02.010 (2015).

62 Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701-1718, doi:10.1002/jcc.20291 (2005).

63 Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* **111**, 7812-7824, doi:10.1021/jp071097f (2007).

64 Stansfeld, P. J. *et al.* MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **23**, 1350-1361, doi:10.1016/j.str.2015.05.006 (2015).

3 Development of Controlled Delipidation Methods for Native Mass Spectrometry Characterisation of Membrane Protein Complexes

Abstract

Membrane proteins are unstable outside their native bilayer environment, so they must be solubilised in membrane mimetics for *in vitro* studies. Detergent micelles are among the most commonly used membrane mimetics, however selection of the appropriate species for a particular application and membrane protein species often requires lengthy screening protocols. Displacement of endogenous lipids with detergent molecules is one of the important considerations for detergent selection, with varying degrees of delipidation desired for different applications.

In this chapter, I develop a method for a controlled and predictable delipidation of membrane proteins by detergents, with evidence from native mass spectrometry characterisation. I introduce a set of detergents, four of which are completely novel, with gradually changing delipidating properties. I demonstrate a standardised purification method applicable to membrane proteins with different molecular weights and oligomeric states, which can be used to predictably preserve the desired degree of protein-phospholipid interactions. Furthermore, I present a separate method for controlled delipidation of lipopolysaccharide.

In addition, I investigate underlying principles behind variable delipidating properties of detergents. Delipidation strengths of detergents are shown to be related to physical parameters, such as hydrophobicity and packing density of

micelles. Finally, the currently proposed mechanism for charge reduction of membrane proteins by detergents during native mass spectrometry experiments is shown to be in agreement with the data obtained from the detergent set in this work. Overall, the results presented in this chapter pave the way for a departure from the current empirical screening of detergents and towards a more 'rational' approach.

3.1 Introduction

Membrane proteins (MPs) are surrounded by a vast number of structurally diverse lipids in their native bilayer environment[1-4]. MP-lipid interactions span a broad range of specificities, from the weakly interacting annular lipids which are in fast exchange with the bulk membrane, to highly specific non-annular lipids, often essential for structure and function of MP complexes[5-9] (see **Section 1.3** for more detail). This diversity presents many challenges for structural biology studies, because most of the existing biophysical techniques produce the best results when performed on highly homogeneous samples[10,11] (**Section 1.4**). This leads to desire for delipidation of endogenous lipids and the use of a more controlled artificial membrane mimetic environment. However, complete removal of the native lipids can be detrimental for MPs, either completely destabilising and aggregating them[12] or causing significant deviations from their native conformations or oligomeric states[13,14]. Therefore, a controlled approach to delipidation can be highly beneficial.

One way to control delipidation is by employing different detergents for MP purification, because every detergent exhibits unique delipidating properties[15]. Detergents with mild delipidating properties, such as *n*-dodecyl-β-d-maltopyranoside (DDM), can co-purify substantial amount of lipids[16]. Detergents

with strong delipidating properties include *n*-octyl-β-d-glucopyranoside (OG), tetraethylene glycol monooctyl ether (C8E4) and n-nonyl-β-d-glucopyranoside (NG) [17]. Complete or partial exchange of the detergent environment from a mild to a strongly delipidating detergent is, therefore, a common way to delipidate MPs lipids in a controlled manner[17,18]. Delipidation strategies including only mild detergents involve prolonged incubation of MPs over hours or even days with high detergent concentrations[19] (see **Section 1.6** for more discussion). Both delipidation methods require detergent and the desired degree of delipidation needs to be determined by trial and error for independent sample batches.

Although it is obvious that the delipidating properties of detergents depend on their molecular structure and working concentration, the properties that make a detergent more or less delipidating are not entirely understood. Consequently, no design guidelines have yet been developed that would allow optimising the structure of detergents for MP delipidation. Recently, a novel family of oligoglycerol detergents (OGDs) was introduced to MP research. The molecular structure of OGDs is highly modular and can be tuned for individual applications[20]. Synthetic protocols enable rapid exchange of OGD building blocks, such as the head group, the linker and the hydrophobic tail. Thus, the architecture of OGD can be synthetically modified to independently optimise average charge state of MPs or the degree of delipidation observed by native mass spectrometry (MS)[20]. However, OGDs with mild delipidating properties turned out to be less suitable for protein purification than OGDs with strong delipidating properties. Furthermore, design guidelines that cover not only OGDs but also other detergent families have not yet been proposed. In this chapter, a novel set of detergents is presented (**Figure 3.1**) (*designed and synthesised by Leonhard Urner*). They were employed to evaluate

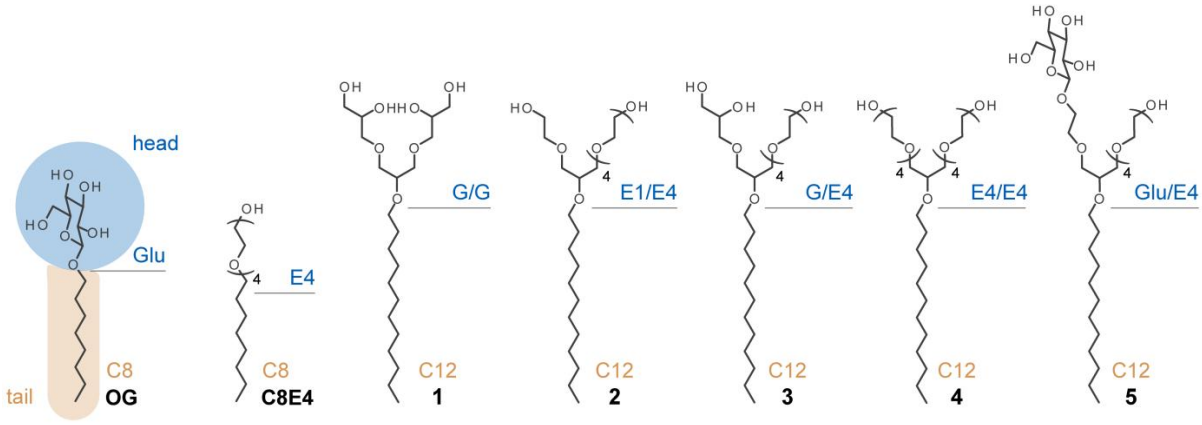


Figure 3.1 Detergents. Novel detergents (**2**-**5**) were designed by altering the head group of the previously reported detergent **1** (ref 20). This set of detergents was designed to have a gradual change in both hydrophobic-lipophilic balance (HLB) and packing parameter. Abbreviations: OG, *n*-octyl-β-d-glucopyranoside; C8E4, tetraethylene glycol monooctyl ether; Glu, glucose; E4, tetraethylene; G, glycerol; E1, monoethylene. Figure made by Leonhard Urner.

whether changing general physical properties of detergents, such as the overall hydrophobicity and shape, enables fine tuning of delipidating properties of detergents more broadly.

3.2 Detergent Design

(*Note: I did not contribute to the detergent design or synthesis, only to the native MS characterisation*). The detergent design for MP delipidation is based on the hypothesis that the amount of lipids taken up by detergent micelles depends on the transport properties of the micelles, and that the transport properties in turn depend on the overall hydrophobicity and shape[21]. To evaluate the applicability of this hypothesis to MP delipidation, a set of detergents with gradually changing hydrophobicity (as indicated by hydrophobic-lipophilic balance (HLB) and the shape (as indicated by the packing parameter) is required.

The HLB concept describes the overall hydrophobicity of a detergent in relation to the mass balance between its water-soluble head group and water-insoluble tail[22]. The expression for the HLB value has a well-known form:

$$HLB = 20 \cdot \left(1 - \frac{MW_{tail}}{MW}\right), \quad (3.1)$$

where MW_{tail} is the molecular weight of the hydrophobic tail and MW is the molecular weight of the detergent. The packing parameter approximates the detergent's molecular shape in relation to the length and volume of the hydrophobic tail and the area of the head group[23]. The expression for the packing parameter has the form:

$$p = \frac{V_{tail}}{l_{tail} \cdot A_{head}}, \quad (3.2)$$

where p is the packing parameter, V_{tail} the volume of the hydrophobic tail, l_{tail} the length of the hydrophobic tail, and A_{head} the area of the head group that is occupied at the interface between the detergent aggregate and solvent. While HLB can be easily calculated from the knowledge of the chemical structure of a particular detergent, determining the packing parameter is less straight-forward, with modelling required to calculate various parameters (see **Materials and Methods**). Both concepts have been used to rationalise the packing density[24,25], morphology[23-26], and transport behaviour of detergent aggregates[25,27]. Therefore, one of the aims of the work presented in this chapter was to evaluate whether these parameters can be linked to delipidation propensities.

In order to investigate the effect of physical parameters of detergents on MP delipidation, a set of detergents was designed (**Figure 3.1**). The previously

reported OGD [G1]-ether-C12 detergent was used as a starting point (shown as detergent **1** in **Figure 3.1**) as it was shown to be compatible with a wide range of proteins as well as with native MS[20,28]. Detergent **1** was previously demonstrated to be relatively strongly delipidating, similarly to C8E4 detergent that is often used in native MS; however, one of the advantages of detergent **1** is its high extraction efficiency and MP stability, while C8E4 does not have these particular useful properties[20]. To gradually change HLB and packing parameters of detergent **1**, the head group was exchanged for building blocks of detergents that are also compatible with membrane proteins, such as mono- and tetraethylene glycols (E1 and E4), glycerol (G), and glucose (Glu) (**Figure 3.1**). The resulting set of detergents 1–5 was complemented with C8E4 and OG, both of which were previously employed for MP purification and native MS experiments[13,17,18,29].

Figure 3.2 shows the values for the HLB (**a**) and the packing parameter (**b**) for OG, C8E4 and detergents **1-5** (plotted in order of the increasing molecular weight (MW) – MWs of detergents are listed in **Table 3.1**). For the species with the same hydrophobic tail (which is true for detergents **1-5**), it is apparent that the smaller the size of the polar head group, the lower the HLB value of the detergent. Therefore, a general decrease in overall hydrophobocity (indicated by the increase in HLB values) is observed in **Figure 3.2a** (with a minor deviation due to the presence of two detergents with shorter tail lengths). An opposite trend is observed for the packing parameter (**Figure 3.2b**), which decreases with MW of the detergent, indicating a shift towards more conical shapes. Overall, the data presented in **Figure 3.2** demonstrate that the overall hydrophobicity and the nature of the shape of OG, C8E4 and detergents **1**-**5** gradually change in the opposite

directions, making this set a good starting point to investigate the mechanism of MP-delipidation by detergents.

Detergent	Molecular weight (g/mol)
OG	229.4
C8E4	306.5
1	408.5
2	480.6
3	510.7
4	612.8
5	642.8

Table 3.1 Molecular weights of detergents.

3.3 Phospholipid removal

The designed OGD proteins were applied to several MPs to estimate their ability to cause delipidation. C8E4 and OG were also used as controls, due to their known ability to effectively remove phospholipids (PLs)[13,18]. Three different MPs were selected, to test the applicability of the findings to proteins of different oligomeric states and MWs: trimeric ammonia channel, AmtB, from *Escherichia coli* (*E.coli*); tetrameric water channel Aquaporin Z (AqpZ) also from *E.coli* and dimeric translocator protein (18kDa) (TSPO) from *Rhodobacter sphaeroides*. Despite the previous report that highlighted the ability of detergent **1** to efficiently extract proteins from the native membranes[20], for the purposes of this work it was decided to follow a different procedure. The proteins were extracted and purified in DDM, which is often considered the 'golden-standard' detergent for this purpose[16]. A detergent exchange was then performed using a 3-ml size-exclusion chromatography (SEC) column, from DDM into detergent of interest at 2x critical micelle concentration (CMC) (see **Materials and Methods** for a more detailed protocol). This way, all of the delipidation experiments were performed on the batches that come from the same purification (and, therefore, the same initial lipid

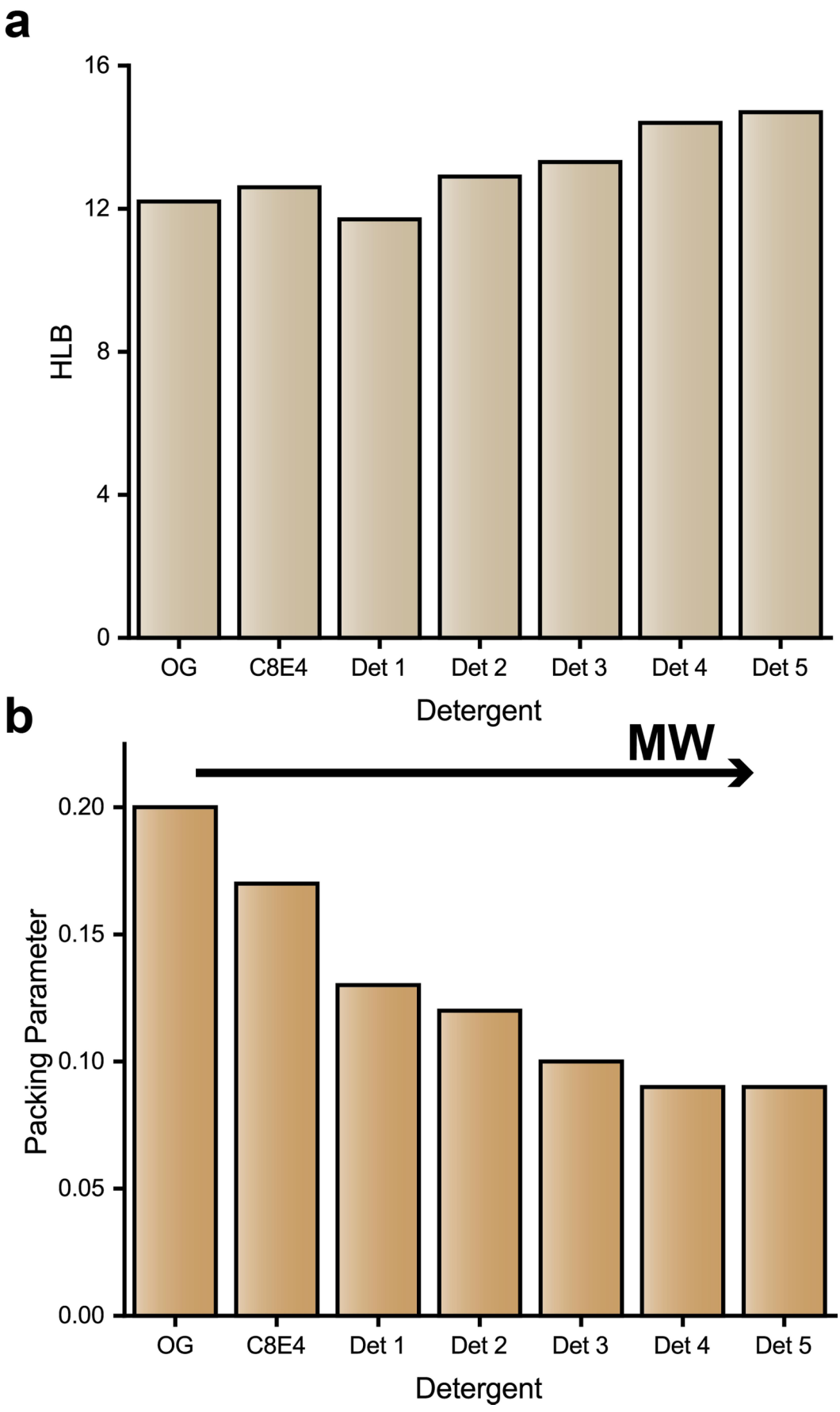


Figure 3.2 HLB and packing parameter. a, Hydrophobic-lipophilic balance **(**HLB) and **b**, packing parameter for each detergent. Detergents plotted in the order of increasing molecular weight (MW). Data provided by Leonhard Urner.

and protein concentrations). This enabled direct comparison of delipidation properties between detergents, without the complications of batch-to-batch variations in lipid concentration due to variable cell-growth conditions; or having to take into account the varying extraction efficiencies of different proteins. To further account for potential variations in the initial lipid concentrations between different purifications, data for AqpZ and AmtB were collected on three biological repeats.

Native MS experiments were performed on a Q Exactive UHMR instrument (which was described in the introductory chapter, **Figure 1.6d**). Removal of detergent adducts from MPs was achieved by increasing the higher-energy C-trap dissociation (HCD) voltage. Spectra of AmtB solubilised by OG, C8E4 and detergents **1**-**5** are shown in **Figure 3.3**. There are several important features which need to be discussed. Firstly, the most abundant charge state is not the same for the different detergents; the difference between OG and detergent **1** and the others is especially large. This happens due to variable charge reduction properties of the detergents used, which are discussed further in the following sections. Secondly, the HCD voltage required to achieve complete detergent removal varies with detergent. This is partly a consequence of the charge differences (as the energy experienced from an electric potential is directly proportional to charge) and partly a reflection of the different gas-phase stabilities of the detergent-protein complexes[13,29]. It should be noted that for all of the detergents investigated here the native trimeric state is the dominant species in the spectrum, implying that the energy required to strip off the micelle is sufficiently low to avoid causing significant structural rearrangement. (*Note: UMHR instruments are currently not compatible with ion mobility (IM). Therefore, the changes in the*

protein folding could not be monitored, so the oligomeric state was used to access stability instead.)

The considerations described above make direct comparison between the delipidating properties more challenging – if some MPs experience more energy in the gas phase than the others, they could lose more lipid as a consequence of that extra energy and not due to the delipidating properties of detergent in solution. The counterargument to that is the fact that no correlation between the average charge state and the degree of delipidation was observed (see **Section 3.5**), in line with previous results obtained for OGDs[20]. The results obtained for C8E4 support this counterargument: C8E4 exhibits strong charge reducing properties and is readily removable at gentle activating conditions, yet no lipid-binding was observed. It should be noted, that for every detergent-protein combination I have tested a range of collision energies, aiming for the value just sufficient to remove the micelle. While this minimal energy is not the same for different detergents, it is possible that the lipid binding is preserved until this stage, with most of the energy being dissipated by the departing detergent molecules – such collisional cooling effects were previously observed for other detergent-MP complexes[30]. Considering the evidence, I conclude that the amount of lipid binding observed is not a gas-phase artefact, and is instead reflective of the delipidating properties of detergents.

It can be seen from the data in **Figure 3.3**, that larger proportions of lipid-bound states of AmtB are detected for detergents with higher MWs. In order to observe this pattern more clearly, the relative amount of *apo* and bound species (calculated as a fraction of all the peaks present in the spectra) was plotted for each detergent (**Figure 3.4a**). Similar plots were also produced for AqpZ and TSPO (**Figures 3.4b** and **c** respectively). All of the ligands present were counted towards the ‘bound’

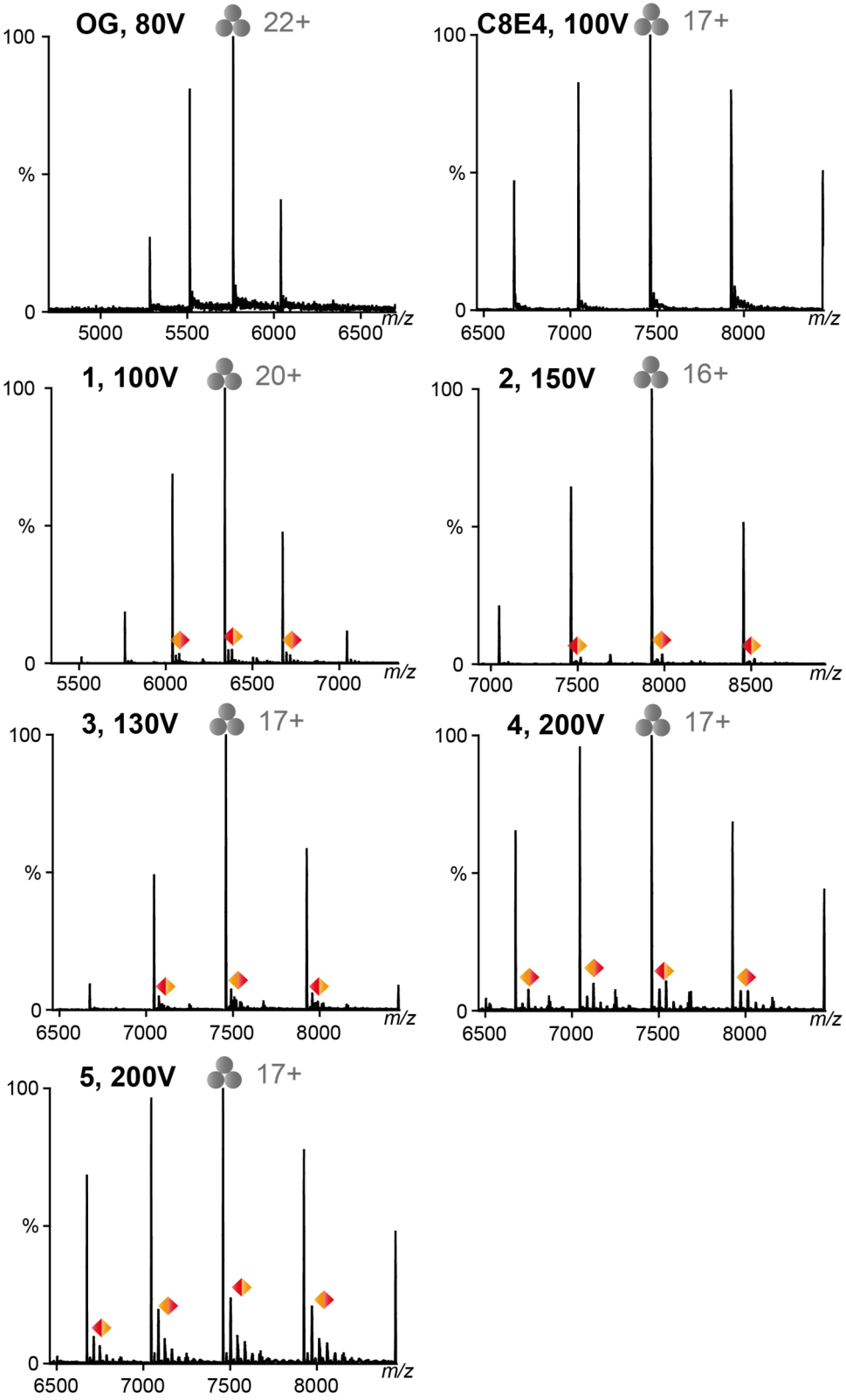


Figure 3.3 AmtB spectra. AmtB trimer (grey), bound to phospholipids of unknown nature (orange/red symbols). The varied amount of delipidation is clearly observable. The voltage in the label refers to the HCD voltage used to produce these spectra. Note the different *m/z* scales. Spectra are zoomed in for clarity, not all charge states are shown. The small unlabelled peak at a higher mass than the phospholipids most likely represents a small amount of LPS binding (based on mass analysis), with a minor contribution of AmtB dimer peaks in some cases.

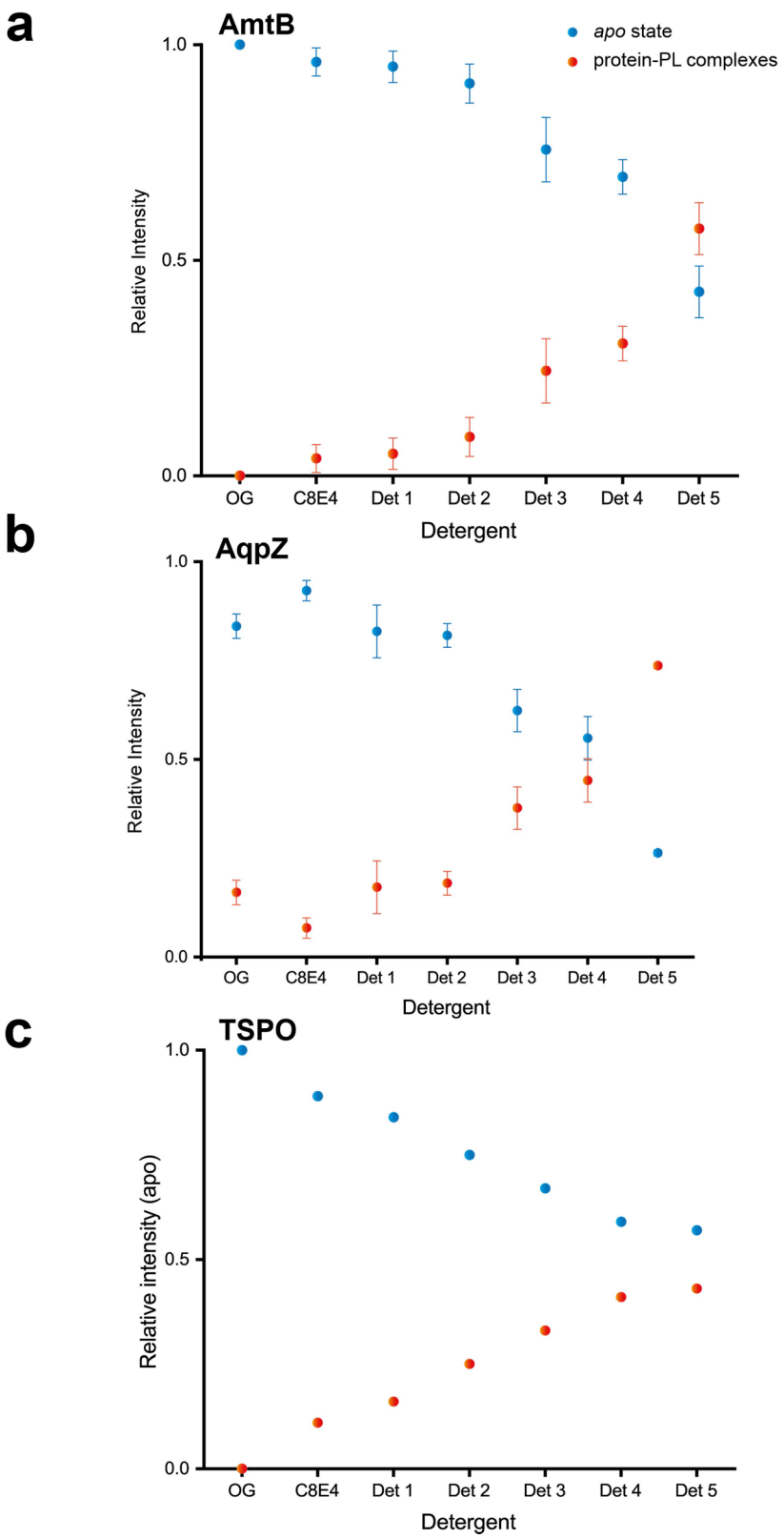


Figure 3.4 Phospholipid removal by detergents. The relative intensities of *apo* and lipid bound states for **a**, AmtB; **b**, AqpZ; and **c**, TSPO. Intensities are calculated relative to the total intensity of all of the peaks in the spectrum. For panels **a** and **b**, values shown are the average of three biological repeats, with error bars representing one standard error of the mean; panel **c** shows the data from a single repeat.

protein intensity (apart from a small peak at a significantly higher mass, which is tentatively assigned to a minor amount of lipopolysaccharide (LPS) binding to the protein based on mass difference, with a very minor contribution of dimeric AmtB; see the next section for more information on LPS). Based on the masses, the ligands were assigned to various PL classes present in *E.coli* cells[3] (which were used for protein expression): namely, phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CDL) or the mixtures of those. It is important to note, that the adduct masses observed were significantly higher (700 Da or larger) than the masses of detergents used. A detailed characterisation of the bound lipids is challenging[16,17,31] and was beyond the scope of this work.

Similar delipidation patterns were observed for all of the proteins used in this study (**Figure 3.4**). This indicates that the delipidation properties of detergents are likely to be transferrable to a wide range of PL-binding MPs. The fact that the intensity of the lipid-bound states changes gradually and predictably between OG, C8E4 and detergents **1**-**5** presents a new opportunity for MP-lipid investigations. Instead of relying on a lengthy 'trial-and-error' screening of detergents and conditions in order to achieve the desired level of delipidation, the amount of co-purified endogenous lipids can be controlled simply by choosing an appropriate detergent from the set shown in **Figure 3.1** during the final stage of the standardised purification protocol (described in **Materials and Methods**). Previously reported mass spectra of MPs obtained upon purification with mildly delipidating OGDs led to noisy spectra[20]. Therefore, the development of mildly delipidating detergents **3**–**5**, which also enable the acquisition of well-resolved and clean MP mass spectra, represents a significant improvement.

Interestingly, no clear correlation is observed between delipidating abilities of OG, C8E4 and detergents **1**-**5**, and the chemical natures of their polar head groups. The most striking example of this observation is detergent **5**, which has a head group that contains E4 and Glu – which are the head groups of C8E4 and OG respectively (**Figure 3.1)**. Unlike these two strongly delipidating detergents, the hybrid detergent **5** was able to preserve many MP-lipid contacts under comparable conditions (Figures **3.3** and **3.4**). Similar observations were made for detergents **3** and **4**, which were also significantly less delipidating than structurally related C8E4 and detergent **1**. Instead of the exact chemical nature of the head group, MP delipidation efficiency appears to correlate with more general structural parameters, such as the HLB and packing parameters (compare **Figures 3.2** and **3.4**). It is possible that both of these parameters contribute to the lipid-displacing ability of a particular detergent. Detergents with lower HLB values are more hydrophobic, and, therefore, are more likely to outcompete lipids bound to MPs. Detergents with lower packing parameters have less conical shapes and form micelles with lower packing densities[24] – meaning that more free space is available inside the micelle to be occupied by lipids. Therefore, the results obtained from delipidation experiments provided a more fundamental understanding of MP delipidation by detergents. The HLB and packing parameter concepts are applicable to a broad range of detergents and may enable the development of general detergent design guidelines in the future.

3.4 Lipopolysaccharide removal

(*Note: some data are shown without repeats, as the experiments were disrupted by the pandemic*) LPS is a glycolipid which makes up the outer leaflet of the outer

membrane of many Gram-negative bacteria, including *E.coli*[32]. Structure of LPS was described in **Chapter 1** (**Section 1.2**, **Figure 1.2d**). Briefly, the essential core LPS consists of Lipid A attached to two 3-deoxy-D-*manno*-oct-2-ulosonic acid (KDO) units – with optional modification by multiple diverse sugars[33,34]. This results in heterogeneity that can produce complicated native mass spectra of LPS-MP complexes, with broad peaks which can obscure other binding events[35]. Moreover, LPS is notoriously difficult to remove either in the gas phase or in solution, with delipidation previously achieved by changing detergent during the extraction step[36]. This strategy is not always applicable, because many MPs have instability issues and can require following highly specific expression and purification protocols[37,38]. For that reason, the ability of detergents presented here to remove LPS was tested.

Mechanosensitive channel of large conductance (MscL) from *E.coli* was used as a model protein for the LPS-removal experiments, as previous native MS experiments revealed that MscL binds LPS when overexpressed in *E.coli* cells[36]. While LPS is located in the outer membrane of bacteria, it is synthesised in the inner membrane and then transported across with assistance of various proteins[39-42]; therefore, interactions of LPS with the inner membrane MscL are not completely surprising. To attempt LPS removal, the same delipidation strategy that was used for PLs in the previous section (2xCMC of detergent introduced during a SEC stage) was also employed here. However, relatively intense LPS-bound peaks were detected, even for a very PL-delipidating detergent **1** (**Figure 3.5a**). Importantly, no clear delipidation pattern was produced this time for the detergents, unlike the PL case (**Figure 3.5b**). This observation implies that no significant LPS removal occurs during the detergent exchange step. The reduced ability of the

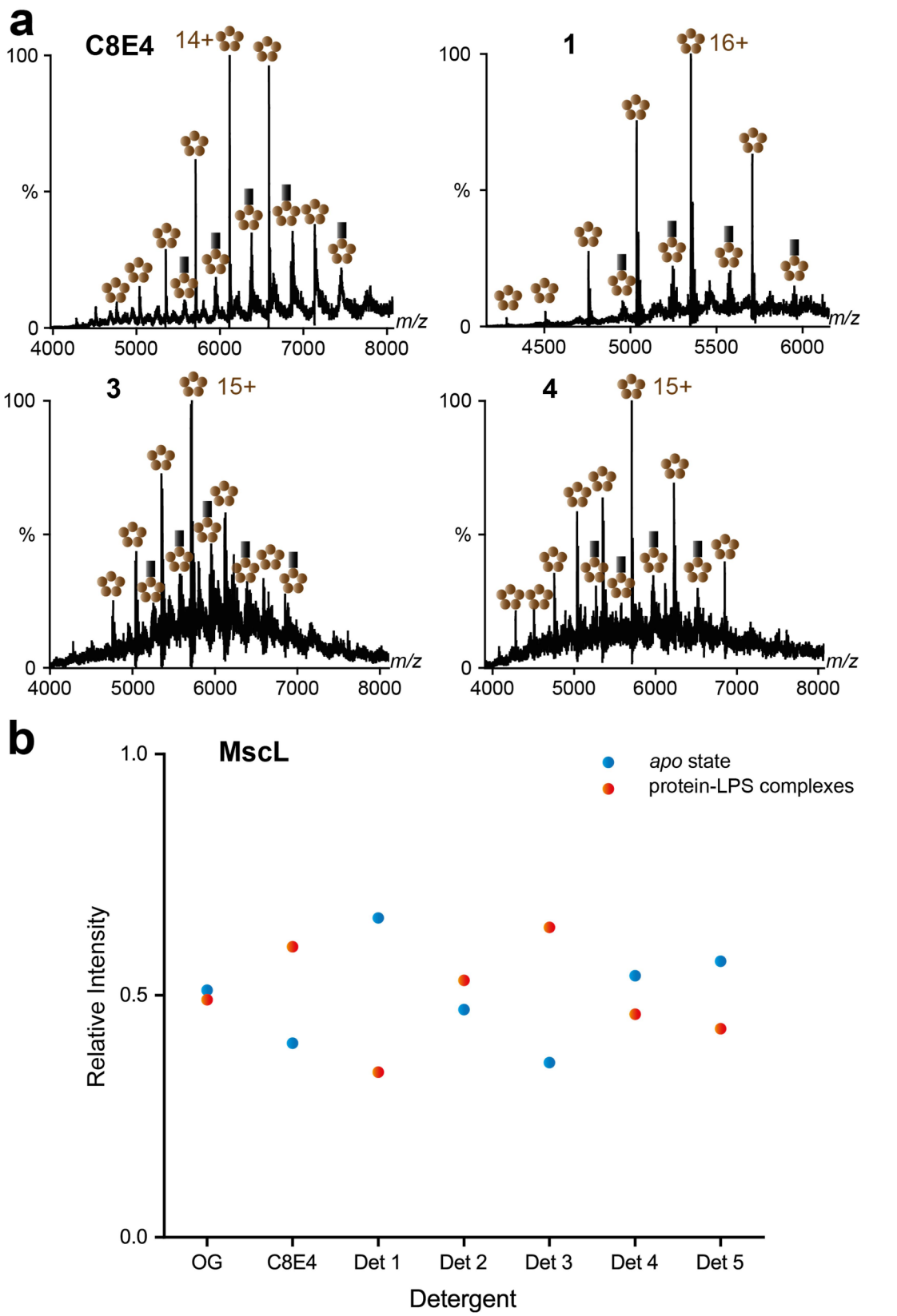


Figure 3.5 Lipopolysaccharide removal by detergents. a, Representative spectra of MscL (brown) in different detergents. LPS-bound peaks (black) are present for all detergents. HCD voltage of 200 V was used in all of the spectra. **b,** The relative intensities of *apo* and lipid bound states for MscL. Intensities are calculated relative to the total intensity of all of the peaks in the spectrum.

PL-delipidation method to remove LPS is likely to be due to the large hydrophobic surface of this glycolipid, as well as its potential to form multiple electrostatic interactions with the protein, resulting in a very high dissociation energy.

An alternative strategy was devised, involving longer exposure of LPS-bound proteins to higher detergent concentrations. A natural way to practically implement this strategy is to introduce detergent at 1% w/v during immobilised metal affinity chromatography (IMAC) purification stage (see **Materials and Methods**). While the protein remains bound to the column, several washing steps with detergent-containing solution can be performed. The number of these washings can be changed to vary the extent of exposure of the protein to high concentrations of detergent.

Detergent **1** was selected for the LPS-delipidation experiments because of its high propensity for delipidation of PL-MP complexes (**Figure 3.4**). Two different model proteins that were known to bind LPS in the gas phase were chosen: the multidrug efflux pump subunit AcrB from *E.coli*, which is the inner-membrane part of the TolC multidrug efflux pump[43]; and the ATP-binding cassette (ABC)-transporter BtuCD from *E.coli*, which is a vitamin B12 importer[35]. First, the LPS-delipidation procedure was applied to AcrB, with between 0 and 40 column volumes (CVs) of detergent **1** at 1% w/v (50xCMC) applied during the IMAC stage as was described above (*Note: the 0 CVs sample also has detergent **1**, which is present in the MS buffer at 2xCMC for all samples; see **Materials and Methods** for a detailed protocol*). Spectra obtained are shown in **Figure 3.6**, for 0 CVs (**a**) and 40 CVs (**b**) of washing with detergent **1**. No resolved peaks were observed when no washing was applied (**Figure 3.6a**), indicating a large number of adducts. Similar unresolved spectra were obtained when the PL-delipidation protocol described in **Section 3.3** was

employed with detergent **1** (data not shown). However, performing an extensive 40 CVs wash with 50xCMC of detergent **1** resulted in well-resolved spectra of trimeric AcrB, with only nominal amount of LPS-bound states present (**Figure 3.6b**). In order to optimise conditions for LPS delipidation, the experiment was repeated at several intermediate CV values (**Figure 3.6c**). A gradual removal of LPS was observed, with the maximum delipidation achieved at around 20 CVs. Importantly, some LPS-MP interactions can be preserved and resolved by native MS by choosing an intermediate number of washes.

Next, the LPS delipidation procedure used for AcrB was tested for BtuCD (**Figure 3.7**). A problem was encountered, as BtuCD complex was found to be dissociating at conditions required to remove the detergent micelle, making the spectra more challenging to analyse (**Figure 3.7a**). This instability was hypothesised to be a result of a relatively high charge of the protein complex, which arises as a consequence of the non-charge reducing nature of detergent **1** (see the next section for further discussion of charge reduction). High charge can cause significant protein instability due to electrostatic repulsion, so low charge states of proteins are generally preferred for native MS[44]. One way to cause charge reduction without introducing any chemical additives to the sample is by switching the nanoelectrospray ionisation (nESI) polarity from positive to negative – this effect was demonstrated to be significant for MPs solubilised in saccharide detergents which are not charge reducing in positive mode, such as DDM and OG[45]. A similar charge reducing effect upon switching the polarity was, therefore, anticipated to be observed for detergent **1.**

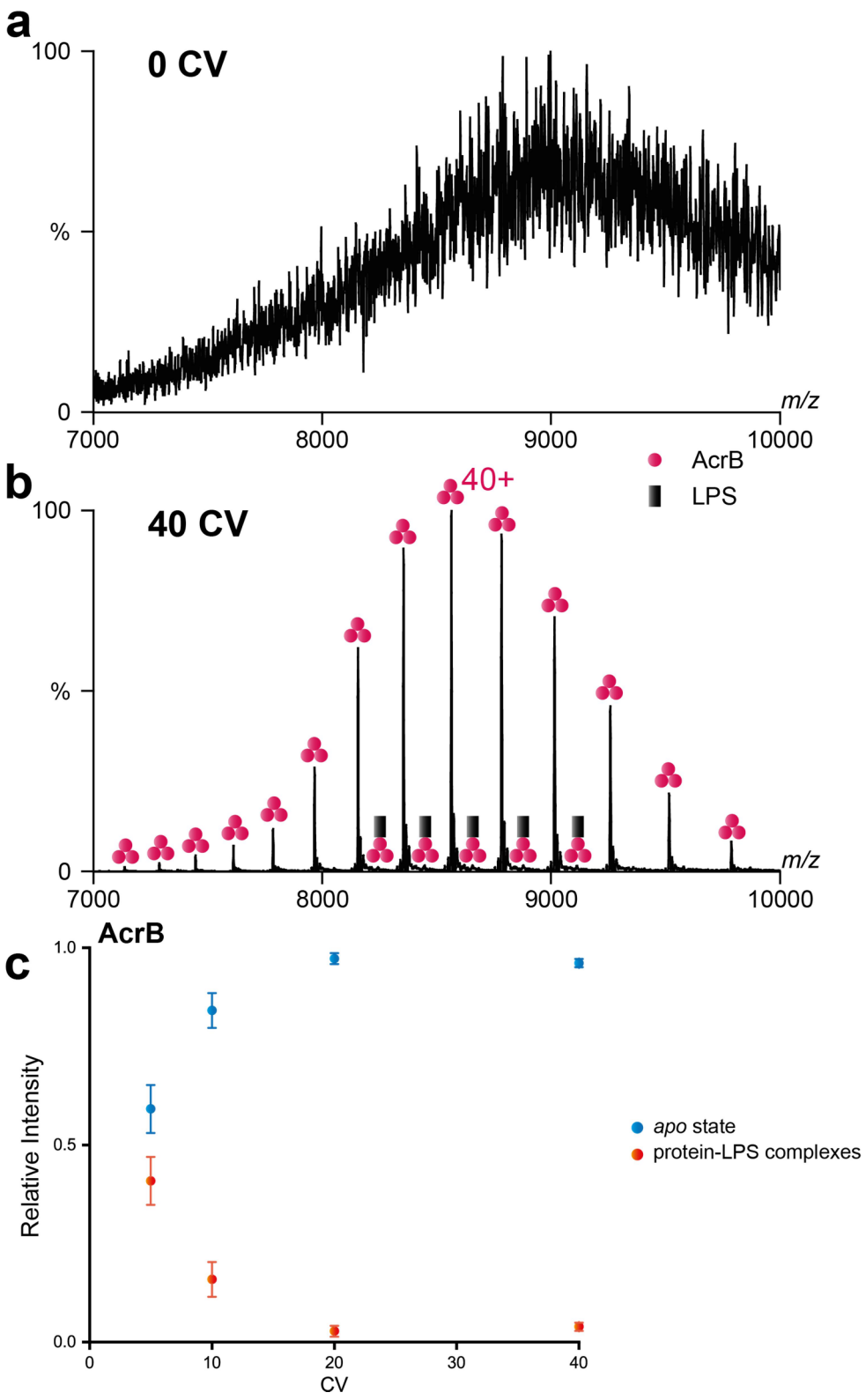


Figure 3.6 Removal of LPS from AcrB by detergent 1. a, Spectrum of AcrB delipidated with 0 CVs of detergent **1**. **b,** Spectrum of AcrB delipidated with 40 CVs of detergent **1**. Spectra in panels **a** and **b** were acquired using the same conditions (300 V HCD energy). **c,** The relative intensities of *apo* and lipid bound states for AcrB. Intensities are calculated relative to the total intensity of all of the peaks in the spectrum. Values shown are the average of two biological repeats, with error bars representing one standard error of the mean. At 0 CVs and 1 CVs no resolved peaks were observed. Abbreviations: CV, column volume; LPS, lipopolysaccharide.

Charge reduction occurs as expected for BtuCD in negative mode with exclusively intact complexes observed for 0 CVs (**Figure 3.7b**) and 40 CVs (**Figure 3.7c**) washed samples. Two major differences are apparent between the BtuCD data and the AcrB data presented earlier in this section. Firstly, even at 0 CVs BtuCD resolved peaks are observed in the mass spectrum (**Figure 3.7b**) (although heavily LPS-bound), while for AcrB only an unresolved 'hump' was visible (**Figure 3.6a**). Secondly, at the highest number of washings (40 CVs) AcrB was almost completely delipidated (**Figure 3.6b**), whereas for BtuCD a small, but significant amount of LPS-bound complex is visible (**Figure 3.7c**). The incomplete delipidation of BtuCD is even more evident in the data shown in **Figure 3.7d**, with a 'plateau' seemingly being reached at around 15% lipidated protein remaining. Taken together, these two observations could be interpreted as the LPS binding being non-specific for AcrB (because practically complete delipidation is achieved, despite a high pre-delipidation amount of LPS) and specific for BtuCD (by the opposite reasoning). In order to test this interpretation, BtuCD was investigated further by complementary methods.

So far, the LPS identity was assumed from the previously reported results of LPS binding being observed for the proteins studied here[35,36], as well as from the mass of the adducts matching the expected mass for bacterial LPS (roughly 3.6-3.7 kDa)[36]. The presence of LPS in solution can be verified by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by silver staining[46]. This method was applied to both AcrB and BtuCD at different stages of LPS-delipidation protocol (**Figure 3.8**). The presence of LPS is indeed detected for both proteins at low delipidation. Interestingly, no LPS is detected by the silver staining above 10 CVs for AcrB and above 0 CVs for BtuCD, despite LPS binding

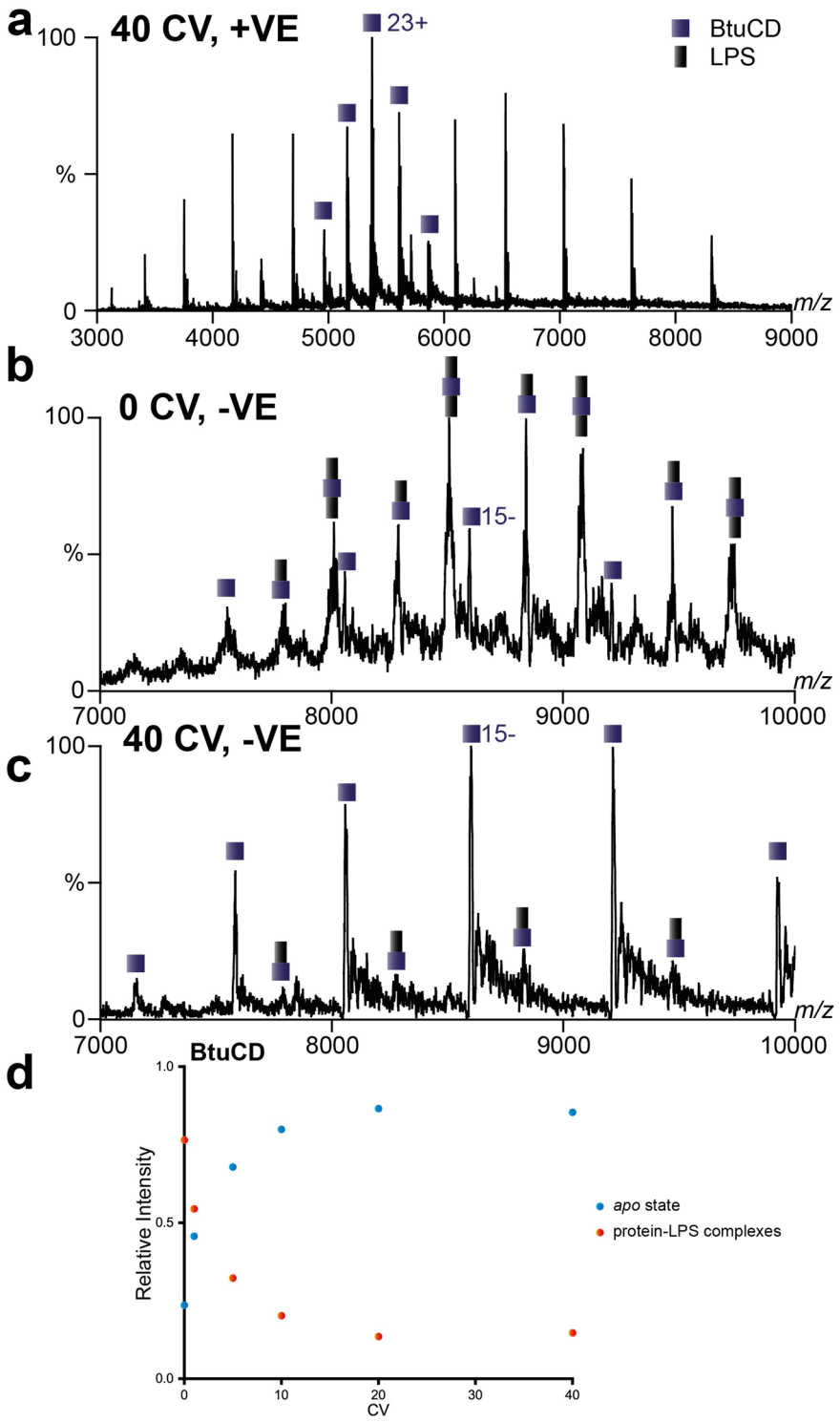


Figure 3.7 Removal of LPS from BtuCD by detergent 1. a, Spectrum of BtuCD in positive (+ve) mode delipidated with 40 CVs of detergent **1**. Significant amount of dissociation occurs, resulting in the presence of various complexes and LPS-bound states (not labelled). Spectrum acquired using 100 V in-source trapping (IST) and 100 V HCD. **b,** Spectrum of BtuCD in negative (-ve) mode delipidated with 0 CVs of detergent **1**. Spectrum acquired using 300 V IST, 300 V HCD. **c,** Spectrum of BtuCD in negative (-ve) mode delipidated with 40 CVs of detergent **1**. Spectrum acquired using 300 V IST, 300 V HCD. **d**, The relative intensities of *apo* and lipid bound states for BtuCD. Intensities are calculated relative to the total intensity of all of the peaks in the spectrum. Abbreviations: CV, column volume; LPS, lipopolysaccharide.

being clearly observed by native MS (**Figures 3.6c** and **3.7d**). This result highlights the utility of native MS for detection of MP-lipid binding interactions, even at relatively low concentrations of lipids.

After confirming the presence of LPS in BtuCD sample, the role of LPS binding on BtuCD function was investigated. In order to achieve this goal, BtuCD was expressed from the ClearColi - a strain of *E.coli* which produces no LPS. (*Note: BtuCD expression from E.coli and the activity assay were performed by Leonhard Urner and Francesco Fiorentino)*. This way, a completely delipidated state of BtuCD was observed by native MS (**Figure 3.9a**). To investigate whether LPS is essential for maintaining the secondary structure of BtuCD in solution, circular dichroism (CD) spectroscopy[47] experiments were conducted – no difference in secondary structure between the ClearColi BtuCD and the standard LPS-bound BtuCD was observed (**Figure 3.9b**). Finally, ATPase activity assay was conducted on the *apo* BtuCD solubilised in detergent **1**, and then repeated upon addition of various lipids (**Figure 3.9c**). No statistically significant change in activity (compared to the *apo* BtuCD) was observed upon binding of PE, CDL or lipid A (which is a component of LPS, see **Figure 1.2d**). However, when KDO_2-lipid A (KLA) was introduced into the sample, a small but significant increase in BtuCD activity of around 8% was detected. Interestingly, a similar increase in activity was caused by PG PL. A common feature of these two lipids, which is not present for the other species in **Figure 3.9c**, is the simultaneous presence of negative charges in the lipid backbone and terminal 1,2-diols in the head group (see **Figure 1.2** in Chapter **1** for the lipid structures). The effect of specific lipids on BtuCD activity is extremely intriguing and is being investigated further by molecular dynamics (MD) simulations

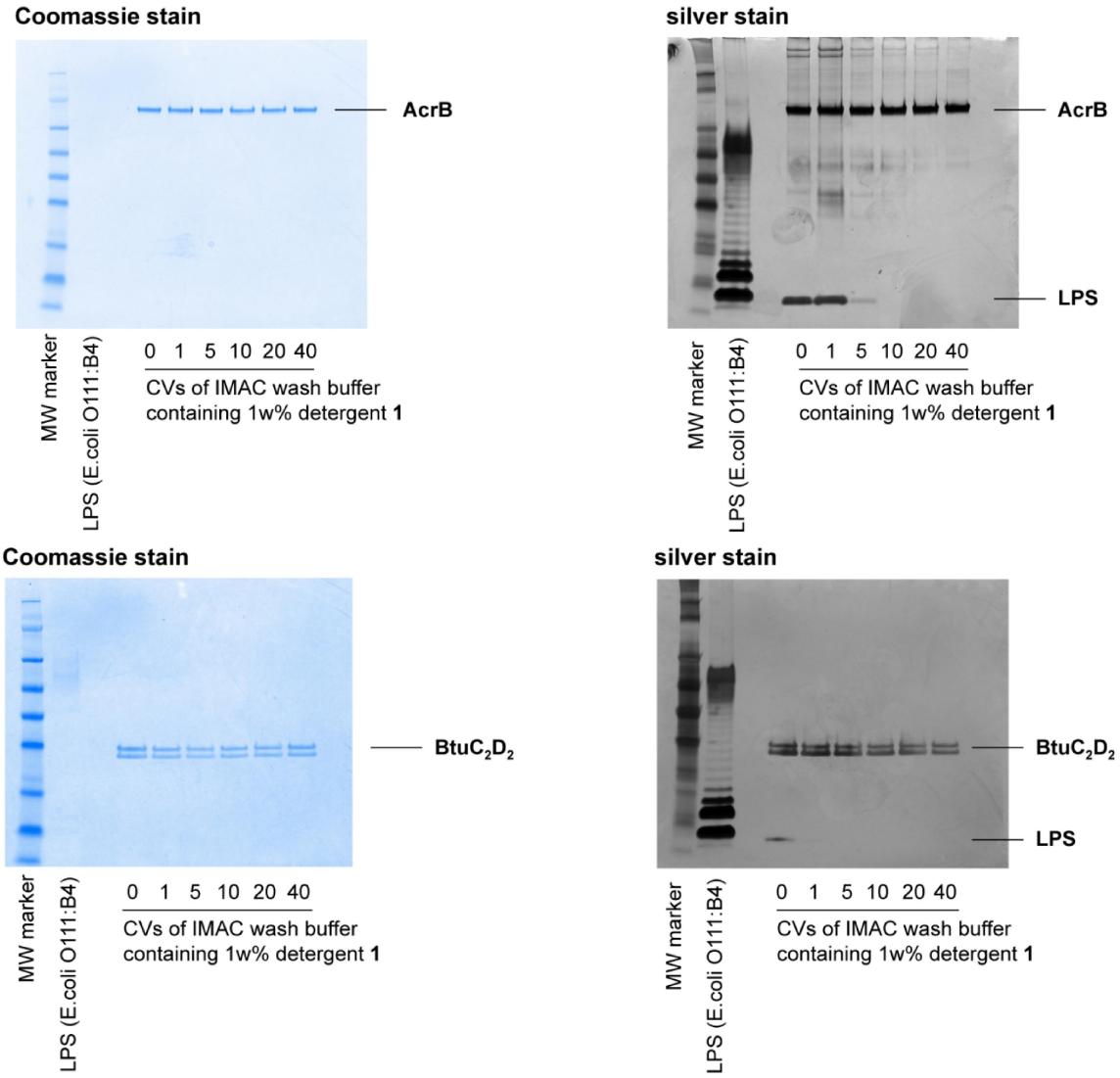


Figure 3.8 SDS-PAGE of AcrB and BtuCD. Coomassie stain (left) sensitive to proteins and silver stain (right) sensitive to LPS of SDS-PAGE of AcrB (top) and BtuCD (bottom). Figure made by Leonhard Urner. Abbreviations: SDS-PAGE, sodium dodecyl sulphate-polyacrylamide gel electrophoresis; LPS, lipopolysaccharide.

(subject of ongoing research). For the purposes of current study on delipidating properties of detergents, the activity increase upon KLA binding to BtuCD implies the presence of specific interactions between LPS and this MP. Therefore, it appears that the LPS-delipidation protocol with detergent **1** can be employed to distinguish between non-specific LPS binding, where complete delipidation occurs as was observed for AcrB (**Figure 3.6c**); and specific LPS binding which is preserved even after extensive washing procedure at 40 CVs, as was observed for

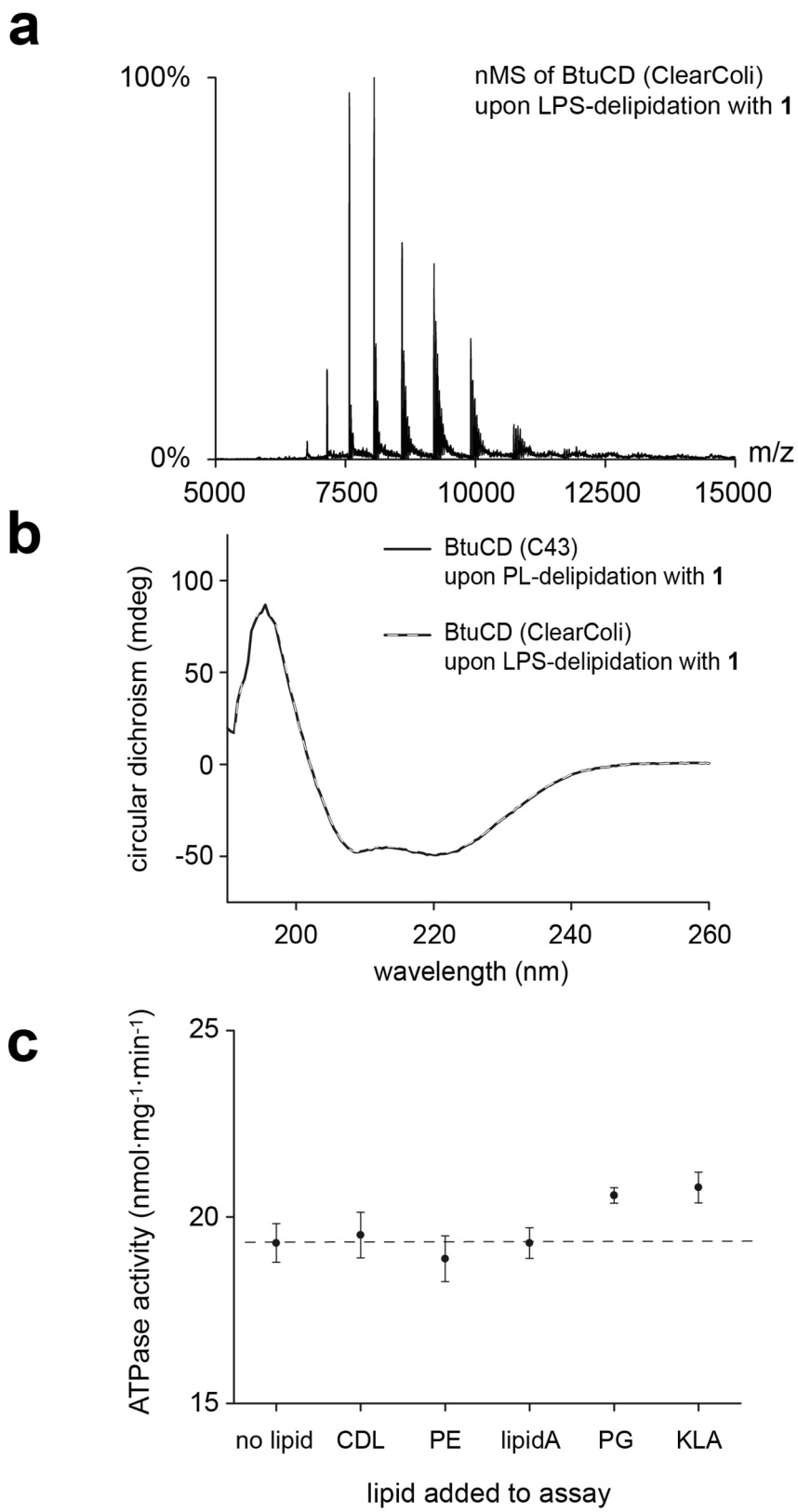


Figure 3.9 Effect of LPS binding on BtuCD. a, BtuCD expressed from ClearColi cells and delipidated with 40CVs of detergent **1**. **b,** Circular dichroism experiments performed on the LPS-bound and *apo* BtuCD revealed no structural differences in solution. **c,** ATPase activity performed on BtuCD expressed from ClearColi, with lipids added exogenously. Values shown are the average of three repeats, with error bars representing one standard error of the mean. Data collected and figure made by Leonhard Urner. LPS, lipopolysaccharide.

BtuCD (**Figure 3.7d**). While drawing definite conclusions from the individual proof-of-principle examples described here is premature, these initial observations are promising.

3.5 Charge Reducing Properties of Detergents

To finalise the characterisation of detergents for native MS, their charge reducing properties were investigated. As was mentioned above, charge reduction is beneficial to MP stability because coulombic repulsion promotes unfolding and even dissociation of protein complexes[13,44] (as was demonstrated here for BtuCD, see **Figures 3.7a** and **3.7c**). In addition, the increased spacing between two adjacent protein peaks in mass spectra of charge reduced MPs facilitates resolution of individual lipid binding events[48].

Detergents dissociate in gas phase as neutral species, so the charge reduction occurs during the nESI process. The proposed mechanism involves detergent molecules sequestering charge from solution and thus reducing available charge for the eventually formed desolvated protein ions[29] (see **Section 1.8** for further discussion). Consequently, a potent charge reducing detergent should have two properties: the ability to carry charge and high hydrophobicity (to be located closer to the surface of a nESI droplet for a higher probability of escape). This can be exemplified by C8E4, which is very charge reducing and has a head group based on polyethylene glycole (PEG) polymer that is known to solvate charges[49]; however without the hydrophobic tail the charge reduction efficiency is significantly lower[29].

Previous studies demonstrated the tunability of charge reducing properties of OGDS by changing the linker between the hydrophobic tail and the polar head group – linkers with higher gas phase basicities were better charge carries (in

positive ionisation mode) and, therefore, resulted in more charge-reduced proteins[20,28]. However, detergents **1**-**5**, OG and C8E4 used in this work all have the same ether linker (**Figure 3.1**). Therefore, the effects of the different head groups of detergents on charge reduction of MPs were investigated here.

As was mentioned above, C8E4 detergent is known to cause charge reduction of MPs; conversely, saccharide detergents, such as OG, are considered to be non-reducing[29]. To investigate whether the presence of E4 head group is determining the amount of charge reduction of MPs by detergents, relative weight percentages of E4 in detergent structure were calculated (**Table 3.2**).

Detergent	Relative % E4
OG	0%
C8E4	62%
1	0%
2	38%
3	40%
4	63%
5	30%

Table 3.2 Relative weight percentages of E4 in detergent structures.

The detergents can be roughly assigned into three categories: 1 - 0% E4 (OG, detergent **1**); 2 – 30-40% E4 (detergents **2**, **3** and **5**); 3 – 60+% E4 (C8E4 and detergent **4**).

The average charge states were calculated at conditions where complete removal of detergent molecules was achieved. While detergents dissociate from MPs as neutral species[29] and thus do not affect the observed charge states directly, a larger number of bound detergent molecules typically remain for the lower charge state of the protein (because activation energy at a given voltage is proportional to charge), meaning that they can be unresolved in the spectra, thus skewing the observed distribution.

The average charge states for AmtB, AqpZ, TSPO and MscL in the presence of detergents from those three groups are shown in **Figure 3.10**. Higher charge states for detergents with 0% E4 (OG and detergent **1**) compared to the other detergents are apparent, while the differences between the intermediate and the high C8E4 percentages are less obvious. Interestingly, within each group more hydrophobic (lower HLB) detergents are consistently more charge reducing than the less hydrophobic ones (a single exception is the order for C8E4 and detergent **4** in the 60+ group of TSPO, but the values only differ by 0.2 of a charge). Together, these observations corroborate the previously proposed charge reduction mechanism[29] described above, where detergents sequester the charge inside nESI droplet and escape them assisted by hydrophobicity.

From a practical point of view, the detergents in **Figure 3.1** can be separated into three broad groups in terms of their charge reducing abilities: no/low charge reduction: OG and detergent **1**; high charge reduction: C8E4 and detergents **2**, **3** and **4**; and intermediate charge reduction for detergent **5** (the dark-grey data point for the 30-40% groups in **Figure 3.10**). Importantly, as I have already mentioned in **Section 3.3**, no obvious correlation exists between charge reduction and the degree of PL delipidation (illustrated for AmtB in **Figure 3.11**); therefore, charge reduction only needs to be considered in terms of MP complex stability when selecting the detergent for a delipidation experiment.

3.6 Summary and Conclusions

In this chapter I have developed and demonstrated a novel method for achieving controlled delipidation of MPs for a subsequent characterisation by native MS. The method includes a set of detergents (**Figure 3.1**), four of which (detergents **2**-**5**)

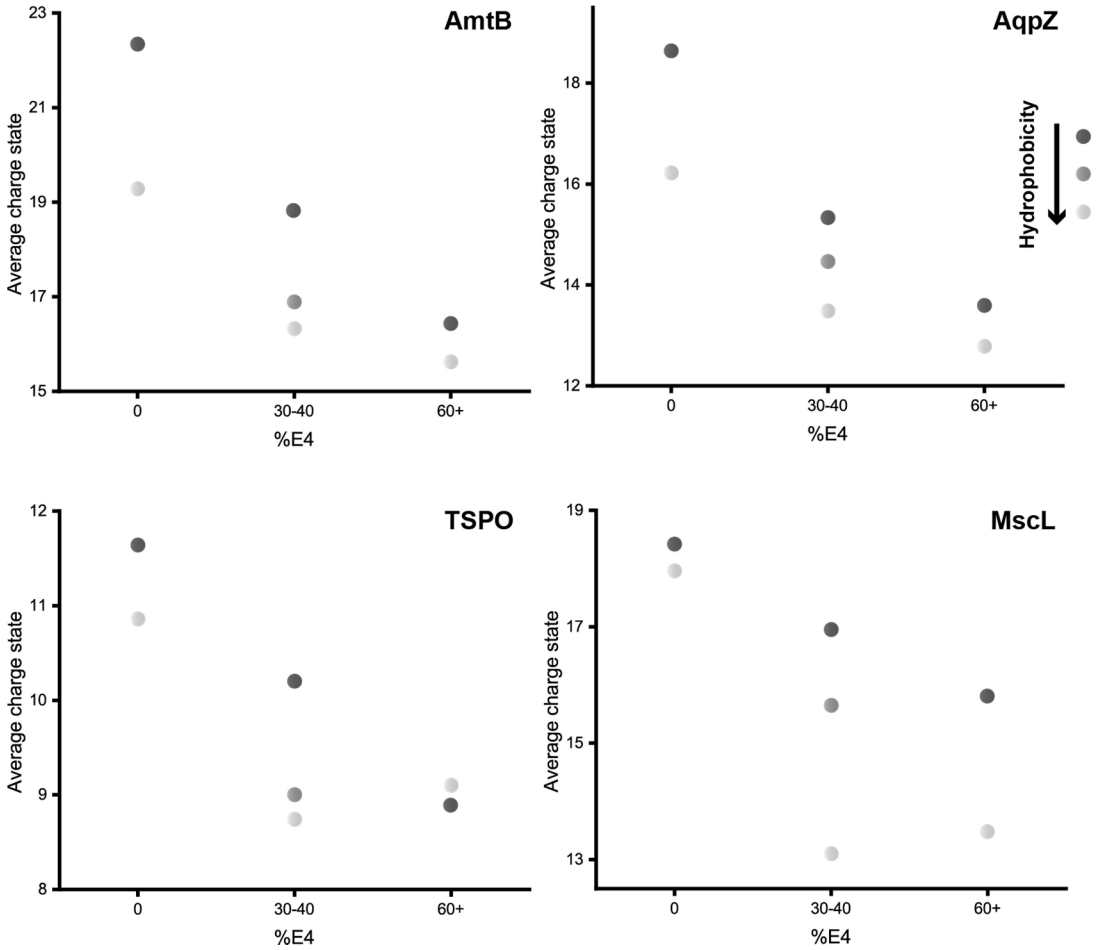


Figure 3.10 Average charge state of proteins in different detergents. Proteins are grouped by the relative weight percentage of tetraethylene glycol (%E4). Average charge states are calculated relative to the total intensity of all of the protein peaks. Hydrophobicity is indicated by HLB (lower HLB value means higher hydrophobicity), the colours should only be compared within one group. HLB, hydrophobic-lipophilic balance.

have not been presented before, as well a standardised purification protocol. The protocol is summarised below (see **Materials and Methods** for more details):

For PL delipidation:

1. Extract and purify the MP in DDM detergent.

2. Perform a detergent exchange over a 3-ml SEC column into MS-compatible buffer containing 2xCMC of one of the detergents shown in **Figure 3.1** (OG, C8E4 or detergents **1**-**5**). If the maximum delipidation is desired, select the leftmost

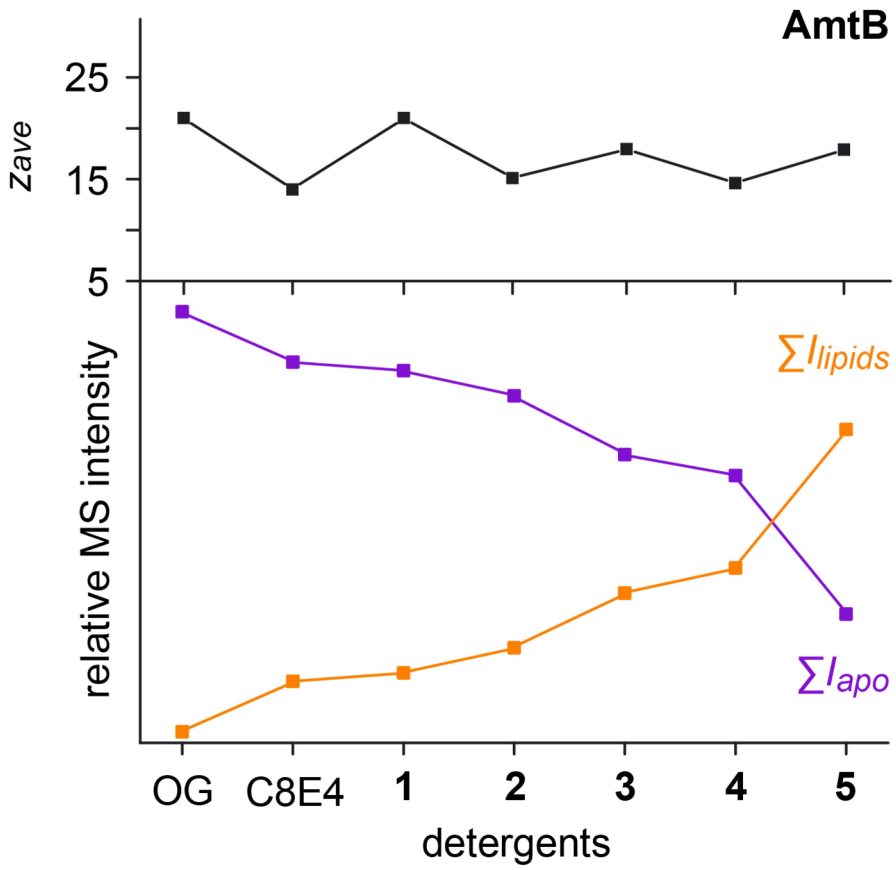


Figure 3.11 Average charge state and delipidation of AmtB in different detergents. No obvious correlation is observed between the average charge state (Zave) of MP and the degree of delipidation caused by the detergent. Figure made by Leonhard Urner.

detergent (OG); each subsequent detergent in the set is gradually less delipidating (but still produces well-resolved native mass spectra).

3. Perform native MS experiments using activation energies which are just sufficient to completely remove any bound detergent molecules.

For LPS delipidation:

1. Extract the MP in DDM detergent.

2. During the IMAC purification stage, use a washing buffer containing detergent **1** at 1% w/v (50xCMC). Use 20 CVs of this buffer for the maximum* removal of LPS, or 5-10 CVs to retain more LPS-bound states.

3. Exchange the protein into MS-compatible buffer (see **Materials and Methods**) containing detergent **1** at 2xCMC.

4. If the protein is unstable during native MS, switching into negative polarity can alleviate the problem.

*complete removal may not be possible for some MPs, potentially indicating specific MP-LPS interactions.

These protocols were demonstrated to produce predictable delipidation for MPs with a range of MWs and oligomeric states, indicating their potential wide-spread applicability. Importantly, the expected oligomeric states of MPs have been preserved by all of the detergents in **Figure 3.1**, reflecting stabilities of MP complexes when solubilised by these detergents.

In addition to the practical applications for a particular set of detergents, I have investigated detergent parameters responsible for MP delipidation more generally. The increase in the degree of PL delipidation was closely matched by a decrease in HLB value of a detergent (reflecting an increase in hydrophobicity) and an increase in packing parameter (corresponding to greater packing density of the micelle). The degree of delipidation was also found to be unrelated to chemical structures of detergent head groups, implying potential general applicability to detergent design. Lastly, charge reducing properties of detergents were shown to be consistent with the currently accepted model of charge reduction.

Taken together, the results presented in this chapter enable the beginning of a transition from the current 'trial-and-error' use of detergents for MP solubilisation and towards a more rational and controlled approach.

Materials and Methods

Packing parameter calculations. In order to calculate the packing parameter of C8E4, OG, **1** – **5** according to equation (2), one needs to calculate the volume of the hydrophobic tail (V_{tail}), the length of the hydrophobic tail (l_{tail}), and the area of the head group that is occupied at the interface between the detergent aggregate and solvent (A_{head}) (equation 3.2). To determine A_{head}, the three-dimensional structure of detergent head groups (without tail) was modelled using a MM2 force field as implemented in the software ChemBio3D v14.0 (PerkinElmer) and calculated their collision cross sections (CCSs) using a projection approximation algorithm[50]. The CCS of a detergent head group was taken as A_{head}. V_{tail} was calculated by using the van der Waals volume calculation method[51] and l_{tail} with ChemBio3D v14.0 (PerkinElmer).

Protein expression and membrane preparation. Membrane protein expression was performed as previously described[13,20]. Plasmids were transformed into C43 (DE3) cells by mixing 1 µL of plasmid solution (plasmid concentration = 100 ng/µL) with a 50 µL aliquot of C43 cells (purchased from Cambridge Bioscience). The cells were incubated on ice for 30 min, heat shocked at 42 °C for 45 s, and cooled on ice for 2 min. LB Broth (450 µL of a 25 g/L aqueous solution) was added. The mixture was shaken at 37 °C (180 rpm) for 1 h. One 50 µL aliquot of this mixture was plated on an agar plate (agar medium composition: 25 g/L LB Broth and 15 g/L agar in water, supplemented with 100 µg/mL ampicillin). The plate was stored

overnight at 37 °C. Up to five colonies were picked and transferred into starter culture medium (5 mL of 25 g/L LB Broth, supplemented with 100 µg/mL of ampicillin). The mixture was shaken with 180 rpm at 37 °C for 8 h. The starter culture was transferred into overnight culture medium (400 mL of 25 g/L LB Broth, supplemented with 100 µg/mL of ampicillin). The mixture was shaken at 37 °C (180 rpm) for 15 hours. The overnight culture was transferred into in 12 L expression medium (12 × 1 L of 25 g/L LB Broth, supplemented with 100 µg/mL of ampicillin). Cells were shaken with 180 rpm at 37 °C until an optical density value at 600 nm between 0.7 and 1.0 was reached. Protein expression was induced by adding isopropyl-ß-D-thiogalactopyranoside (IPTG) (12 × 1 mL of a 0.5 M aqueous solution) and the cells were shaken at 37 °C (180 rpm) for another 4 hours. Cells from a 12 L expression batch were harvested by centrifugation (5,000 × g, 10 min), suspended in 100 mL buffer (20 mM Tris, 300 mM NaCl, 20% v/v glycerol, pH=7.4, supplemented with two protease inhibitor tablets), and lysed using a M-110 PS Microfluidiser (Microfluidics). After supernatant clarification (20,000 × g, 20 min, 4 °C), the membranes were pelleted down (100,000 × g, 2 h, 4 °C) and homogenised in 6 mL buffer B (20 mM Tris, 100 mM NaCl, 20% v/v glycerol, pH=7.4, supplemented with one protease inhibitor per 50 mL). The membrane suspension was separated into 2 mL aliquots, which were frozen in liquid nitrogen and could be stored at - 80 °C for up to two years.

Membrane protein purification. A membrane aliquot (2 mL) was added to a mixture of 9 mL buffer B and 1 mL DDM stock solution (10w% DDM in MilliQ water). The suspension was agitated for one hour at a temperature of 4 °C before the supernatant was clarified by centrifugation (4,000 × g, 30 min, 4 °C). The supernatant was purified by IMAC as described below.

First, the protein-containing supernatant was mixed with the IMAC resin as follows: 7 mL of nickel-nitrilotriacetic acid (Ni-NTA) agarose suspension (50%, Quiagen) were mixed with 23 mL of MilliQ. The supernatant was clarified by centrifugation (4,000 × g, 2 min, 4 °C), discarded, and the procedure was repeated. The ethanol-free resin was suspended in 5 mL IMAC wash buffer (50 mM Tris, 200 mM NaCl, 20 mM imidazole, 10% v/v glycerol, 2xCMC of DDM, pH = 8) and the protein-containing supernatant was added. The mixture was agitated at 4 °C for 15 minutes and then loaded into an empty gravity flow column (14 cm high, 1.5 x 1.2 cm polypropylene columns from Bio-Rad). Once the liquid passed though the column, the IMAC resin was washed with 45 mL IMAC wash buffer, 45 mL of a IMAC wash/elute buffer mixture (9/1 v/v, 2xCMC of DDM, pH = 8), and the protein was eluted with 10 mL IMAC elute buffer (50 mM Tris, 200 mM NaCl, 250 mM imidazole, 10% v/v glycerol, 2xCMC of DDM, pH = 8). The eluted membrane protein was concentrated with centrifugal filters (Amicon®). The molecular weight cut-off (MWCO) of the centrifugal filters was adjusted to the molecular weight of the proteomicelle formed with the expected membrane protein oligomer (MWCO = 50 kDa for TSPO, MsCl-GFP; MWCO = 100 kDa for AqpZ-GFP, AmtB-MBP, AcrB and $BtuC_2D_2$). The volume was reduced to 5 mL and His-tagged Tobacco Etch Virus (TEV) protease was added to GFP- or MBP-tagged membrane proteins .The mixture was transferred into dialysis cassettes (MWCO = 3.5 kDa) and dialysed for 16 hours at 4 °C against dialysis buffer (50 mM Tris, 200 mM NaCl, 20 mM imidazole, 10% v/v glycerol, 2xCMC of DDM, 5 mM 2-mercaptoethanol, pH = 8). His-tagged membrane proteins were dialysed under similar conditions but without the use of TEV. The dialysed protein mixture was then purified by reverse IMAC. For this purpose, the IMAC column was washed with 20 mL dialysis buffer. The

dialysed protein mixture was passed over the column and the flow-through was collected. The column was washed with another 5 mL of dialysis buffer and the flow-through was collected. The combined flow-throughs were concentrated in centrifugal filters until a protein concentration between 30 µM and 50 µM was reached. The protein solutions were separated into 45 µL aliquots, frozen in liquid nitrogen, and could be stored at – 80 °C for up to one year.

Gradual phospholipid delipidation. Gradual PL delipidation was achieved with a detergent exchange from DDM to OG, C8E4, or **1** – **5** over a 3 mL SEC column (Superdex 200 10/300GL column). The column was equilibrated with 1.2 CVs of ammonium acetate solution (200 mM, pH = 6.8, 2xCMC of either OG, C8E4 or detergents **1**- **5**). A 45 µL aliquot of purified protein was used for each detergent exchange. The protein was eluted over 1.5 CVs of detergent-containing ammonium acetate solution at a flow rate of 0.2 mL/min. The main protein fractions were combined and concentrated using Amicon® Ultra 0.5 mL centrifugal filters to a final volume of about 20 to 30 µL.

Lipopolysaccharide delipidation. To achieve LPS delipidation, IMAC resin was prepared as follows: an empty bio-spin column (Bio-Rad) was loaded with 500 µL of Ni-NTA agarose suspension (50%, Quiagen). The resin was washed with 500 µL MilliQ water, 500 µL IMAC elute buffer (50 mM Tris, 200 mM NaCl, 250 mM imidazole, 10% v/v glycerol, 1w% of detergent 1, pH = 8), and 500 µL IMAC wash buffer (50 mM Tris, 200 mM NaCl, 20 mM imidazole, 10% v/v glycerol, 1w% of detergent **1**, pH = 8). The column was loaded with two 45 µL aliquots of His-tagged membrane protein. To identify how many CVs of IMAC wash buffer are needed to delipidate LPS, six columns were prepared in parallel. The columns were washed with different CVs of IMAC wash buffer: 0, 1, 5, 10, 20, or 40. Subsequently, the

columns were washed with 1 mL of an IMAC wash/elute buffer mixture (9/1 v/v, 2xCMC of detergent **1**, pH = 8). The proteins were eluted with 650 µL IMAC elute buffer (50 mM Tris, 200 mM NaCl, 20 mM imidazole, 10% v/v glycerol, 2xCMC of detergent **1**, pH = 8). The relative LPS concentration was monitored by SDS PAGE silver stain analysis and the samples were concentrated to a final volume of 20 to 30 µL using Amicon® Ultra 0.5 mL centrifugal filters. The sample buffer was exchanged to ammonium acetate solution (200 mM, pH = 6.8, 2xCMC of detergent **1**) using 75 µL Zeba™ Spin Desalting columns (MWCO = 7 kDa, Thermo Fisher Scientific). The buffer exchange was done twice.

Native mass spectrometry. The samples were analysed using a modified Q-Exactive mass spectrometer using the following instrumental parameters: injection flatapole: 7.9 V, inter flatapole lens: 6.9 V, bent flatapole: 5.9 V, transfer multipole: 4 V, capillary voltage: 1.2 kV, source temperature: 100 °C, IST: 100 V, HCD cell pressure: 9×10^{-10} mBar, noise level parameter: 3, microscans: 10, and resolution: 17,500 (unless stated otherwise in the text). HCD voltage was adjusted for each MP-lipid combination as specified in the text. Intensities of the *apo* states and protein-lipid complexes were extracted from Xcalibur V2.2. Molecular masses of membrane protein ions and bound ligands were calculated using Navia Beta v0.5 (https://d-que.github.io/navia/) and Origin V9.1. Relative intensities of *apo* states and protein-lipid complexes as well as average charge state values of membrane protein ions were calculated with Origin V9.1.

Acknowledgments

First and foremost, I would like to thank Leonhard Urner for involving me in this fascinating project. Leo designed and synthesised the detergents **1**-**5** and carried

out MP expression (with my assistance for TSPO) and purification (with my assistance for the final detergent exchange steps). He also performed the silver staining, CD spectroscopy and the ATPase activity assay on BtuCD. I would also like to thank Francesco Fiorentino for expressing BtuCD from ClearColi and for his assistance with the activity assay.

References

1 Marsh, D. Protein modulation of lipids, and vice-versa, in membranes. *Biochim Biophys Acta* **1778**, 1545-1575, doi:10.1016/j.bbamem.2008.01.015 (2008).
2 Santos, A. L. & Preta, G. Lipids in the cell: organisation regulates function. *Cell Mol Life Sci* **75**, 1909-1927, doi:10.1007/s00018-018-2765-4 (2018).
3 Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol Rev* **40**, 133-159, doi:10.1093/femsre/fuv008 (2016).
4 Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Biol* **19**, 281-296, doi:10.1038/nrm.2017.138 (2018).
5 Bechara, C. & Robinson, C. V. Different modes of lipid binding to membrane proteins probed by mass spectrometry. *J Am Chem Soc* **137**, 5240-5247, doi:10.1021/jacs.5b00420 (2015).
6 Paschkowsky, S., Oestereich, F. & Munter, L. M. Embedded in the Membrane: How Lipids Confer Activity and Specificity to Intramembrane Proteases. *J Membrane Biol* **251**, 369-378, doi:10.1007/s00232-017-0008-5 (2018).
7 Musatov, A. & Sedlak, E. Role of cardiolipin in stability of integral membrane proteins. *Biochimie* **142**, 102-111, doi:10.1016/j.biochi.2017.08.013 (2017).
8 Lee, A. G. Lipid-protein interactions. *Biochem Soc Trans* **39**, 761-766, doi:10.1042/BST0390761 (2011).
9 Contreras, F. X., Ernst, A. M., Wieland, F. & Brugger, B. Specificity of intramembrane protein-lipid interactions. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a004705 (2011).
10 Birch, J. *et al.* The fine art of integral membrane protein crystallisation. *Methods* **147**, 150-162, doi:10.1016/j.ymeth.2018.05.014 (2018).
11 Thonghin, N., Kargas, V., Clews, J. & Ford, R. C. Cryo-electron microscopy of membrane proteins. *Methods* **147**, 176-186, doi:10.1016/j.ymeth.2018.04.018 (2018).
12 Loll, P. J. Membrane proteins, detergents and crystals: what is the state of the art? *Acta Crystallogr F Struct Biol Commun* **70**, 1576-1583, doi:10.1107/S2053230X14025035 (2014).
13 Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172-+, doi:10.1038/nature13419 (2014).
14 Chorev, D. S. *et al.* Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science* **362**, 829-834, doi:10.1126/science.aau0976 (2018).
15 Ilgu, H. *et al.* Variation of the detergent-binding capacity and phospholipid content of membrane proteins when purified in different detergents. *Biophys J* **106**, 1660-1670, doi:10.1016/j.bpj.2014.02.024 (2014).

16 Gault, J. *et al.* Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. *Nat Methods* **17**, 505-508, doi:10.1038/s41592-020-0821-0 (2020).
17 Gupta, K. *et al.* Identifying key membrane protein lipid interactions using mass spectrometry. *Nat Protoc* **13**, 1106-1120, doi:10.1038/nprot.2018.014 (2018).
18 Bolla, J. R. *et al.* A Mass-Spectrometry-Based Approach to Distinguish Annular and Specific Lipid Binding to Membrane Proteins. *Angew Chem Int Ed Engl* **59**, 3523-3528, doi:10.1002/anie.201914411 (2020).
19 Bechara, C. *et al.* A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nat Chem* **7**, 255-262, doi:10.1038/Nchem.2172 (2015).
20 Urner, L. H. *et al.* Modular detergents tailor the purification and structural analysis of membrane proteins including G-protein coupled receptors. *Nat Commun* **11**, 564, doi:10.1038/s41467-020-14424-8 (2020).
21 Thota, B. N., Urner, L. H. & Haag, R. Supramolecular Architectures of Dendritic Amphiphiles in Water. *Chem Rev* **116**, 2079-2102, doi:10.1021/acs.chemrev.5b00417 (2016).
22 Pasquali, R. C., Taurozzi, M. P. & Bregni, C. Some considerations about the hydrophilic–lipophilic balance system. *Int. J. Pharm.* **356**, 44-51 (2008).
23 Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. Theory of Self-Assembly of Hydrocarbon Amphiphiles into Micelles and Bilayers. *J. Chem. Soc., Faraday Trans. 2* **72**, 1525-1568 (1976).
24 Nagarajan, R. Molecular Packing Parameter and Surfactant Self-Assembly: The Neglected Role of the Surfactant Tail. *Langmuir* **18**, 31-38 (2002).
25 Trappmann, B. *et al.* A New Family of Nonionic Dendritic Amphiphiles Displaying Unexpected Packing Parameters in Micellar Assemblies. *J. Am. Chem. Soc.* **132**, 11119-11124 (2010).
26 Thota, B. N. S., v. Berlepsch, H., Böttcher, C. & Haag, R. Towards engineering of self-assembled nanostructures using non-ionic dendritic amphiphiles. *Chem. Commun.* **51**, 8648-8651 (2015).
27 Akula, S., Gurram, A. K. & Devireddy, S. D. Self-Microemulsifying Drug Delivery Systems: An Attractive Strategy for Enhanced Therapeutic Profile. *Int. Sch. Res. Notices.* **2014**, 1-11 (2014).
28 Urner, L. H., Maier, Y. B., Haag, R. & Pagel, K. Exploring the Potential of Dendritic Oligoglycerol Detergents for Protein Mass Spectrometry. *J Am Soc Mass Spectrom* **30**, 174-180, doi:10.1007/s13361-018-2063-2 (2019).
29 Reading, E. *et al.* The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase. *Angew Chem Int Edit* **54**, 4577-4581, doi:10.1002/anie.201411622 (2015).
30 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, doi:10.1021/ja401736v (2013).
31 Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat Methods* **13**, 333-336, doi:10.1038/nmeth.3771 (2016).
32 Emiola, A., George, J. & Andrews, S. S. A Complete Pathway Model for Lipid A Biosynthesis in Escherichia coli. *PLoS One* **10**, e0121216, doi:10.1371/journal.pone.0121216 (2014).
33 Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635-700, doi:10.1146/annurev.biochem.71.110601.135414 (2002).
34 Whitfield, C. & Trent, M. S. Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* **83**, 99-128, doi:10.1146/annurev-biochem-060713-035600 (2014).

35 Fiorentino, F., Bolla, J. R., Mehmood, S. & Robinson, C. V. The Different Effects of Substrates and Nucleotides on the Complex Formation of ABC Transporters. *Structure* **27**, 651-659 e653, doi:10.1016/j.str.2019.01.010 (2019).

36 Reading, E. *et al.* The Effect of Detergent, Temperature, and Lipid on the Oligomeric State of MscL Constructs: Insights from Mass Spectrometry. *Chem Biol* **22**, 593-603, doi:10.1016/j.chembiol.2015.04.016 (2015).

37 Prive, G. G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397, doi:10.1016/j.ymeth.2007.01.007 (2007).

38 Carpenter, E. P., Beis, K., Cameron, A. D. & Iwata, S. Overcoming the challenges of membrane protein crystallography. *Curr Opin Struct Biol* **18**, 581-586, doi:10.1016/j.sbi.2008.07.001 (2008).

39 Botos, I., Noinaj, N. & Buchanan, S. K. Insertion of proteins and lipopolysaccharide into the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* **372**, doi:10.1098/rstb.2016.0224 (2017).

40 Simpson, B. W., May, J. M., Sherman, D. J., Kahne, D. & Ruiz, N. Lipopolysaccharide transport to the cell surface: biosynthesis and extraction from the inner membrane. *Philos Trans R Soc Lond B Biol Sci* **370**, doi:10.1098/rstb.2015.0029 (2015).

41 Clairfeuille, T. *et al.* Structure of the essential inner membrane lipopolysaccharide-PbgA complex. *Nature* **584**, 479-483, doi:10.1038/s41586-020-2597-x (2020).

42 Schultz, K. M. & Klug, C. S. Characterization of and lipopolysaccharide binding to the E. coli LptC protein dimer. *Protein Sci* **27**, 381-389, doi:10.1002/pro.3322 (2018).

43 Kobylka, J., Kuth, M. S., Muller, R. T., Geertsma, E. R. & Pos, K. M. AcrB: a mean, keen, drug efflux machine. *Ann N Y Acad Sci* **1459**, 38-68, doi:10.1111/nyas.14239 (2020).

44 Mehmood, S. *et al.* Charge Reduction Stabilizes Intact Membrane Protein Complexes for Mass Spectrometry. *J Am Chem Soc* **136**, 17010-17012, doi:10.1021/ja510283g (2014).

45 Liko, I., Hopper, J. T., Allison, T. M., Benesch, J. L. & Robinson, C. V. Negative Ions Enhance Survival of Membrane Protein Complexes. *J Am Soc Mass Spectrom* **27**, 1099-1104, doi:10.1007/s13361-016-1381-5 (2016).

46 Tsai, C. M. & Frasch, C. E. A Sensitive Silver Stain for Detecting Lipopolysaccharides in Polyacrylamide Gels. *Anal Biochem* **119**, 115-119, doi:Doi 10.1016/0003-2697(82)90673-X (1982).

47 Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876-2890, doi:10.1038/nprot.2006.202 (2006).

48 Patrick, J. W. & Laganowsky, A. Generation of Charge-Reduced Ions of Membrane Protein Complexes for Native Ion Mobility Mass Spectrometry Studies. *J Am Soc Mass Spectrom*, doi:10.1007/s13361-019-02187-6 (2019).

49 Consta, S. & Malevanets, A. Manifestations of charge induced instability in droplets effected by charged macromolecules. *Phys Rev Lett* **109**, 148301, doi:10.1103/PhysRevLett.109.148301 (2012).

50 Von Helden, G., Hsu, M. T., Gotts, N. & Bowers, M. T. Carbon Cluster Cations with up to 84 Atoms - Structures, Formation Mechanism, and Reactivity. *J Phys Chem-Us* **97**, 8182-8192, doi:DOI 10.1021/j100133a011 (1993).

51 Zhao, Y. H., Abraham, M. H. & Zissimos, A. M. Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. *J Org Chem* **68**, 7368-7373, doi:10.1021/jo034808o (2003).

4 Investigating Lipid-Controlled Dimerisation of a Mammalian Ion Exchanger by Native Mass Spectrometry and Complementary Biophysical Methods

Abstract

Na^+/H^+ exchangers (NHEs) are important ion transporters, found in all kingdoms of life. In humans, dysfunctions of NHEs are linked to many diseases, including cancers. Despite their physiological importance, structural information on mammalian NHEs was lacking due to their poor stabilities in detergents. Here, a first high-resolution structure of *horse* NHE9 obtained by cryogenic electron microscopy (cryo-EM) is presented.

In this work, I employ native mass spectrometry (MS) to probe the effect of lipids on dimerisation of NHE9. I show that dimers are exclusively observed in the presence of brain lipids, while a monomer/dimer equilibrium exists in their absence. Based on that information, a thermal shift assay revealed the enhanced stability of NHE9 in the presence of negatively charged phosphoinositide lipids (PIPs). Further native MS analysis, performed on the mutant of NHE9 with proposed PIP binding site disrupted, confirmed the importance of these lipids for NHE9 dimerisation. This finding led to a hypothetical mechanism, stating that the role of PIPs is to regulate NHE9 function by controlling dimerisation.

Overall, the combined results obtained by me and collaborators contribute towards significant advancement of our understanding of NHE9 and NHEs in general.

4.1 Introduction

(*Note: this chapter was of a highly collaborative nature. I have performed native mass spectrometry (MS) experiments only; please see the acknowledgments section at the end of this chapter for a detailed list of contributions. A large part of the data presented here can be found in a recent publication[1].)*

Na^+/H^+ exchangers (NHEs) are ion transporters found in all kingdoms of life[2-5], and are probably some of the oldest protein transport forms[6,7]. NHEs function by coupling transport of H^+ ions to counter transport of ions such as Na^+ and Li^+[8,9]. In mammals, there are 13 distinct NHE orthologues, differing in tissue localisation and substrate preferences[5] (**Figure 4.1**). Dysfunction of NHEs has been linked to many diseases, including cancer, heart failure and diabetes[4,10]. However, NHE6 and NHE9 are the only two isoforms with known disease-causing mutations[4]. Despite their obvious importance, structures and molecular details of NHEs are poorly understood, largely due to their instabilities in detergent environments.

Structurally, all NHEs are expected to form functional homodimers[11]. Each monomer consists of a long cytosolic C-terminal domain that regulates ion exchange activity, and a transporter module domain that performs ion exchange[4,5]. The transporter module is believed to contain a dimerisation domain and an ion-transport domain (based on the structures of bacterial homologues[12]) – the ion transport domain functions by performing large 'elevator-like movements', while the dimerisation domain remains stationary[12,13].

The instability in detergent is likely to arise from the large unstructured C-terminal cytosolic domain. Therefore, this study was focused on NHE9, which has one of the shortest such domains compared to the other NHEs[1]; in addition, a potential

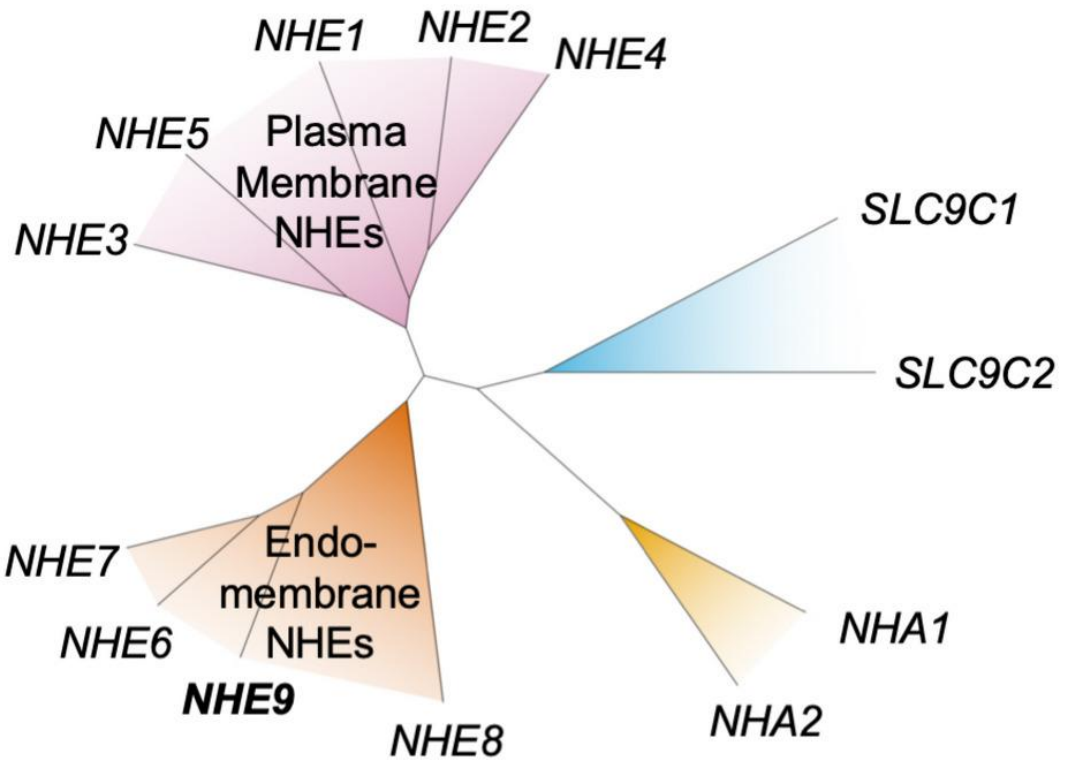


Figure 4.1 NHE orthologues. Most NHEs are present either in plasma membranes or in endosome membranes. Figure adapted from ref (1)

interpretation of the disease-causing mutation could arise from the high resolution structure. *Horse* NHE9, which has 95% sequence identity to *human* NHE9 was shown to be the most stable of the available options and, therefore, chosen for the cryogenic electron microscopy (cryo-EM) studies[1] (see **Section 1.4** for a brief description of cryo-EM).

4.2 Obtaining High Resolution Cryo-EM Structure of NHE9

(*Note: details of cryo-EM experimentation can be obtained from the publication[1]*). Construct with a partially truncated C-terminal tail was used for the initial structure determination experiment (NHE9*, residues 8-574) to improve protein yield and stability. However, even this truncated version of the C-terminal domain caused complications during cryo-EM structural determination, with no density at all being

observed for that 93-residue long domain for either of the monomeric subunits (**Figure 4.2a**). A mutant with the C-terminal domain (CTD) completely removed (NHE9 ΔCTD) was created. Importantly, this mutant maintained the same activity as NHE9* (at the same time, activity was abolished for a control mutant, with a conservative 'ND' motif substituted to alanine residues) (**Figure 4.2b**). For NHE9 ΔCTD mutant a cryo-EM structure at resolution of 3.2Å was obtained (**Figure 4.2c**). The NHE9 monomer consists of 13 transmembrane (TM) segments, making it more similar to bacterial homologues like NapA[12], rather than the typically used NHE models that only have 12 TM segments, more resembling bacterial NhaA[5]. Since the ion-binding site is heavily conserved across NHEs, this finding can improve the modelling for the other NHEs as well[5].

The dimeric region of NHE9 is highlighted in orange in **Figure 4.2c**. The total buried interfacial area for the dimer was estimated to be between 1700 and 2000 Å. Previous studies using native mass spectrometry (MS) linked the dimeric surface area of Na^+/H^+ antiporters to their ability to form dimers without assistance from lipids: NhaA that has a small (around 700 Å) interface required negatively charged cardiolipin (CDL) in order to form the dimer, while NapA with a larger (around 1800 Å) subunit interface could form successfully dimerise without any lipids being required[14,15]. By that reasoning, NHE9 could be expected to form *apo* dimer without the need for stabilising lipids. However, an interesting observation was made regarding positioning of TM3. Looking at the bottom panel of **Figure 4.2c**, the end of TM3 on each subunit can be observed to move away from the interface, exposing a hydrophobic surface. Interestingly, a similar arrangement creating a lipid binding site exists for a structurally related protein *Pa*NhaP[16], as well as for a distantly related citrate transporter CitS[17]. Coarse-grained molecular

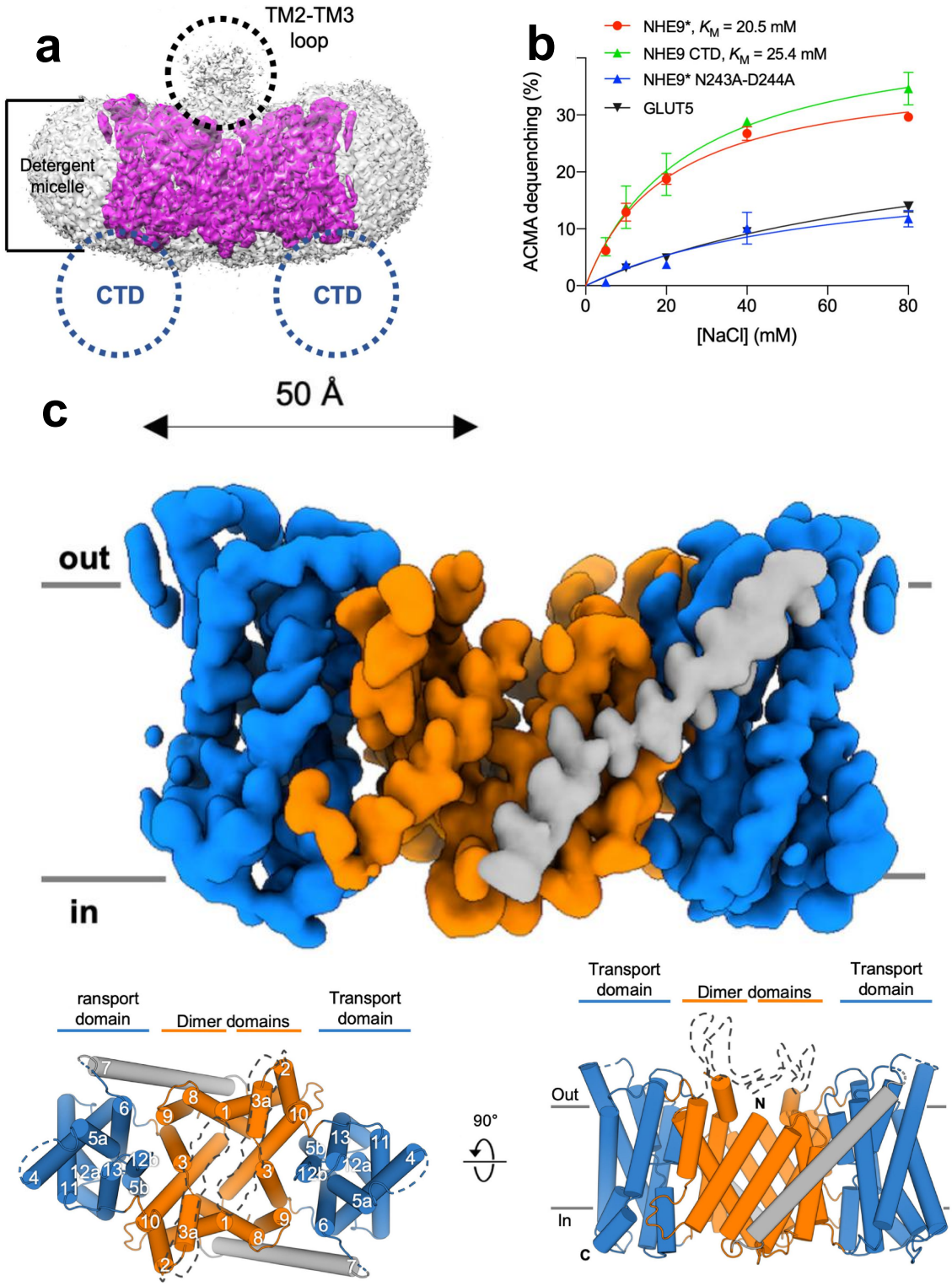


Figure 4.2 Cryo-EM structure of NHE9. a, Cryogenic electron microscopy (cryo-EM) structure of NHE9* - C-terminal domain (CTD) and the loop between transmembrane segments TM2-TM3 are missing for each monomer. **b,** Michaelis-Menten kinetics as detected by ACMA dequenching following substrate addition; values shown are the average of three technical repeats, with error bars representing one standard error of the mean. **c**, Cryo-EM density map (top) and cartoon representation (bottom) of the NHE9 ΔCTD homodimer; the dimer domain is shown in orange. Figure adapted from ref (1)

dynamics (CG-MD) simulations were employed to obtain more insight into the nature of the potential lipid binding event, however lipids were unable to diffuse to the interface[1]. In order to investigate the plausible lipid-induced dimerisation of NHE9, native MS experiments were performed.

4.3 Native Mass Spectrometry Reveals the Role of Lipids for NHE9 Dimerisation

Native MS is a method that can observe different oligomeric states of proteins and their individual interactions with lipids or ligands in a single experiment[18]. This technique is, therefore, well-suited for characterisation of lipid-induced oligomerisation[14,15]. Here, native MS experiments were performed on the NHE9* construct. While the TM3 movement was observed for the NHE9 ΔCTD mutant, the main difference between the two structures was the higher resolution for the latter, with no significant differences in segment positioning. NHE9* was expressed in *Saccharomyces cerevisiae*, which does not have the same lipid composition as the mammalian origin. Since NHE9 is known to be highly expressed in the pre-frontal cortex of the brain[19], brain lipids were added back to NHE9* during purification. The resulting mass spectrum (obtained on Q Exactive UHMR, see **Figure 1.6d** for description) is shown in **Figure 4.3**. A large 'hump' is observed, however some individual peaks are clearly resolved, with a calculated mass of roughly 3.5 kDa larger than would be expected for the NHE9* dimer. These observations indicate that large amounts of lipids are present in solution, many of which are seen bound to the protein. It should be noted that the highest activation energy available for the instrument (300 V in both in-source trapping (IST) and higher-energy C-trap dissociation (HCD), as well as elevated source temperatures

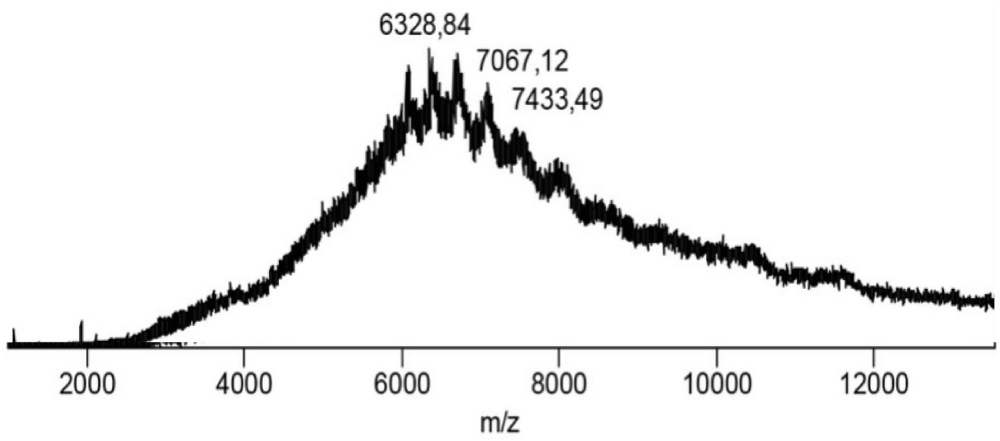


Figure 4.3 Native MS of NHE9 in the presence of brain lipids.
Figure adapted from ref (1)

of 200-300 C) had to be used just to achieve the minimal level of peak resolution observed in **Figure 4.3**; therefore, no further activation is possible to remove some of the lipid adducts.

In order to obtain more interpretable spectra, NHE9* sample with no added lipids was used (**Figure 4.4a**). As expected, the peaks are more resolved than they were for the lipidated sample (**Figure 4.3**). Importantly, this time both monomers and dimers were observed in the same spectrum, whereas only dimers were present for NHE9* in the presence of brain lipids. Additionally, dimers with two different masses were observed, with a mass difference of around 650 Da. The lower-mass dimer is roughly 3.5 kDa larger than would be expected for the *apo* dimer, similarly to the 'brain lipid' NHE9*.

Although the peaks in **Figure 4.4a** are clearly more resolved than those in **Figure 4.3**, signal-to-noise ratio (S/N) is still not very high, hindering data analysis. The observed 'hump' can be partly attributed to the presence of co-expressed lipids, but contribution from detergent is also likely. NHE9 was solubilised in a mixture of *n*-dodecyl-β-d-maltopyranoside (DDM) and cholesteryl hemisuccinate (CHS)

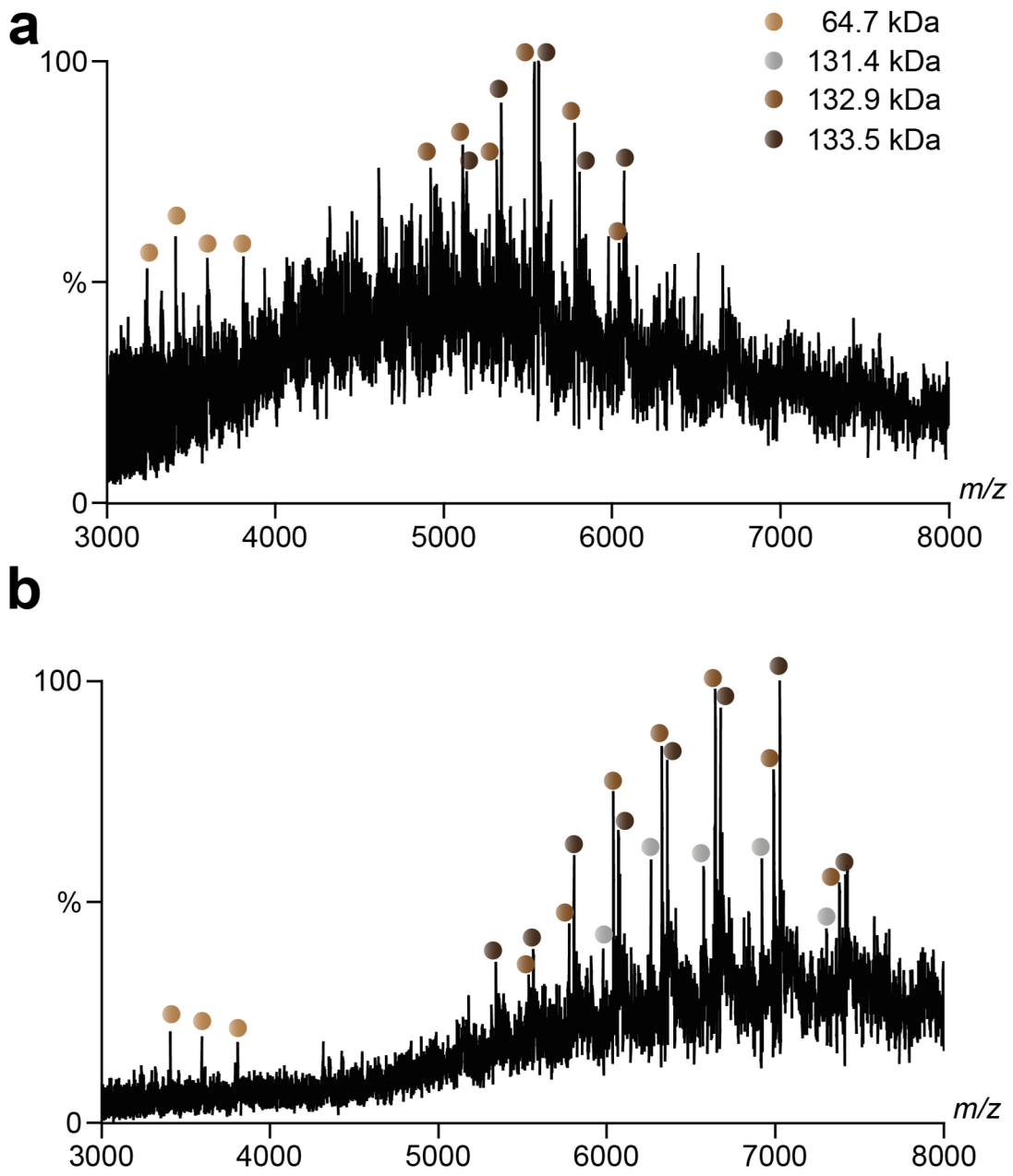


Figure 4.4 Native MS of NHE9 in the absence of brain lipids. a, Before and **b,** after the 'wash' step, consisting of dilution and reconcentration of the sample with ammonium acetate/2xCMC DDM solution prior to the buffer exchange stage. CMC, critical micelle concentration; DDM, *n*-dodecyl-β-d-maltopyranoside.

detergents, and then buffer exchanged into ammonium acetate (AA) buffer containing DDM at 2x critical micelle concentration (CMC) (see **Materials and Methods**). While these detergents are known to be stabilising in solution and are commonly used to study unstable MPs, such as G protein coupled receptors (GPCRs)[20], DDM is not an ideal choice for native MS due to high energies required

to dissociate it from MPs in the gas-phase[21]; in addition, DDM has a strong propensity to form charged clusters with a wide range of masses that can overlap with other peaks in a spectrum and obscure important information[22]. This effect is especially strong when the concentration of DDM is inadvertently increased during a sample concentration stage. Consequently, reducing the concentration of DDM in the sample can be beneficial.

My initial strategy was to replace DDM with a more native MS-compatible detergent (namely tetraethylene glycol monooctyl ether, C8E4, at 2xCMC) during the buffer exchange step. This method results in an incomplete detergent exchange, forming mixed DDM-C8E4 micelles, which are both sufficiently stabilising for the protein and easily removable during native MS experiments (see **Chapter 5** for an example of this effect). Unfortunately, addition of C8E4 resulted in spectra mostly dominated by DDM peaks, with no NHE9* visible (data not shown). This observation implies that even partial exchange into a harsh detergent is sufficient to destabilise this protein.

I then decided to perform a 'wash' step; to attempt decreasing the concentrations of DDM, CHS and any non-specific endogenous lipids present in solution. This 'wash' step consisted of diluting the sample with AA/DDM (2xCMC) buffer and then reconcentrating it with a 100k molecular weight cut-off (MWCO) concentrator (see **Materials and Methods** for more details). Inclusion of this step resulted in significant improvement of mass spectral quality (**Figure 4.4b**). This enabled the observation of another species of NHE9* dimer, with a mass roughly 1.5 kDa below the previously observed variants (**Figure 4.4a**). The absence of this species in the 'unwashed' spectrum can be explained by its relatively low intensity, probably resulting in it being 'buried' in the noise. Another notable feature is the apparent

shift towards lower charge states (higher mass-to-charge ratio (*m/z*) values) after the 'wash' step. Considering that no new chemicals were added to the sample, the shift is unlikely to arise from charge reduction, but instead from the lower charge states not being resolved in the 'unwashed' spectra (**Figure 4.4a**). The energy that the ion experiences at a given voltage is directly proportional to its charge; therefore, lower-charge MPs preserve a larger number of adducts, which can result in an overlap and lead to those species being unresolved. The observation of a previously undetected NHE9 species, as well as the appearance of more charge states, highlights the utility of the 'wash' step, so it was performed for all subsequent native MS experiments on NHE9. It is important to note, that the outcome of the 'wash' step proved to be highly unreliable, often resulting in no improvement, or even reduction in spectral quality compared to the 'unwashed' sample. This caused significant sample losses, and was extremely time-consuming. Therefore, I would only recommend this method when no alternative way of improving the sample quality (such as an extra size-exclusion chromatography (SEC) step, for example) can be employed.

The unpredictable nature of the 'wash' step also resulted in variations between the spectra obtained from the different batches of the sample, which were often difficult or impossible to reproduce. This issue was exacerbated by the expected variations in native mass spectra between the different nanoelectrospray emitters (needles), which are prepared in-house and have unique shapes and sizes[23]. This effect proved to be very significant for NHE9, with spectral quality often changing drastically when two different needles were used, even for the samples coming from the same batch. One example of sample-to-sample variation can be observed in the spectrum shown in **Figure 4.5a**, which was obtained by following exactly the

same procedure that was used to produce the spectrum shown in **Figure 4.4b**, but applied to a different 'washed' batch of the sample. While the dimer is similar in appearance (apart from the charge state distribution, which could be a 'needle effect'), a second form of monomer can be clearly observed only in the spectrum shown in **Figure 4.5a**. This monomer species is of a higher mass than the one discussed earlier, representing either a ligand or lipid-bound state or some covalent modification of the protein. Interestingly, the mass of the monomer matched the half-mass of one of the dimer species (132.9 kDa) quite well. The peaks representing the 'new' monomer are broad in appearance (**Figure 4.5b**); these features often arise because of the binding of endogenous lipids[24], which have inherently heterogeneous masses due to a large chemical diversity of their hydrophobic tails[25] (see **Section 1.2** for further discussion). However, for NHE9* even the lower-mass monomer peaks are also broad, so it is difficult to determine the nature of the mass difference with confidence.

In summary, no *apo* dimers of NHE9* were observed in any spectra (**Figures 4.3-4.5**), with all of the observed dimers having more mass than would be expected from a sum of two monomeric masses. This could indicate the requirement for ligand or lipid binding for NHE9 dimer formation; however, further investigation of the nature of this possible binder by native MS in combination with various activation methods (such as HCD) was not possible, because the highest instrumental potentials had to be used just to produce the spectra. Therefore, complementary methods were employed to identify potential binding partners.

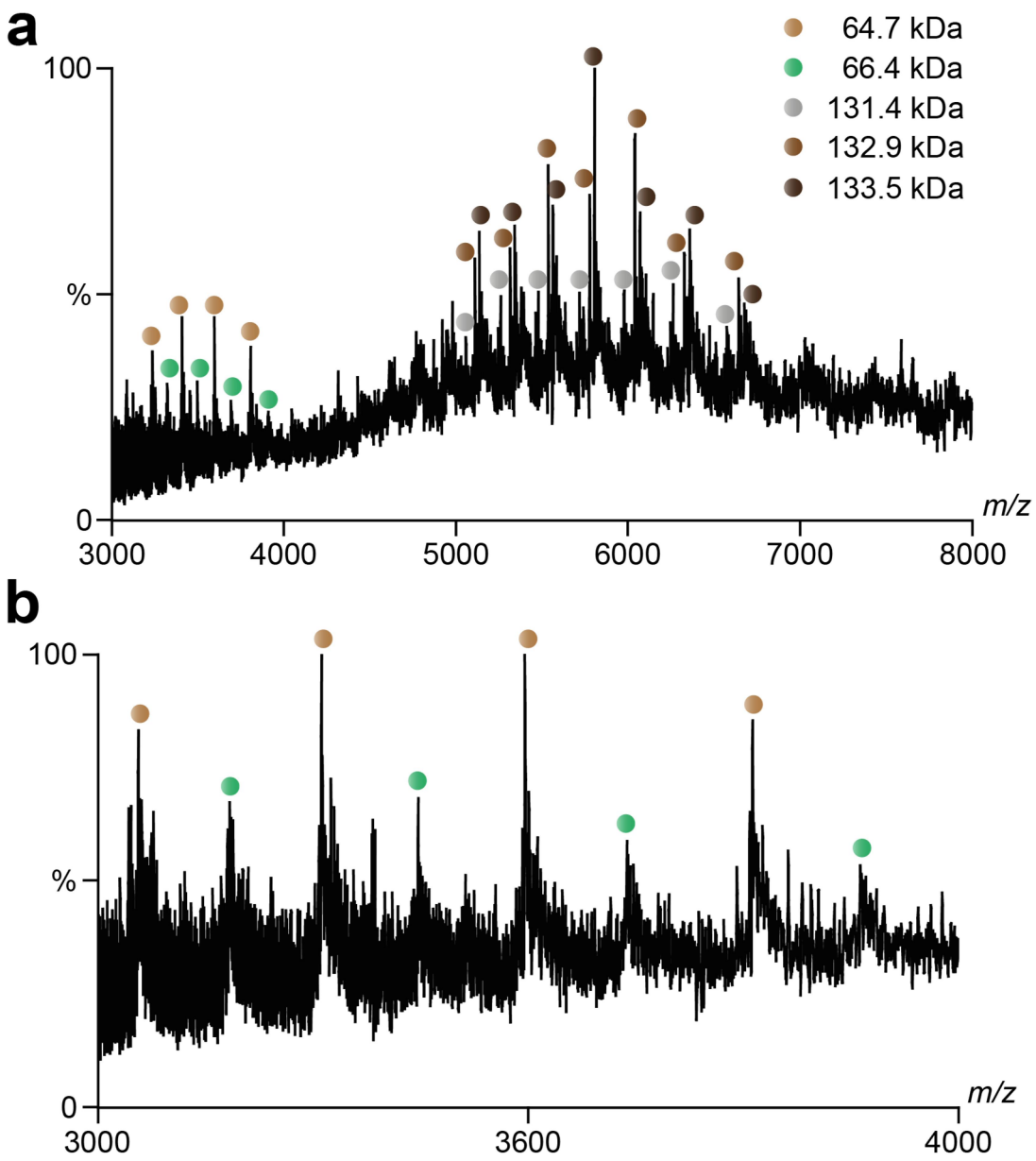


Figure 4.5 Native MS of NHE9 with two monomeric species. a, Spectrum of NHE9* after the 'wash' step. A second monomeric species is observed. **b,** Zoomed-in view of the two monomers. Unlabelled peaks were assigned to impurities.

A green fluorescent protein (GFP)-based thermal shift assay was previously able to correctly identify the specific nature of CDL-binding to NhaA[26], with CDL known to be required for dimerisation of that protein[15]. This assay works by gradually increasing the temperature of a protein-containing solution and observing the changes in fluorescence by SEC; a characteristic transition occurs at a particular melting temperature, T_m, that indicates protein denaturing and forming aggregates,

with oligomeric structures dissociating as a result of this denaturation at the same time[26]. Among several endogenously added lipids, only phosphoinositides, PIP_2 and PIP_3 (described in **Section 1.2**), were discovered to be significantly stabilising by this assay (**Figure 4.6 a** and **b**). Interestingly, PIP_2 was previously shown to be essential for NHE1 function[5,27,28], and NHE3 can be directly and reversibly activated by PIP_3[27]. For NHE1, this occurs partly due to the interaction of the lipid with CTD[5]; however in this study NHE9 ΔCTD mutant was shown to have the same thermal stabilisation by PIP_2 as the NHE9* construct, meaning that CTD is not responsible for the specific binding observed (data not shown). By taking into account native MS data that suggested a possible role of lipids for NHE9 dimerisation, a potential lipid-binding site in the dimer interface region was considered. However, no positively charged residues were observed there in the cryo-EM structure, which would be expected for a specific interaction with negatively charged PIPs[29]. The attention was drawn towards TM2-TM3 loop, which is not observed in the cryo-EM structure due to its dynamic nature (**Figure 4.2a**). Intriguingly, TM2-TM3 loop contains three positively charged lysine residues. Moreover, it is predicted to form β-strands (**Figure 4.6c**), in a similar fashion to the β-hairpins in NhaA that mediate dimerisation by binding to the negatively charged CDL[14,15,26,30]. In order to test whether or not PIPs actually bind at the TM2-TM3 loop, a triple mutant was constructed, with all of the lysine residues substituted for glutamine. The triple mutant of NHE9* was significantly less stabilised by PIP_2 compared to NHE9* (**Figure 4.6d**), confirming the role of the TM2-TM3 loop in PIP binding.

Native MS was again employed to further validate the observed results. First, I focused on the NHE9* construct, to investigate whether adding PIP_2 to the sample

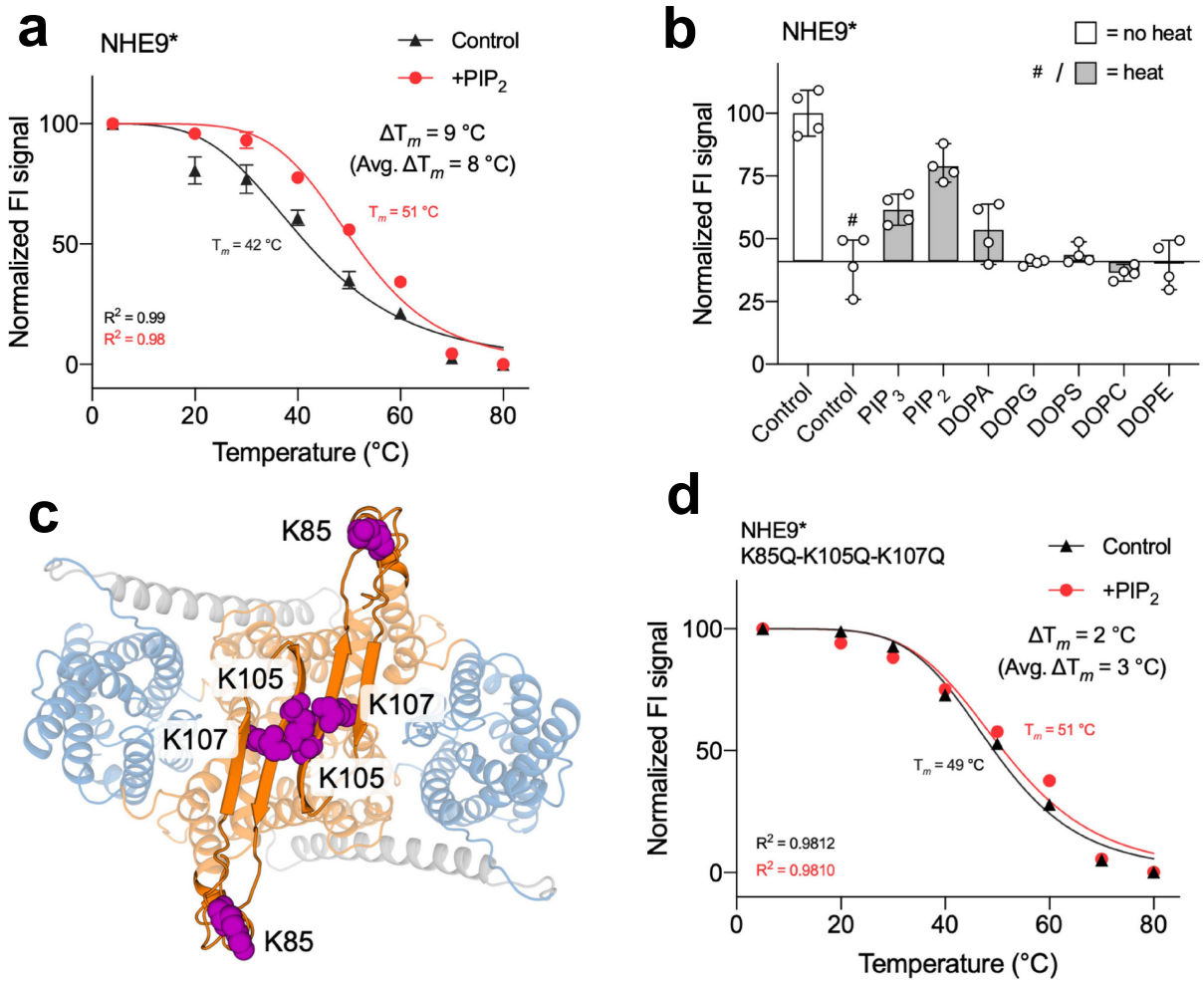


Figure 4.6 Thermal shift assay of NHE9. **a,** Thermal shift stabilisation of purified dimeric NHE9*-GFP in the presence of PIP_2 (red) compared to PIP_2-free (black). Data presented are normalised mean FSEC peak fluorescence as mean values ± data range of n = 2 technical repeats; the average ΔT_m presented is calculated from n = 2 independent titrations. **b,** Thermal stabilisation of purified dimeric NHE9*-GFP by lipids. Normalised mean FSEC peak fluorescence before heating (open bars), and after heating and centrifugation (grey bars) in the presence of different lipids. Data presented are mean values ± data range of n = 4 experiments. **c,** Cartoon representation of NHE9, including the model of the loop TM2-TM3. **d,** Same as **a**, but for the dimeric NHE9* triple mutant. Figure adapted from ref (1).

would result in an observable shift in oligomeric state distribution. Starting from a completely delipidated monomeric NHE9* and exogenously adding PIP_2 at various concentrations would produce a clearly interpretable outcome. Therefore, a method to delipidate NHE9* was desired. While I described a novel delipidation method in **Chapter 3** (which works by exchanging the samples into specific

detergents) that could be potentially applicable to this problem, that method was not attempted on NHE9*. There are two main reason for that: firstly, these projects were performed in parallel, so the delipidation method had not yet been completely established at the time; secondly, the delipidation protocol described in **Chapter 3** needs to be applied during purification stages, so relatively large sample volumes (around 50 μl) are required – since I did not perform protein expression and purification myself, such a large volume constituted the majority of sample I had available. Considering the poor stability of NHE9* when solubilised in detergents, exchanging the entire sample into a previously untested detergent would be unwise. Instead, a delipidation method that could be applicable to small volumes (roughly below 5 μl) was required. One method that fulfils this criterion involves addition of a strongly delipidating detergent, such as n-nonyl-β-d-glucopyranoside (NG) or *n*-octyl-β-d-glucopyranoside (OG) to the sample directly before native MS characterisation[31]. Therefore, I have added NG at 0.5% concentration to NHE9* and monitored the results by native MS. Remarkably, the addition of NG consistently resulted in complete disappearance of all NHE9* peaks from the spectra (data not shown). Since I have added NG from a concentrated stock and the protein was only diluted by 5% in volume, the disappearance of NHE9* can be attributed to the instability of this protein in the presence of NG. This outcome is consistent with the results of C8E4 addition described above and further highlights the instability of NHE9 in any but the mildest detergents.

Another way to monitor the effect of PIP_2 on NHE9 dimerisation is to add this lipid at a range of concentrations directly to the sample and monitor the shift in the monomer/dimer ratio[15]. Following this approach, I have added PIP_2(4,5) to NHE9* (**Figure 4.7a**). It should be noted that the addition of extra lipids to the already

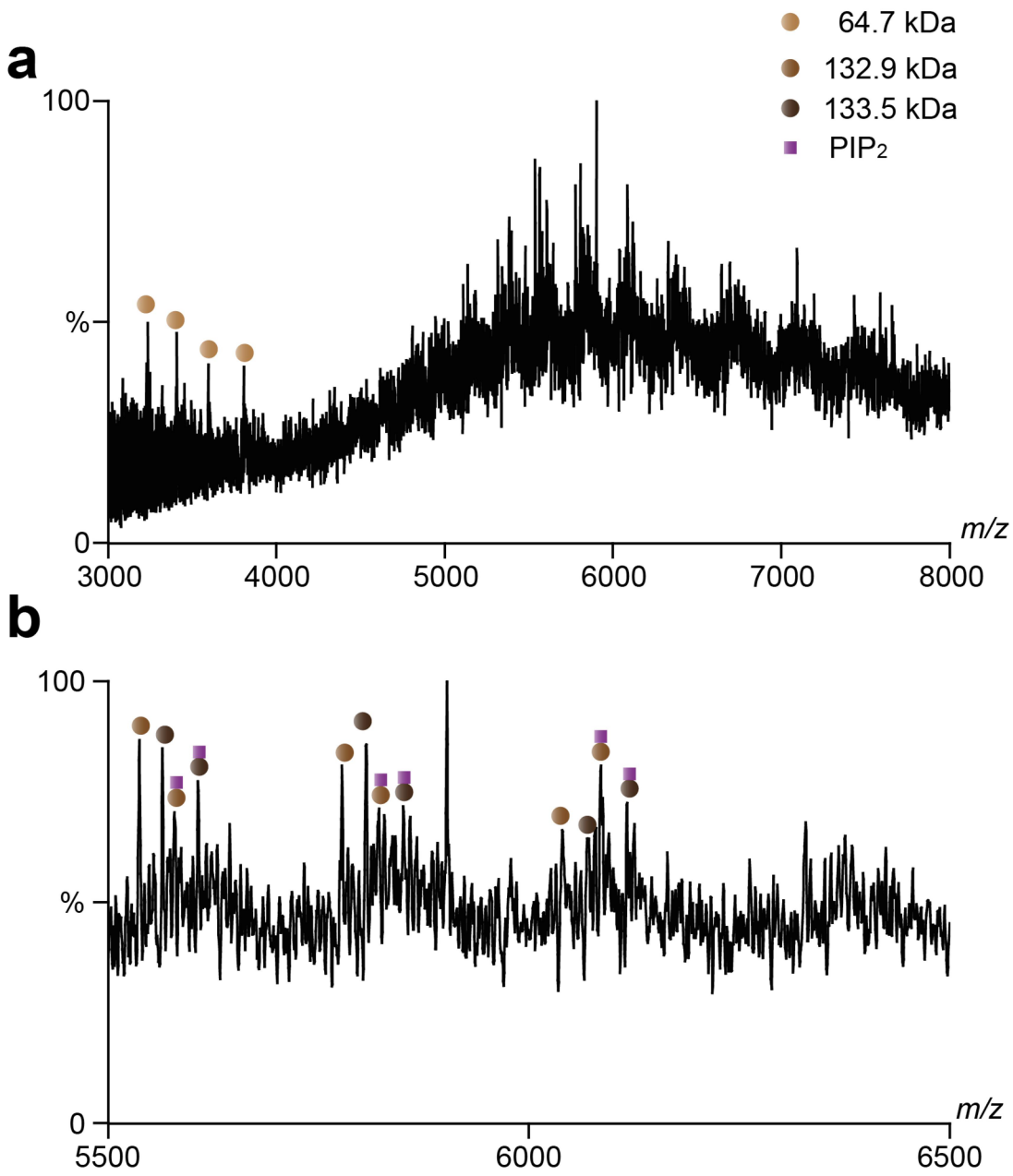


Figure 4.7 Native MS of NHE9 with the addition of PIP2. a, Spectrum of NHE9* with PIP_2(4,5) added, dimeric species are labelled in the zoomed-in spectrum. **b,** Zoomed-in view of NHE9* dimers bound to PIP_2. While signal-to-noise ratio is low, the peaks corresponding to PIP binding are consistently present for three consecutive charge states of the dimer.

complicated and noisy spectra resulted in a further quality deterioration. Nevertheless, the binding of PIP_2 to the NHE9* dimer could be observed (**Figure 4.7b**), although it is not possible to establish whether there are any changes in dimer/monomer ratio of the protein. While some of the peaks are barely

distinguishable from background noise, their consistent presence for three consecutive charge states enables assignment with a certain degree of confidence; however, the low S/N of the spectra in the presence of PIP_2 is evident. This fact, together with reproducibility difficulties discussed above, prevented further investigation of PIP_2-bound NHE9* beyond the confirmation of PIP_2 binding to the dimer. It should be noted that no corresponding binding of the lipid to monomer was observed, although it is challenging to determine if this is the consequence of a lower spectral intensity of monomers compared to dimers (leading to the bound peaks being buried in the noise) or the indication of preferential binding of PIP_2 to the dimer.

Next, I investigated the triple mutant of NHE9*. Using the same sample preparation as the one performed on NHE9* (a 'wash' step followed by buffer exchange procedure, both into AA/DDM at 2xCMC) native mass spectra of this mutant could be acquired (**Figure 4.8a).** Only monomers and no dimers could be observed, corroborating the solution-phase stability assay and confirming the proposed role of TM2-TM3 in dimerisation. It should be noted, that no dimers were observed when repeating the experiment from different needles and sample batches; this included starting from the samples with PIP_2 added to the protein solution prior to the 'wash' step. In some of the spectra a species at 130.0 kDa was appearing, which would be close to a dimeric species with some adducts. However, I have also observed contaminants at 43.3 kDa and 86.6 kDa, which together with the 130.0 kDa matched very well to be a monomer, dimer and trimer of some impurity. I was able to verify this while attempting to achieve further S/N improvement by performing a 'wash' step with AA only (no DDM) followed by a standard AA/DDM buffer exchange, in hope that sufficient concentration of DDM would be left in the

sample to preserve NHE9 stability. Unfortunately, this was not the case and neither monomers nor dimers of NHE9*-triple mutant could be observed after such treatment (**Figure 4.8b**). However, the species at 130.0 kDa and at 86.6 kDa were still present (as well as another contaminant), indicating that the observed peaks most likely belong to a soluble protein. Although this result was achieved inadvertently, the potential utility of this method to assist with identification of species in complex spectra of membrane proteins is evident. For the purposes of this study, confirmation of the assignment of the species with the mass of 130.0 kDa to an impurity strengthened the conclusion that no dimers could be observed for the triple mutant of NHE9*.

4.4 Summary and Conclusions

In this chapter, the first ever high resolution structure of a mammalian NHE, namely *horse* NHE9, was demonstrated. The structure revealed the 13 TM topology similar to NapA, which also allowed to improve modelling of the other members of NHE family. The dimer interface was observed, with the end of TM3 pointing away from the interface and exposing a potential lipid binding site.

Native MS was employed to investigate the role of lipids in NHE9 dimerisation. When brain lipids were present in the sample only NHE9 dimers were observed with bound adducts. In the absence of brain lipids, both monomers and dimers were produced; the dimers also had higher masses than would be expected for the *apo* species. Taken together, these observations suggested the importance of a binding partner for NHE9 dimerisation. A combination of in-solution GFP-based thermal shift assay, structure analysis and mutation determined specific binding of negatively charged phophoinositides, PIP_2 and PIP_3 to an interfacial TM2-TM3

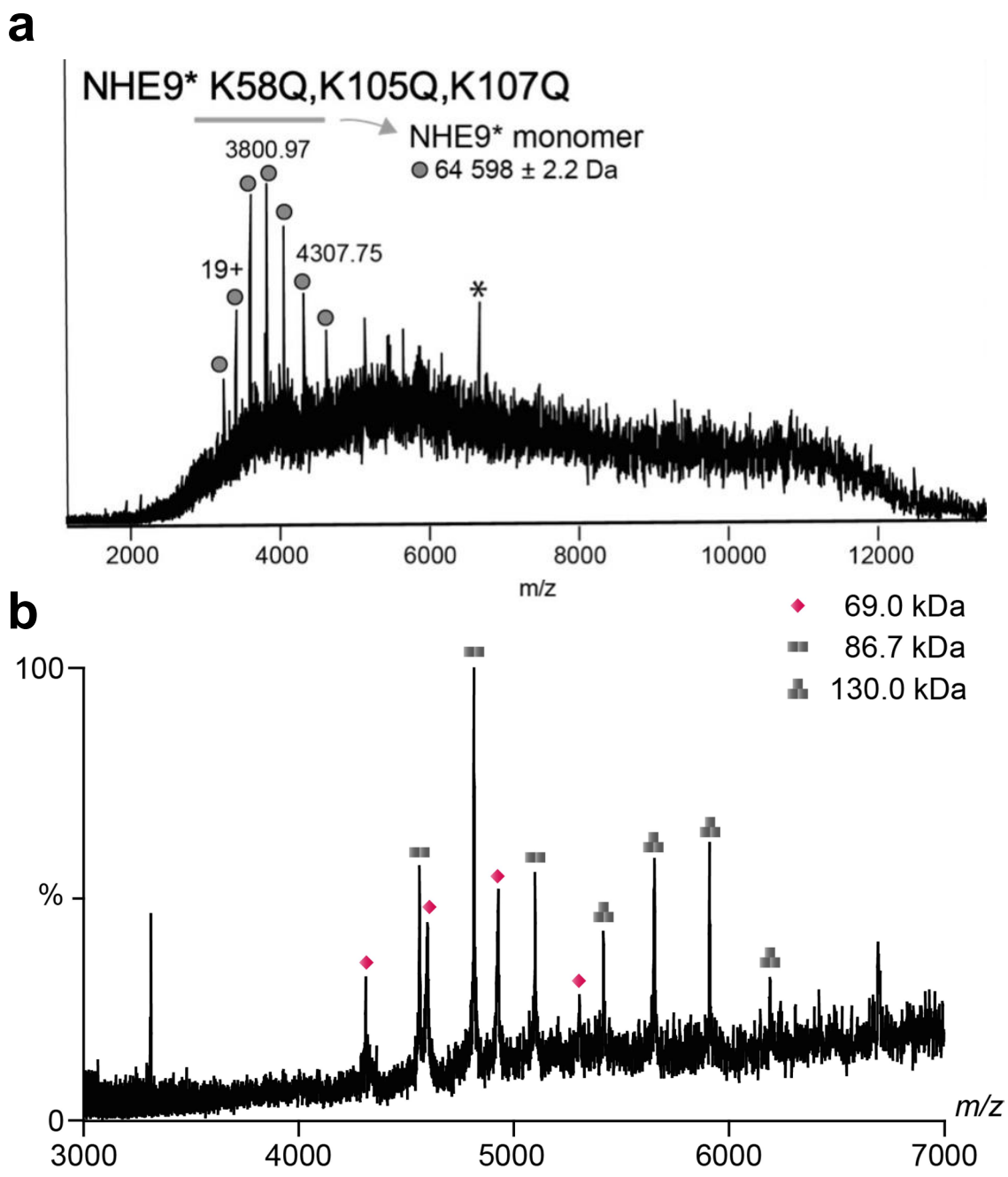


Figure 4.8 Native MS of NHE9 triple mutant. a, Spectrum of the triple mutant of NHE9*, no dimers are observed. The peak labelled with * was assigned to a contaminant. Panel **a** adapted from ref (1). **b,** Spectra produced after the 'no DDM wash', where no NHE9 monomers are observed. The species with a 130 kDa mass is therefore assigned as a contaminant rather than the NHE9* triple mutant dimer.

loop. The observed binding of NHE9* dimers to PIP_2, and the absence of dimers formed by the loop mutant were confirmed by native MS analysis.

A schematic of a proposed mechanism of NHE9 function is shown in **Figure 4.9**. PIP binding facilitates formation of a stable NHE9 dimer (**Figure 4.9, left**). The dimerisation domain then acts as an anchor, while the transporter domain performs the ion transport by large, elevator-like movements (**Figure 4.9, right**). Given this proposed mechanism and the fact that the dimer is the functional unit of NHE9, it is tempting to speculate that PIPs allosterically modulate NHE9 activity by controlling dimerisation. This hypothesis is supported by the known ability of PIPs to activate NHE1 and NHE3[5], as well as a similar role of negatively charged CDL for activity of NhaA[32]. Moreover, the 'NhaA fold' shared by NHE9 is responsible for the formation of a unique dimer interface[13] – implying that this fold could have been 'designed' by evolution specifically to control transporter function by modulation of dimerisation. This hypothesis provides an explanation for NHE9 not forming stable dimers without assistance from the lipids despite a large interfacial area, in contradiction to the current model[15]. It might be possible that in some special cases the overall structure of the oligomeric interface has to be considered when predicting stability of oligomeric states of MPs, rather than just the interfacial area and the number of salt bridges[15].

Overall, the findings presented in this chapter significantly advance the mechanistic understanding of mammalian NHEs. It should be noted that in this chapter I mostly focused on the results related to dimerisation, as this was the area to which I contributed by performing native MS experiments. As a consequence, I have left out many of the impressive insights into NHE9 structure obtained by collaborators, such as the nature of the ion binding site and characterisation of the disease-causing mutation of NHE9 – this information can be obtained from the recent publication[1].

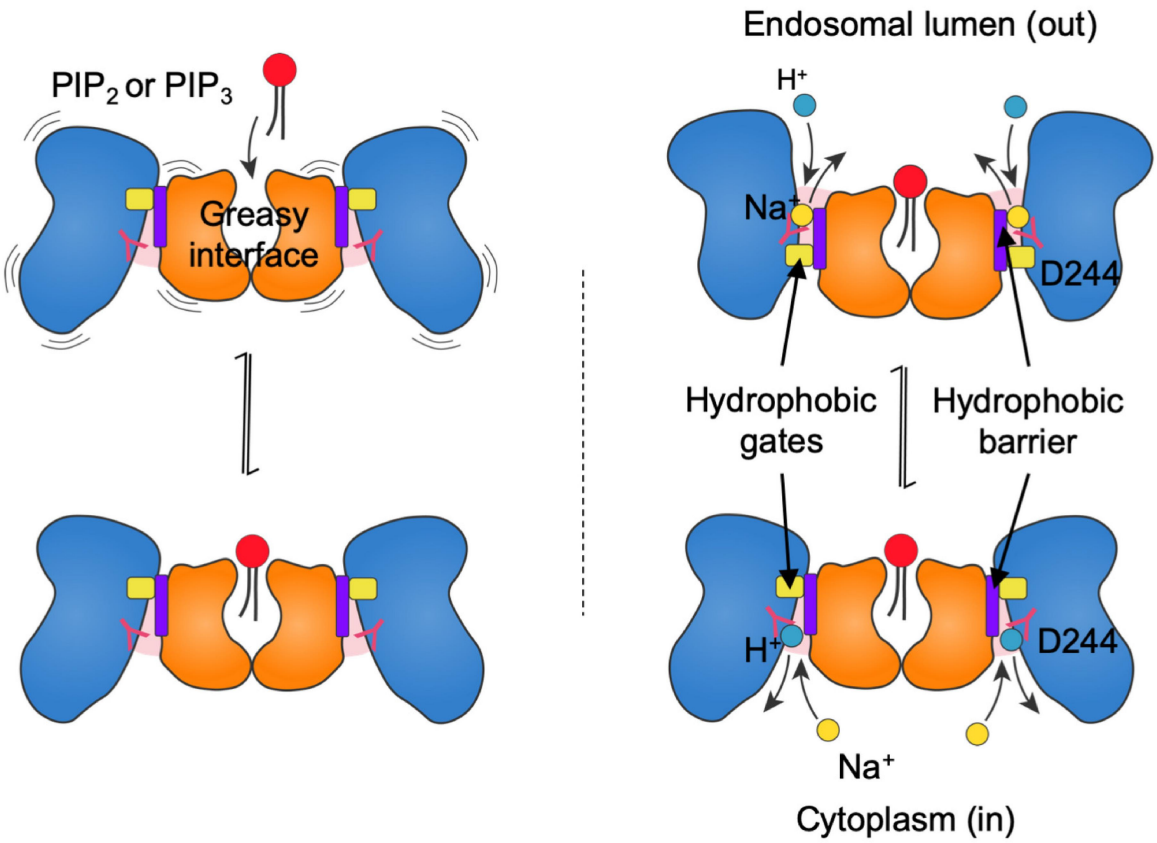


Figure 4.9 NHE9 elevator access mechanism. Schematic representation of the NHE9 transport mechanism, highlighting the role of negatively charged phosphoinositide lipids (PIP_2 or PIP_3) to stabilise homodimerisation. Figure adapted from ref (1).

Materials and Methods

Non-native MS methods, including protein expression and purification details, can be obtained from the recent publication[1].

Sample preparation for native MS. The 'wash' step was applied to some of the samples prior to buffer exchange, as specified in the text. 10 µl of protein was diluted with 90 µl of 200mM AA solution (Sigma Aldrich), containing 2CMC of DDM (Generon). This solution was reconcentrated using a 100K MWCO concentrator. This step was repeated two times, before performing the buffer exchange step.

The buffer exchange was performed using Zeba Spin desalting columns (Thermo Scientific), equilibrated with 200mM AA solution, containing 2xCMC of DDM.

For the lipid binding experiments, aqueous solution of $PIP_2(4,5)$ at 320 μM concentration was diluted 60 times with AA/2xCMC DDM buffer and then added to protein at 1L:4 P volume ratio immediately before native MS analysis.

Native mass spectrometry. Nanoelectrospray ionisation was initiated from borosilicate needles, pulled and gold-coated in house as previously described[23]. All experiments were performed on Q Exactive UHMR (**Figure 1.6d**). The following parameters were used: capillary voltage, 1.4 kV; S-lens RF, 100%; source temperature, 200^{o}C; in-source trapping (IST), 300V; HCD collision energy, 300 V; HCD cell pressure, 1×10^{-9} mbar; collision gas, argon. Data were analysed using the Thermo Excalibur software package.

Acknowledgements

I would like to thank Michael Landreh and David Drew for involving me with this project and for many engaging discussions. I thank Iven Winkelmann for sending me all of the different NHE9 samples. I would like to thank Kin (Hugh) Hoi for providing the aqueous $PIP_2(4,5)$ solution. Finally, I would like to thank all of the collaborators who took part in this project. Contributions: Iven Wilkelmann and Pascal Meier carried out cloning, expression screening and sample preparation for cryo-EM, and experiments for functional analysis. Iven Wilkelmann, Pascal Meier and Rei Matsouka carried out cryo-EM data collection and map reconstruction. Rei Matsuoka and David Drew carried out model building. Laura Orellana, Chenou Zhang, Ricky Sexton and Oliver Beckstein carried out MD simulations and PC analysis.

References

1 Winkelmann, I. *et al.* Structure and elevator mechanism of the mammalian sodium/proton exchanger NHE9. *Embo J*, e105908, doi:10.15252/embj.2020105908 (2020).

2 Orlowski, J. & Grinstein, S. Diversity of the mammalian sodium/proton exchanger SLC9 gene family. *Pflugers Arch* **447**, 549-565, doi:10.1007/s00424-003-1110-3 (2004).

3 Donowitz, M., Ming Tse, C. & Fuster, D. SLC9/NHE gene family, a plasma membrane and organellar family of Na(+)/H(+) exchangers. *Mol Aspects Med* **34**, 236-251, doi:10.1016/j.mam.2012.05.001 (2013).

4 Fuster, D. G. & Alexander, R. T. Traditional and emerging roles for the SLC9 Na+/H+ exchangers. *Pflugers Arch* **466**, 61-76, doi:10.1007/s00424-013-1408-8 (2014).

5 Pedersen, S. F. & Counillon, L. THE SLC9A-C MAMMALIAN Na+/H+ EXCHANGER FAMILY: MOLECULES, MECHANISMS, AND PHYSIOLOGY. *Physiol Rev* **99**, 2015-2113, doi:10.1152/physrev.00028.2018 (2019).

6 Lane, N. Proton gradients at the origin of life. *Bioessays* **39**, doi:10.1002/bies.201600217 (2017).

7 Lane, N. & Martin, W. F. The Origin of Membrane Bioenergetics. *Cell* **151**, 1406-1416, doi:10.1016/j.cell.2012.11.050 (2012).

8 West, I. C. & Mitchell, P. Proton/sodium ion antiport in Escherichia coli. *Biochem J* **144**, 87-90, doi:10.1042/bj1440087 (1974).

9 Murer, H., Hopfer, U. & Kinne, R. Sodium/proton antiport in brush-border-membrane vesicles isolated from rat small intestine and kidney. *Biochem J* **154**, 597-604 (1976).

10 Ueda, M. *et al.* Up-regulation of SLC9A9 Promotes Cancer Progression and Is Involved in Poor Prognosis in Colorectal Cancer. *Anticancer Res* **37**, 2255-2263, doi:10.21873/anticanres.11562 (2017).

11 Brett, C. L., Donowitz, M. & Rao, R. Evolutionary origins of eukaryotic sodium/proton exchangers. *Am J Physiol Cell Physiol* **288**, C223-239, doi:10.1152/ajpcell.00360.2004 (2005).

12 Lee, C. *et al.* A two-domain elevator mechanism for sodium/proton antiport. *Nature* **501**, 573-577, doi:10.1038/nature12484 (2013).

13 Drew, D. & Boudker, O. Shared Molecular Mechanisms of Membrane Transporters. *Annu Rev Biochem* **85**, 543-572, doi:10.1146/annurev-biochem-060815-014520 (2016).

14 Landreh, M. *et al.* Integrating mass spectrometry with MD simulations reveals the role of lipids in Na+/H+ antiporters. *Nat Commun* **8**, 13993, doi:10.1038/ncomms13993 (2017).

15 Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **541**, 421-424, doi:10.1038/nature20820 (2017).

16 Wohlert, D., Kuhlbrandt, W. & Yildiz, O. Structure and substrate ion binding in the sodium/proton antiporter PaNhaP. *Elife* **3**, doi:10.7554/eLife.03579 (2014).

17 Xu, X. *et al.* Structural insights into sodium transport by the oxaloacetate decarboxylase sodium pump. *Elife* **9**, doi:10.7554/eLife.53853 (2020).

18 Bolla, J. R., Agasid, M. T., Mehmood, S. & Robinson, C. V. Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. *Annu Rev Biochem* **88**, 85-111, doi:10.1146/annurev-biochem-013118-111508 (2019).

19 Zhang-James, Y. *et al.* Effect of disease-associated SLC9A9 mutations on protein-protein interaction networks: implications for molecular mechanisms for ADHD and autism. *Atten Defic Hyperact Disord* **11**, 91-105, doi:10.1007/s12402-018-0281-x (2019).

20 Jones, A. J. Y., Gabriel, F., Tandale, A. & Nietlispach, D. Structure and Dynamics of GPCRs in Lipid Membranes: Physical Principles and Experimental Approaches. *Molecules* **25**, doi:10.3390/molecules25204729 (2020).

21 Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172-+, doi:10.1038/nature13419 (2014).

22 Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat Protoc* **8**, 639-651, doi:10.1038/nprot.2013.024 (2013).

23 Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* **2**, 715-726, doi:10.1038/nprot.2007.73 (2007).

24 Gault, J. *et al.* Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. *Nat Methods* **17**, 505-508, doi:10.1038/s41592-020-0821-0 (2020).

25 Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Biol* **19**, 281-296, doi:10.1038/nrm.2017.138 (2018).

26 Nji, E., Chatzikyriakidou, Y., Landreh, M. & Drew, D. An engineered thermal-shift screen reveals specific lipid preferences of eukaryotic and prokaryotic membrane proteins. *Nat Commun* **9**, 4253, doi:10.1038/s41467-018-06702-3 (2018).

27 Abu Jawdeh, B. G. *et al.* Phosphoinositide binding differentially regulates NHE1 Na+/H+ exchanger-dependent proximal tubule cell survival. *J Biol Chem* **286**, 42435-42445, doi:10.1074/jbc.M110.212845 (2011).

28 Aharonovitz, O. *et al.* Intracellular pH regulation by Na+/H+ exchange requires phosphatidylinositol 4,5-bisphosphate. *J Cell Biol* **150**, 213-224, doi:DOI 10.1083/jcb.150.1.213 (2000).

29 McLaughlin, S., Wang, J., Gambhir, A. & Murray, D. PIP(2) and proteins: interactions, organization, and information flow. *Annu Rev Biophys Biomol Struct* **31**, 151-175, doi:10.1146/annurev.biophys.31.082901.134259 (2002).

30 Lee, C. *et al.* Crystal structure of the sodium-proton antiporter NhaA dimer and new mechanistic insights. *J Gen Physiol* **144**, 529-544, doi:10.1085/jgp.201411219 (2014).

31 Bolla, J. R. *et al.* A Mass-Spectrometry-Based Approach to Distinguish Annular and Specific Lipid Binding to Membrane Proteins. *Angew Chem Int Ed Engl* **59**, 3523-3528, doi:10.1002/anie.201914411 (2020).

32 Rimon, A., Mondal, R., Friedler, A. & Padan, E. Cardiolipin is an Optimal Phospholipid for the Assembly, Stability, and Proper Functionality of the Dimeric Form of NhaA Na+/H+ Antiporter. *Sci Rep-Uk* **9**, doi:10.1038/s41598-019-54198-8 (2019).

5 Identifying Specific Protein-Lipid Interactions for Human Leukotriene Biosynthetic Enzymes by Native Mass Spectrometry

Abstract

Cysteinyl leukotrienes (LTs) are established modulators of asthma in humans. They are synthesised by the members of membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG) superfamily, namely LTC_4 synthase (LTC4S) and microsomal glutathione S-transferase 2 (MGST2), making these proteins attractive drug targets. However, their interactions with phospholipids (PLs) have not been characterised, missing potentially essential information about structurally and functionally important lipids. Here, I apply native mass spectrometry (MS) to alleviate this problem.

For LTC4S, I confirm that the expected oligomeric state and the native fold of this protein are preserved in the gas phase with the assistance of ion mobility (IM) spectrometry. I then apply collision-induced unfolding (CIU) to distinguish between specific and non-specific interactions between LTC4S and PLs. The data obtained was in agreement with a solution-phase assay, providing further confidence in the biological relevance of these findings.

For MGST2, I demonstrate that this protein is unstable in the absence of particular endogenous lipids. I employ a mixed detergent micelle to create conditions that are simultaneously compatible with native MS and do not cause MGST2 aggregation. Finally, I identify the presence of certain post-translational modifications (PTMs)

that could shed the light on the nature of the unusual reactivity of one-third of the ligand-binding sites for MGST2.

5.1 Introduction

Leukotrienes (LTs) are inflammation modulators, predominantly found in various types of leukocytes[1]. LTs are synthesised from a common precursor, LTA_4 (which is in turn synthesised from arachidonic acid)[2] (**Figure 5.1a**). LTs can be categorised into two major classes based on the synthetic pathway from LTA_4: one class consists of LTB_4 which is synthesised from LTA_4 by LTA_4 hydrolase (LTA4H); the other class involves three cysteinyl LTs: LTC_4, LTD_4 and LTE_4 - with LTC_4 synthesised by conjugation of LTA_4 and glutathione (GSH) by a membrane protein (MP) called LTC_4 synthase (LTC4S) (**Figure 5.1b**) and then converted into the other cysteinyl-LTs[2,3] (**Figure 5.1a**). Cysteinyl-LTs are associated with several diseases in humans; in particular, they were known for several decades to be established asthma mediators[2,4]. Their role in asthma lead to the development of drugs, often referred to as 'lukasts'[3], that inhibit one of the G protein-coupled receptors (GPCRs) involved in the cysteinyl-LT signalling pathway, namely $CysLT_1$ – however, these drugs are not effective for approximately 40% of the patients[5], most likely because four other GPCRs can also interact with cysteinyl-LTs in complex ways (**Figure 5.1c**)[6]. Therefore, an alternative treatment would be to inhibit the synthesis of LTs, instead of the downstream signalling – LTC4S is a natural target for that approach, as it is responsible for the formation of LTC_4, which is the precursor to all of the cysteinyl-LTs[3,6,7] (**Figure 5.1a**).

LTC4S is a part of membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG) superfamily and particularly of the subfamily including three

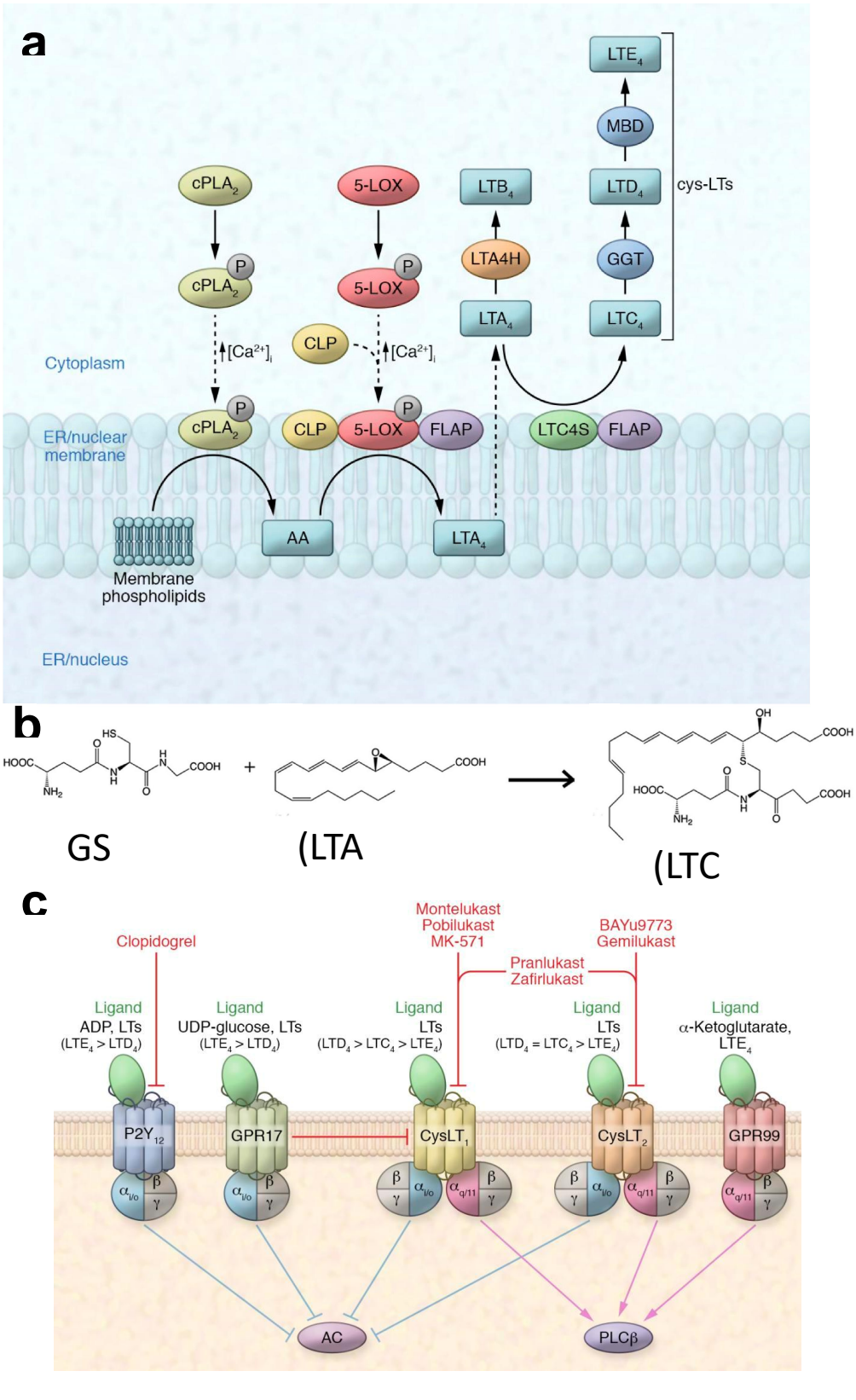


Figure 5.1 Cysteinyl Leukotrienes. a, Synthetic pathway for the formation of leukotrienes (LTs). **b,** The reaction performed by LTC_4 synthase (LTC4S): glutathione (GSH) and LTA4 are joined to form LTC4, the precursor of all cysteinyl LTs (cys-LTs). **c,** Five G protein coupled receptors (GPCRs), involved in LTC4S signalling. Panel **a** adapted from ref (3); panel **c** adapted from ref (6).

trimeric proteins: LTC4S, microsomal glutathione S-transferase 2 (MGST2) and five-lipoxygenase-activating protein (FLAP). As was described above (**Figure 5.1a**), LTC4S is responsible for production of LTC_4 from LTA_4 and GSH; within a cell, it is located in the endoplasmic reticulum and in the outer leaflet of the nuclear envelope[8]. MGST2 is also able to synthesise LTC_4 in a similar way to LTC4S; the main difference being that while LTC4S shows activity for all of its three active sites, MGST2 only shows 1/3 of activity, despite actually having all three of the sites occupied[9]. Finally, FLAP's main role is to activate 5-lipoxygenase (5-LOX) for the synthesis of LTA_4[7], however, evidence of FLAP interacting directly with LTC4S also exists[10]. In this study, I have employed native mass spectrometry (MS) to investigate any effect phospholipids (PLs) might have on the structure or function of LTC4S and MGST2.

5.2 Characterisation of LTC4S by Native MS

(*Note: protein expression and purification were performed by Madhuranayaki Thulasingam, unless stated otherwise in the text. Some of the experiments were disrupted by global pandemic, as specified in the text.)*

I have chosen to use two different mass spectrometers to acquire the data presented here: modified Synapt G1, with a drift tube (DT) ion mobility (IM)-cell (see **Figure 1.6c**) and Q Exactive UHMR (see **Figure 1.6d**). UHMR can be used to acquire mass spectra at high resolution, however, it has no IM compatibility (see **Sections 1.12** and **1.13** for further discussion). The main advantage of the DT is the ability to calculate collisional cross section (CCS) values for proteins without the need for calibration – these values can then be compared to a theoretical CCS

calculated from the crystal structure to determine whether the protein has retained its native-like conformation in the gas phase (**Section 1.13**).

In order to perform native MS experiments, I have exchanged LTC4S from *n*-dodecyl-β-d-maltopyranoside (DDM) detergent into tetraethylene glycol monooctyl ether (C8E4) detergent (see **Materials and Methods**). While DDM is a stabilising detergent in solution, the high energy required to remove it in the gas-phase makes it less suitable for native MS purposes; C8E4, conversely, is a relatively harsh detergent, but is easily removed by relatively gentle gas-phase activation[11-13]. The 'harsh' nature of C8E4 may result in structural rearrangements for some proteins; therefore, the initial experiments were performed with the aim to verify LTC4S stability. A spectrum of LTC4S solubilised in C8E4 is shown in **Figure 5.2**. Importantly, the expected trimeric state of the protein is observed[14], meaning that LTC4S can maintain its native oligomeric state in C8E4 detergent. Monomeric species are also observed at high intensity. Considering the absence of a corresponding dimer, which would be expected for a typical collision-induced dissociation (CID) pattern[15], this was assigned to the presence of some monomeric LTC4S in solution.

Having established the presence of the native oligomeric state of LTC4S, the ability of the protein to maintain a native-like fold was tested next. The crystal structure[14] (PDB: 2UUI) was used to estimate a theoretical CCS value by corrected projected approximation (PA) model (see **Section 1.13**; also **Materials and Methods**). It should be noted that the crystal structure has two extended helical 'arms' (**Figure**

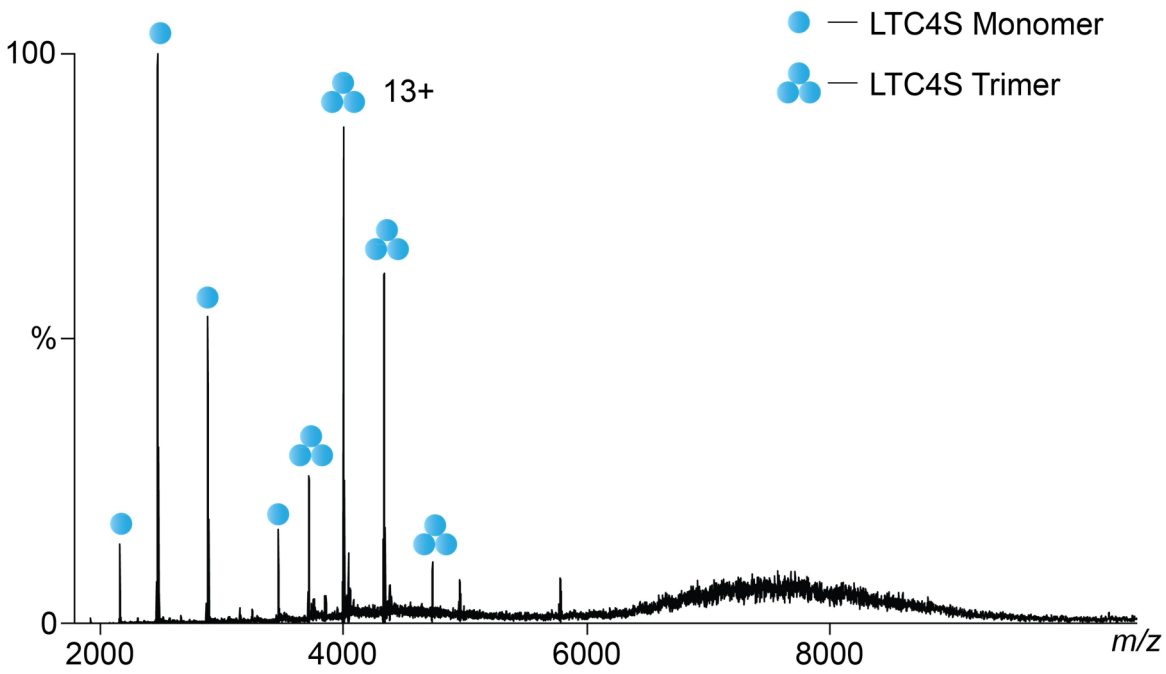


Figure 5.2 Mass spectrum of LTC4S in C8E4 obtained on UHMR. Expected oligomeric state of the trimer is observed, as well as some monomer. The 'hump' at high *m/z* is unassigned.

5.3, **left**) which are extremely unlikely to remain in that position in the absence of a bilayer. CCS calculations are performed on the structure as it is, without prior relaxation[16], so the CCS value produced from this structure can be expected to be slightly too large (was calculated to be 3913 Å^2). In order to also obtain a lower bound for the CCS value, I have generated a structure with the extended 'arms' helices removed (**Figure 5.3**, **right**) using PyMOL software: the CCS value obtained from this structure was calculated to be 3332 Å^2). A reasonable native-like structure of LTC4S would be expected to have CCS somewhere in between these two extreme cases. The experimental value calculated for the 11+ charge state of LTC4S in C8E4 was roughly in the middle of the expected range (3655 Å^2), leading to a conclusion that the protein was likely to maintain a native-like fold.

After confirming the presence of a native-like LTC4S in the gas phase, collision-induced unfolding (CIU) experiments were performed. The theory behind CIU was

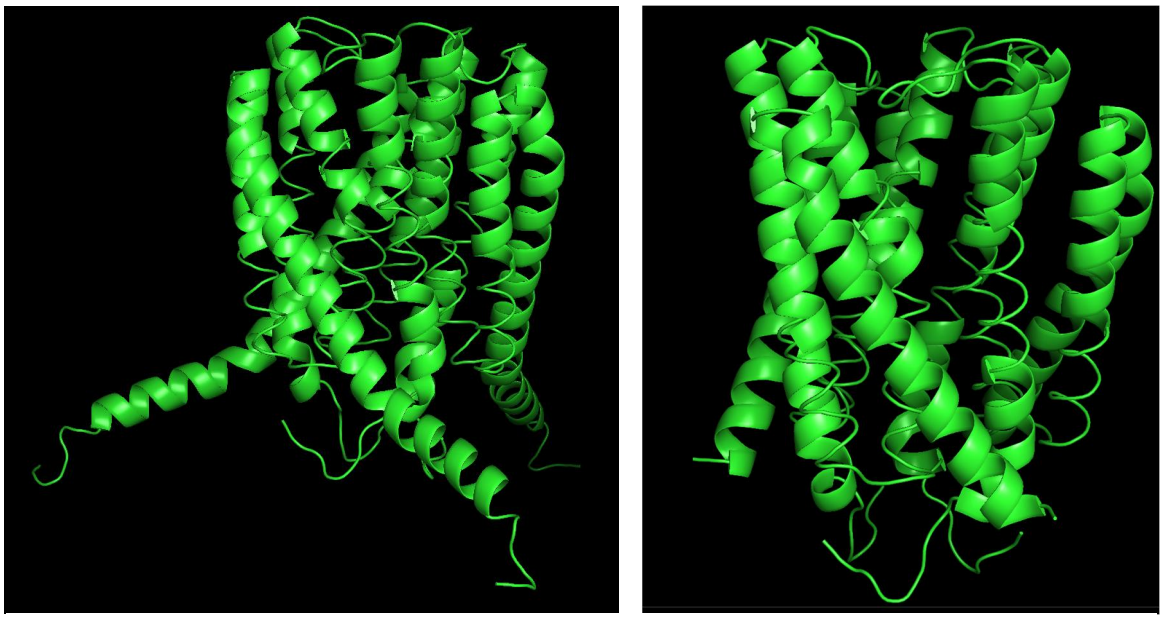


Figure 5.3 Crystal structure of LTC4S. Left, full-length crystal structure (PDB: 2UUI); right, truncated structure, with the two extended helices removed.

discussed in **Section 1.14**, and some practical applications of this method were demonstrated in **Chapter 2**. Briefly, a protein is gradually unfolded in the gas-phase by increasing collision voltage; after a certain threshold in internal energy is reached an unfolding event is observed; the presence of a bound ligand or lipid can induce stabilisation to the protein; stabilisation values can be compared to identify structurally important binding partners[11,17].

Considering the location of LTC4S in the endoplasmic reticulum, I have selected PL classes that would be expected to be naturally interacting with LTC4S in its native environment, namely phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI) and phosphatidylserine (PS)[18]. Some degree of protein stabilisation usually occurs for any species bound, even non-specific[19]; therefore, I have also selected two control lipids, which would not normally be encountered by LTC4S in the native membrane: phosphatidylglycerol (PG) and cardiolipin (CDL)[18]. The lipids can be grouped in terms of their charge: PG, PI, PS and CDL are negatively charged, while PC and PE are neutral. I have also selected

one of the substrates of LTC4S, GSH (the other substrate, LTA_4 is highly unstable (with a half-life of around 10s)[20] and cannot survive on the time scale of a CIU experiment (around 30 minutes for a complete energy ramp).

The unfolding pattern for the 12+ charge state of LTC4S trimer is shown in **Figure 5.4a**. Two distinct unfolding events are clearly visible (black ellipses). Also, a greater stability of PC-bound-LTC4S compared to the *apo* trimer is highlighted for the first transition (red). It should be noted, that the data for the lipid bound and the *apo* unfolding were obtained from the same spectrum (without quadrupole selection) – while performing the experiment this way has a potential drawback of lipids falling off and interfering with the results, the effect was judged to be small for single binding events[21]; whereas the benefits of having both species in the same spectrum include greater intensity, as no ion loss occurs during a quadrupole selection step, and, more importantly, reducing the error due to the variations between runs due to fluctuations of temperature and pressure.

The calculated stabilisation values (see **Materials and Methods**) are shown in **Figure 5.4b** for 1 and 2 lipids bound to the 12+ charge state of LTC4S trimer. Several notable features need to be mentioned:

Firstly, the stabilisation with respect to unfolding induced by the substrate (GSH) is small compared to the PLs. This can be attributed to the nature of a substrate binding pocket that is located at the subunit interface, rather than within individual subunits, and thus providing limited stabilisation with respect to unfolding (see **Chapter 2**, and in particular **Figure 2.1** for further discussion of this phenomenon).

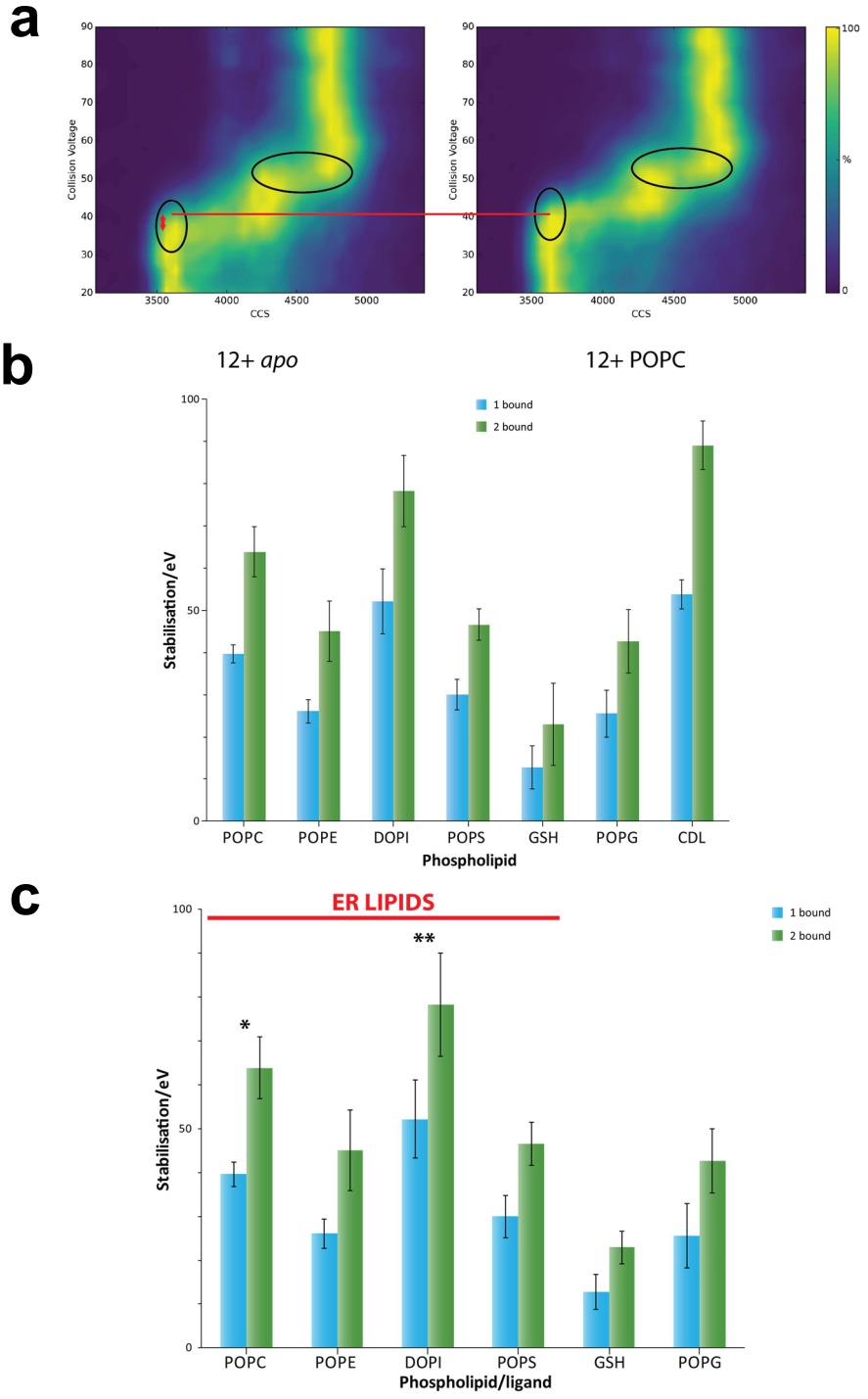


Figure 5.4 Collision-induced unfolding of LTC4S. a, The unfolding profile for *apo* and POPC-bound 12+ LTC4S trimer, with unfolding events highlighted in black and stabilisation difference between *apo* and lipid-bound trimer highlighted in red. Data were acquired at 2V increments. **b,** The average stabilisation for 1 or 2 bound events to LTC4S trimer (12+). **c,** Same as in **b**, with a one-way ANOVA test result shown. * indicates significance at 0.10 confidence level; ** indicated significance at 0.05 confidence level; both compared to the control lipid (POPG) for both 1 and 2 lipids bound (as indicated by the Tukey post hoc test). For panels **b** and **c**, values shown are the average of three repeats performed from different needles, with error bars representing one standard deviation.

Secondly, the two species with the visually largest stabilisation values are PI and CDL, which are both negatively charged (**Figure 1.2a**). The third-highest stabilisation, however, comes from PC, which is zwitterionic (**Figure 1.2a**). A large stabilisation induced by CDL is surprising as it is not expected to be present in the native membrane environment of LTC4S[18]. This could potentially be attributed to the 'double' nature of CDL, which has close to double the mass, the charge and the number of head groups and acyl tails compared to the other PLs (**Figure 1.2a**). In fact, when comparing 1 bound CDL against 2 bound PLs, the values appear more similar, in agreement with this explanation. However, it should be noted that at least one counterexample exists, with data for MscL not showing this effect[11]. Regardless of the exact nature of the observed phenomenon, a very large stabilisation induced by CDL makes it unsuitable for control purposes, so it was not included in further analysis.

Thirdly, PI species had a slightly different hydrophobic tail composition to the other PLs: while all other species have 1-palmitoyl-2-oleoyl (PO, 16:0-18:1) composition, PI has 1,2-dioleoyl (DO, 18:1-18:1) acyl chains. This was only done for practicality, as POPI lipid was not available at the time. While such a small difference in the acyl chain composition is unlikely to account for the large stability of LTC4S trimer in the presence of PI, a control experiment was planned, with POPI and DOPI compared against each other – this experiment had to be postponed due to the global pandemic.

While differences in stabilities between species are visually apparent, the error bars are also relatively large. In order to verify statistical significance of the results discussed in this section, a one-way analysis of variance (ANOVA) test was performed on the lipids shown in **Figure 5.4b**, excluding CDL (**Figure 5.4c**). By

ANOVA (and the Tukey post hoc test) only PI was significantly more stabilising than the control lipid (PG) at 0.05 confidence level; although PC was also found to be significantly more stabilising than PG, but at 0.10 confidence level instead.

To verify the results observed by CIU in the gas phase, a solution-based activity assay was performed (**Figure 5.5**) (*Note: these data were acquired by Madhuranayaki (Madhu) Thulasingam*). Interestingly, the observed activities for the lipid increased in the same order as gas-phase stabilities (the apparent higher activity of the *apo* trimer could be an artefact, with lipids obscuring some of the fluorescence signal just by being in solution). It should be noted that performing the in-solution activity assay proved to be very challenging, with lengthy condition screening required in order to obtain reproducible results (data not shown); whereas CIU experiments can be performed in a reliable way in a matter of days, with a similar outcome. This highlights the power of the native MS-based CIU method when applied to the task of investigating MP-lipid interactions.

Another experiment was planned, investigating simultaneous binding of GSH and PLs to LTC4S by CIU to see whether any allosteric effects could be observed; these studies could not performed because of the lockdown.

5.3 Characterisation of MGST2 by Native MS

Having shown utility of native MS to probe LTC4S-PL interactions, I have attempted performing similar experiments on MGST2. A spectrum of MGST2 after

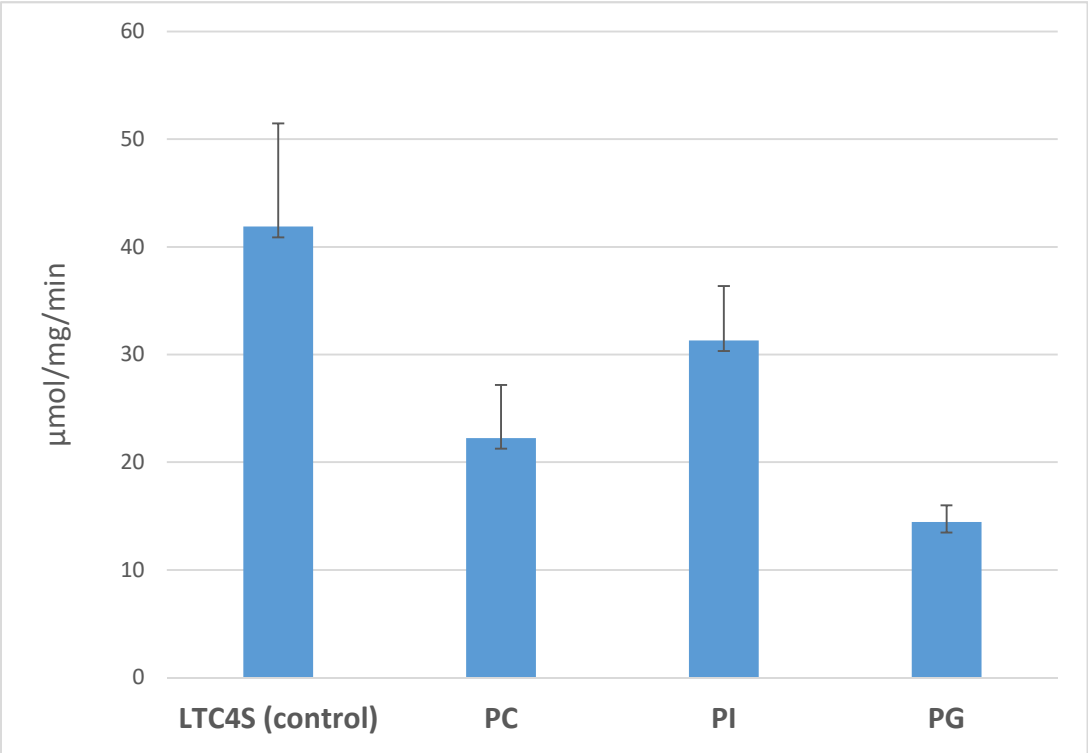


Figure 5.5 Solution-based activity assay of LTC4S. Values shown are the average of three repeats, with error bars representing one standard deviation. Data acquired and figure made by Madhuranayaki Thulasingam.

detergent exchange from DDM into C8E4 is shown in **Figure 5.6a**. While MGST2 also maintains the expected trimeric state, the spectrum is a lot more challenging to interpret due to the presence of a broad adduct. While the mass of the adduct is approximately what would be expected for bound PL species, the exact mass is difficult to quantify due to heterogeneity. Importantly, addition of endogenous lipids to this sample in order to perform CIU experiments would not be possible, as the lipids would simply overlap with the broad endogenous adduct peak. If the energy is increased to attempt removal of the bound species, the dissociation of the trimer is observed (**Figure 5.6b**). An interesting feature of this dissociation is that the dimer mass differs from the two monomer masses by an extra 58 Da. This observation is discussed in more detail later in this section.

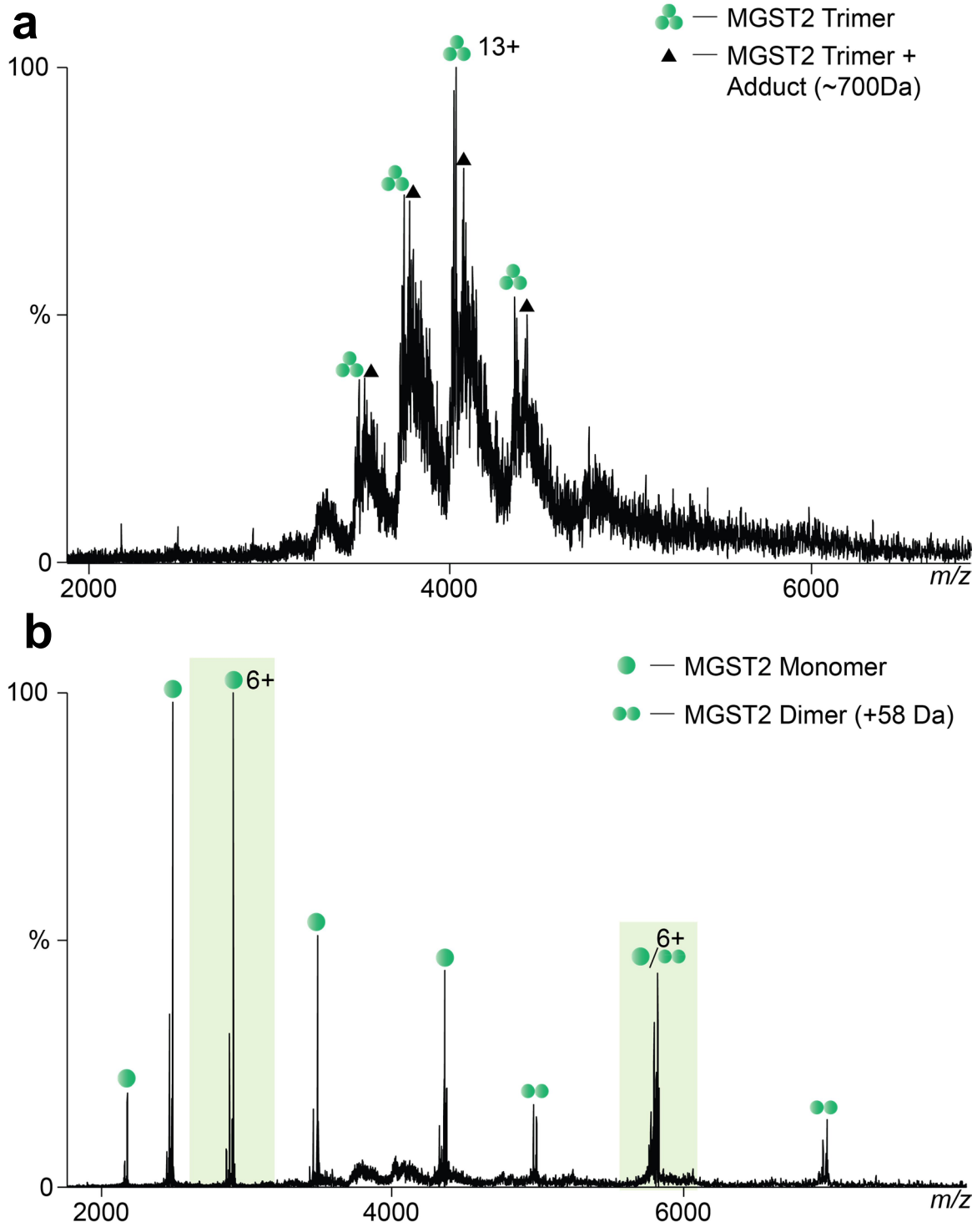


Figure 5.6 Mass spectra of MGST2 in C8E4 obtained on UHMR. a, Expected trimeric state of MGST2 is observed, with a heterogeneous adduct bound. **b,** Applying activation energy results in trimer dissociation, no *apo* trimer is observed. The peaks highlighted in green are further analysed in **Figure 5.7**.

It was determined that the difficulty in calculating exact mass of the adduct peak was not only hindered by a potential heterogeneity of the binding partners, but in

part by MGST2 itself. Some degradation can be observed for a monomer, with either one or two C-terminal amino acids being lost (as well as a few other smaller mass variations) (**Figure 5.7a**). While the amount of degraded protein is relatively small and does not significantly affect the spectral quality for the monomer, different combination of the proteoforms results in a much more convoluted spectrum for the dimer (**Figure 5.7b)** and this problem is exacerbated for the trimer. In order to obtain more interpretable spectra, I have attempted to perform MGST2 purification myself (from the *Pichia Pastoris* cells, containing overexpressed MGST2 and grown by Madhu) using a modified protocol (see **Materials and Methods**). Perhaps the most notable modification to the purification protocol was an introduction of an ultracentrifugation step – this way, all of the soluble proteases are removed from the MP sample, so less degradation can be expected to occur. No cleavage forms were visible in a spectrum of a freshly purified MGST2 monomer (obtained in DDM), confirming this prediction (**Figure 5.7c**).

Remarkably, this new sample was not stable in C8E4, with no MGST2 surviving the detergent exchange. To alleviate this problem, I have replaced C8E4 with a structurally related detergent hexaethylene glycol monooctyl ether (C8E6). This time, well-resolved MGST2 trimer peaks were observed (**Figure 5.8a**). Interestingly, only very minor presence of any adducts was detected for the trimer (**Figure 5.8b**), indicating suitability of this sample for addition of exogenous lipids and CIU experiments. However, MGST2 turned out to be unstable in C8E6 in the long term and a significant drop is spectral quality was observed after a single freeze-thaw cycle (**Figure 5.8c**). Surprisingly, several other attempts of performing detergent exchange from DDM into C8E6 did not work and resulted in the protein crashing out (similarly to the C8E4 case described above). I then decided to test

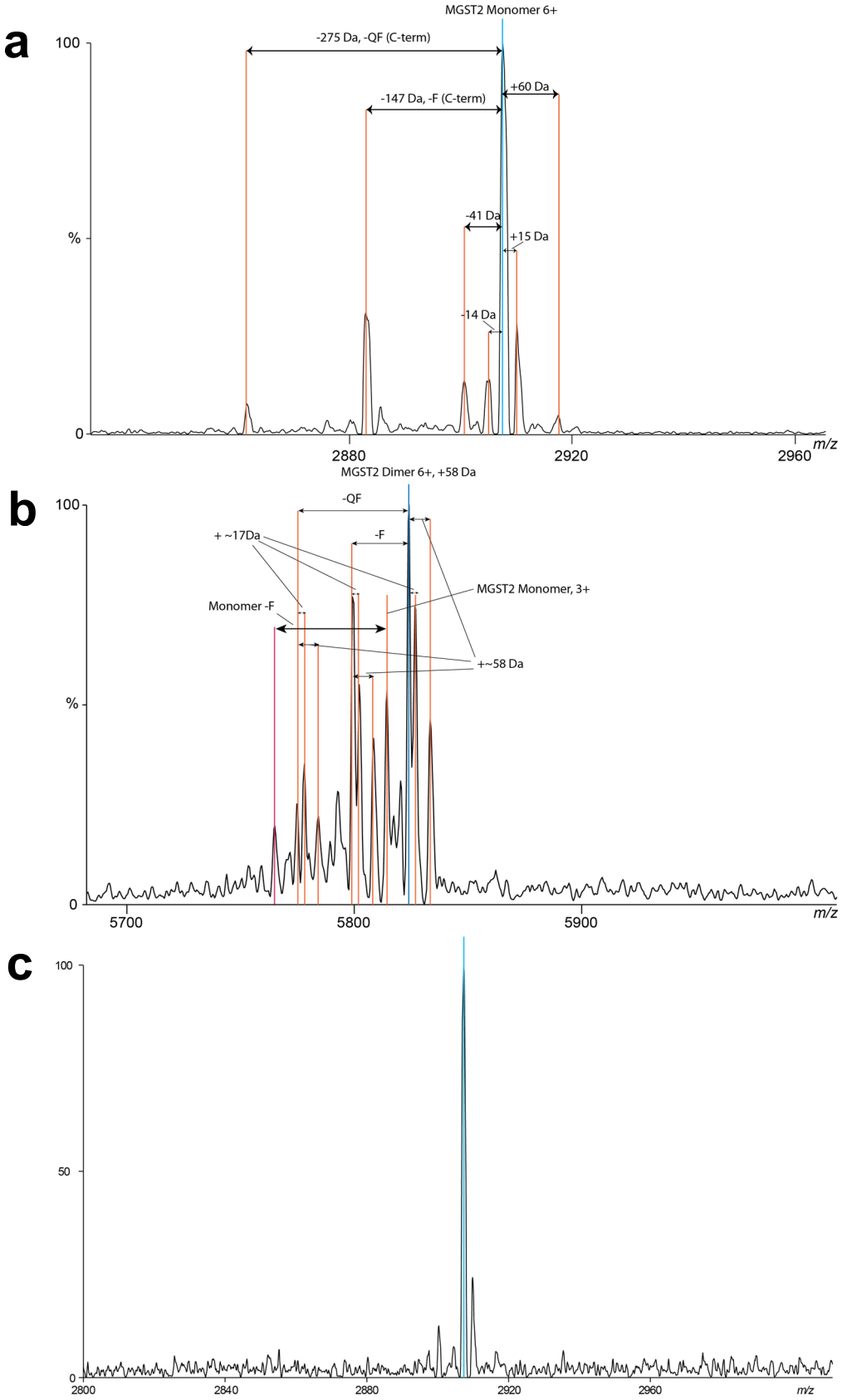


Figure 5.7 Heterogeneity of MGST2. a, Different proteoforms present for the MGST2 monomer. **b,** Different proteoforms present for the MGST2 dimer. **c**, Different proteoforms present for the MGST2 monomer (new purification).

the utility of one of oligoglycerol detergents (OGDs), namely [G1]-mix ([G1]-ether-C12 with regioisomers in the headgroup[22]). Interestingly, MGST2 was stable in this detergent, however only monomers and dimers were observed in the spectra (**Figure 5.9**). The masses of these species were again very interesting. There are two dimer masses, one with one extra 57-58 Da in mass (as was observed earlier, **Figure 5.6b**) and one with two extra such units. No dimer with no extra mass exists. At the same time, the vast majority of monomer is at a single mass value, with a very minor population of the extra 57-58 Da. Considering the fact that a similar pattern was produced from two separate samples solubilised in different detergents, this appears to be the common pattern for MGST2. From the data, it is tempting to speculate that each trimer contains two 'heavy' units (probably modified with some post-translational modification (PTM)) and one 'light' unit; also the light unit appears to be detaching in preference to one of the heavy ones, possibly indicating the interfacial nature of such modification. While this is only a speculation, it would be interesting to characterise the nature of those PTMs (if they are actually PTMs) by proteomics. If this finding will be confirmed by a direct observation, it could potentially shed more light on how the activity of MGST2 is controlled to only involve one out of the three active sites[9].

Finally, I have addressed the puzzling nature of MGST2 exchange into C8E6, which worked well once but never again. I hypothesised that this might imply that incomplete detergent exchange happened to occur during the first purification, allowing MGST2 to retain some stability. This would also be in agreement with the

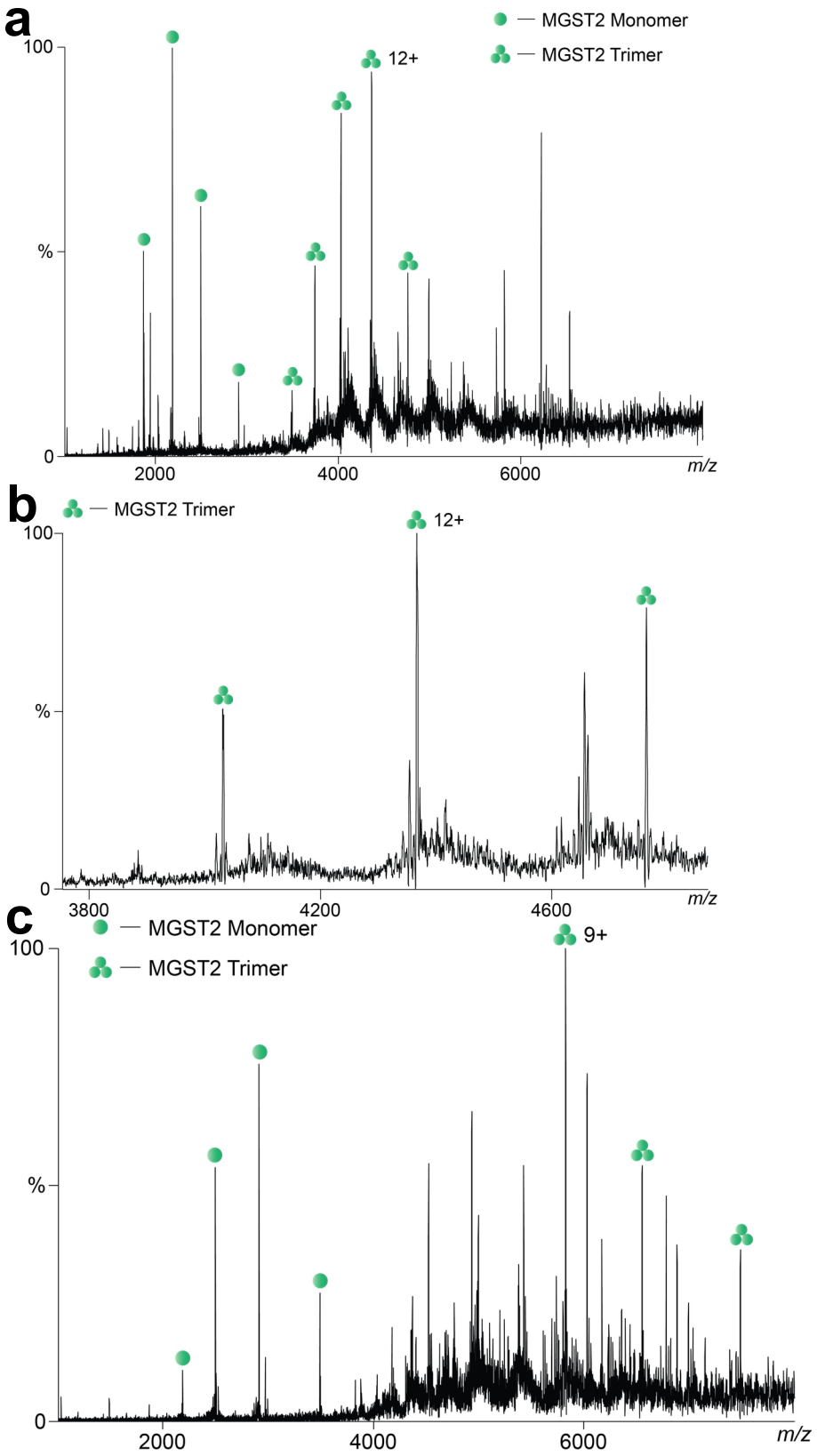


Figure 5.8 Mass spectra of MGST2 in C8E6 obtained on UHMR. a, A spectrum observed before freeze-thaw. **b,** Zoomed-in view on the trimer, only minor amounts of adducts are visible. **c**, A spectrum observed after freeze-thaw, spectral quality deteriorated significantly.

observation that MGST2 was stable in C8E4 before my purification and not after, as I have lowered the DDM concentration in the sample in order to improve spectral quality. By this reasoning, performing incomplete detergent exchange over a buffer exchange column (biospin, BS), instead of by using size-exclusion chromatography (SEC) could lead to the formation of a stable MGST2 trimer. To test this hypothesis, I performed a partial detergent exchange from DDM into C8E4 (**Figure 5.10a**). After 1 BS, MGST2 trimer was resolvable, but the spectrum was dominated by a large number of DDM clusters. After 2 BSs (**Figure 5.10b**), a clearly resolved MGST2 trimer was visible. It should be noted that a broad binding peak was still observed. In fact, resolved trimers were not obtainable on a Synapt instrument by this technique, resulting in wide, unresolved humps instead (data not shown). Overall, the results of this section suggest that a specific bound partner, most likely lipid, is required for MGST2 stability, as some amount of bound species is observed for the trimer in all spectra, and where the peak intensity is low MGST2 is unstable, as was the case for C8E6 (**Figure 5.8c**) .

Unfortunately, I was unable to characterise MGST2 further because of the lockdown. Even if Synapt resolution will turn out not to be high enough to resolve the trimeric peaks, meaning that CIU is impossible, some further experiments can be performed on UHMR. A more detailed characterisation of the bound species can be achieved. One interesting experiment would be to add exogenous lipids to the sample one by one, to see if the trimer intensity would increase significantly for

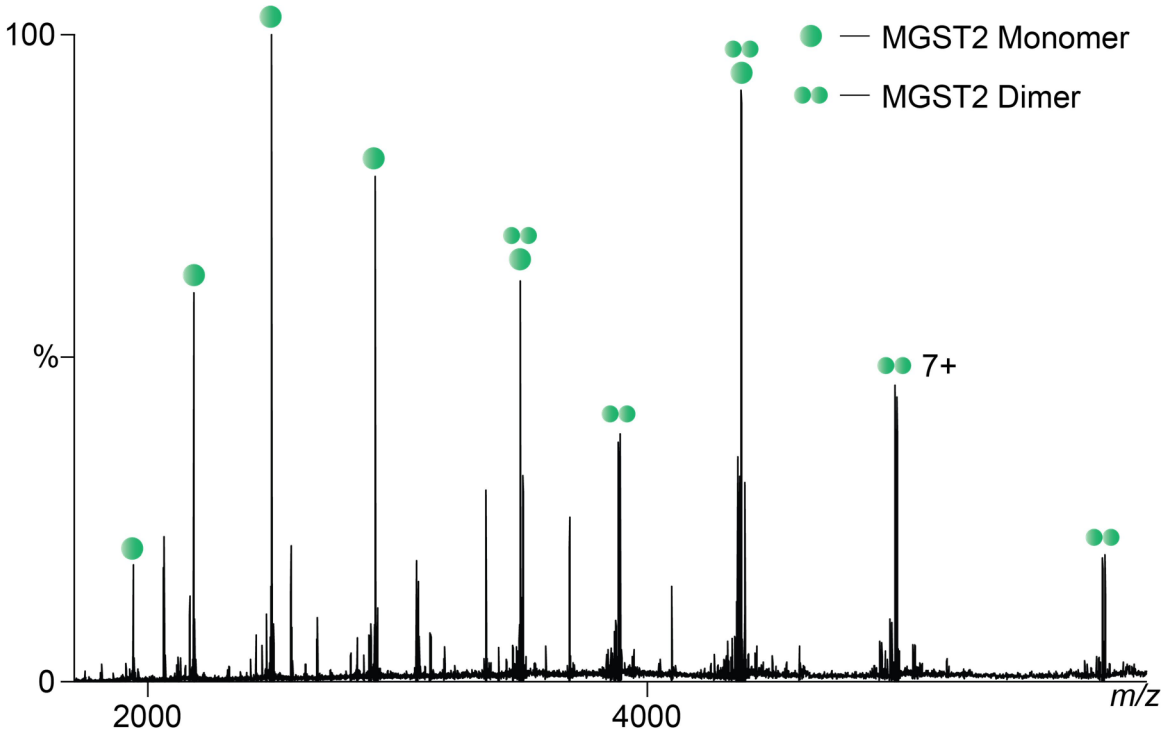


Figure 5.9 Mass spectra of MGST2 in [G1]-mix obtained on UHMR. Monomers and dimers of MGST2 are clearly visible, but no trimers can be observed. Unlabelled peaks were assigned to a contaminant.

any of them – this could allow indirect identification of the proposed specific lipid required for MGST2 stability.

Summary and Conclusions

In this chapter, I employed native MS to investigate two members of the MAPEG superfamily: LTC4S and MGST2. For LTC4S, I showed that the protein maintained both the expected oligomeric state and a native-like fold after being solubilised in C8E4 detergent and transferred to the vacuum of a mass spectrometer. I then exogenously added PL species that are expected to be present in the native membrane of LTC4S and monitored unfolding stabilisation by CIU. PI lipid resulted in the largest stabilisation, followed by PC – this order was then also observed by an activity assay in solution.

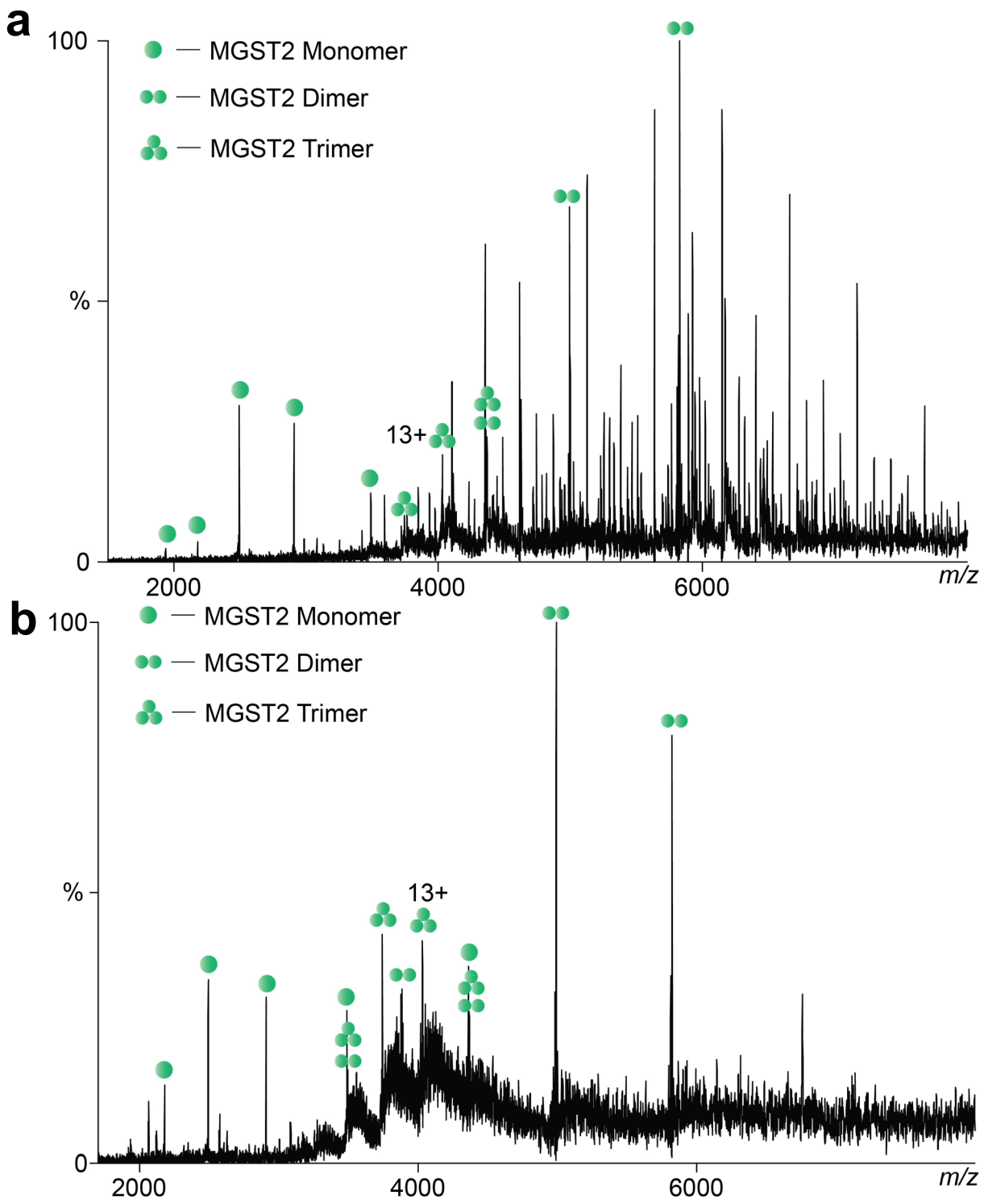


Figure 5.10 Mass spectra of MGST2 in DDM-C8E4 mixed micelle. a, Spectrum of MGST2 after 1 biospin into C8E4 buffer, unlabelled peaks were assigned to DDM clusters. **b,** Spectrum of MGST2 after 2 biospins into C8E4 buffer.

MGST2 was a lot more challenging to work with than LTC4S. After trying several detergent exchange strategies, I identified mixed micelles as a compromise

method, both protecting MGST2 in solution and enabling native MS. Every time MGST2 trimer was observed in a spectrum, a bound-state was also present (although sometimes at low intensity). Finally, mass analysis revealed that the trimeric subunits of MGST2 are not all identical, but differ by the presence or an absence of a certain modification. This discovery could potentially be linked to the 1/3 active site reactivity of MGST2.

Materials and Methods

Membrane protein expression and purification. LTC4S and MGST2 were expressed in *Pichia Pastoris* cells and purified as previously described[23,24] (by Madhu). The purification protocol used to create a more homogenous MGST2 sample was performed as follows:

The cell pellets corresponding to 2L of expression were resuspended in lysis buffer (100mM Tris pH7.8, 100mM KCl, 10% glycerol, supplemented with two protease inhibitor cocktail tablets (Roche)) and lysed by 10-11 passes at 30k pressure using M-110 PS Microfluidiser (Microfluidics). After pelleting the cell debris at 20000g for 20 min, membranes were collected by ultracentrifugation at 100000g for 2h, resuspended in membrane resuspension buffer (25 mM Tris, 100mM NaCl, 20% Glycerol, pH 7.8, supplemented with 1 protease inhibitor cocktail table) and frozen and stored at -80^{o}C.

On the next day, the membranes were thawed out and the proteins were extracted in 0.5% w/v deoxycholine and 1% w/v Triton X-100 for 3 hours with gentle agitation at 4^{o}C. Centrifuged at 20000g for 45 min. This was followed by purification by immobilised metal affinity chromatography (IMAC) on Akta Start, using a 5-ml His-Trap column. Proteins were washed with 20 column volumes (CVs) of buffer A (25

mM Tris (pH 7.8), 0.5M NaCl, 10% glycerol, 0.02% w/v DDM, 5 mM β-mercaptoethanol (BME), 20 mM imidazole), then by 20 CVs of 7% buffer B (same as Buffer A, but with 300 mM imidazole), then eluted with 100% buffer B. Washed three times with SEC buffer (25mM Tris (pH 8.0), 100mM NaCl, 10% glycerol and 0.02% DDM) on 50k MWCO concentrator, then did a SEC on 24-ml column into SEC buffer.

Sample preparation for native MS. Complete detergent exchange was performed on a 24-ml SEC column equilibrated with 200 mM ammonium acetate (AA)/2xCMC detergent of interest. Partial exchange was performed using Zeba Spin desalting columns (Thermo Scientific), equilibrated with 200mM AA/2xCMC C8E4 solution once or twice as indicated in the text.

Native MS. Nanoelectrospray ionisation was initiated from borosilicate needles, pulled and gold-coated in house as previously described[25].

Synapt G1 data. The following parameters were used: Capillary voltage: 1.4 kV, Sampling cone: 200 V, Extraction cone: 1 V, Trap CE: 20-90 V (for CIU experiments), Trap gas flow (Argon): 10 ml/min, IMS gas flow (Helium): 60 ml/min, Trap WV: 100, Trap WH: 1.0, Transfer WV: 100, Transfer WH: 10.0, Backing Pressure: 5 mbar. To calculate the CCS of LTC4S, drift voltages from 40 V to 75 V were used.

UHMR data. The following parameters were used: Capillary voltage: 1.2 kV, Capillary temperature: 30 ºC, S-lens RF level: 200, in-source trapping (IST): 50 V, HCD: 70-120 V.

Data analysis. CIU experiments. Data were processed with PULSAR software as previously described[17]. The final conversion of data into a model was performed using an in-house python script (provided by Tim Allison).

CCS calculation. LTC4S CCS was calculated from the drift-tube IM data using PULSAR software[17]. IMPACT software[16] was used to calculate CCS values from crystal structures using projection approximation (PA), which were then corrected using equation (1.6):

$$CCS_{calc} = 1.14 CCS_{PA} \left(\frac{M_{exp}}{M_{PDB}}\right)^{\frac{2}{3}},$$

Where M_{exp} is the molecular weight of the protein used in the experiment and M_{PDB} is the molecular weight of the protein in the crystal structure.

ANOVA testing. One-way ANOVA tests were performed in Origin Pro software. Assumption of normality was verified by Shapiro-Wilks test (at 0.05 significance level); assumption for homogeneity of variance was verified by Levene's test (at 0.05 significance level). Post hoc Tukey HSD test was used (at 0.05 significance level or 0.10 significance level) to identify differences compared to the control lipid.

Activity assay. Detailed protocol for activity assay can be obtained from Madhuranayaki Thulasingam.

Acknowledgments

I would like to thank Madhuranayaki (Madhu) Thulasingam and Jesper Haeggstrom for involving me in this fascinating project. Madhu expressed and purified all of the proteins (apart from where stated otherwise in the text), and also performed the solution-based assay on LTC4S in the presence of lipids. I would

like to thank Joseph Gault for many engaging discussions. Finally, I would like to thank Leonhard Urner for providing [G1]-mix detergent.

References

1 Haeggstrom, J. Z. & Wetterholm, A. Enzymes and receptors in the leukotriene cascade. *Cell Mol Life Sci* **59**, 742-753, doi:10.1007/s00018-002-8463-1 (2002).

2 Haeggstrom, J. Z. & Funk, C. D. Lipoxygenase and leukotriene pathways: biochemistry, biology, and roles in disease. *Chem Rev* **111**, 5866-5898, doi:10.1021/cr200246d (2011).

3 Haeggstrom, J. Z. Leukotriene biosynthetic enzymes as therapeutic targets. *J Clin Invest* **128**, 2680-2690, doi:10.1172/Jci97945 (2018).

4 Samuelsson, B. Leukotrienes: mediators of immediate hypersensitivity reactions and inflammation. *Science* **220**, 568-575, doi:10.1126/science.6301011 (1983).

5 Malmstrom, K. *et al.* Oral montelukast, inhaled beclomethasone, and placebo for chronic asthma. A randomized, controlled trial. Montelukast/Beclomethasone Study Group. *Ann Intern Med* **130**, 487-495, doi:10.7326/0003-4819-130-6-199903160-00005 (1999).

6 Yokomizo, T., Nakamura, M. & Shimizu, T. Leukotriene receptors as potential therapeutic targets. *J Clin Invest* **128**, 2691-2701, doi:10.1172/Jci97946 (2018).

7 Thulasingam, M. & Haeggstrom, J. Z. Integral Membrane Enzymes in Eicosanoid Metabolism: Structures, Mechanisms and Inhibitor Design. *J Mol Biol* **432**, 4999-5022, doi:10.1016/j.jmb.2020.07.020 (2020).

8 Niegowski, D. *et al.* Structure and inhibition of mouse leukotriene C4 synthase. *PLoS One* **9**, e96763, doi:10.1371/journal.pone.0096763 (2014).

9 Ahmad, S. *et al.* Trimeric microsomal glutathione transferase 2 displays one third of the sites reactivity. *Biochim Biophys Acta* **1854**, 1365-1371, doi:10.1016/j.bbapap.2015.06.003 (2015).

10 Strid, T. *et al.* Distinct parts of leukotriene C(4) synthase interact with 5-lipoxygenase and 5-lipoxygenase activating protein. *Biochem Biophys Res Commun* **381**, 518-522, doi:10.1016/j.bbrc.2009.02.074 (2009).

11 Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172-+, doi:10.1038/nature13419 (2014).

12 Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat Protoc* **8**, 639-651, doi:10.1038/nprot.2013.024 (2013).

13 Reading, E. *et al.* The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase. *Angew Chem Int Edit* **54**, 4577-4581, doi:10.1002/anie.201411622 (2015).

14 Martinez Molina, D. *et al.* Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase. *Nature* **448**, 613-616, doi:10.1038/nature06009 (2007).

15 Benesch, J. L. P. Collisional Activation of Protein Complexes: Picking Up the Pieces. *J Am Soc Mass Spectr* **20**, 341-348, doi:10.1016/j.jasms.2008.11.014 (2009).

16 Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J. & Benesch, J. L. Collision cross sections for structural proteomics. *Structure* **23**, 791-799, doi:10.1016/j.str.2015.02.010 (2015).

17 Allison, T. M. *et al.* Quantifying the stabilizing effects of protein-ligand interactions in the gas phase. *Nat Commun* **6**, doi:10.1038/ncomms9551 (2015).

18 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, doi:10.1038/nrm2330 (2008).

19 Landreh, M., Costeira-Paulo, J., Gault, J., Marklund, E. G. & Robinson, C. V. Effects of Detergent Micelles on Lipid Binding to Proteins in Electrospray Ionization Mass Spectrometry. *Anal Chem* **89**, 7425-7430, doi:10.1021/acs.analchem.7b00922 (2017).

20 Niegowski, D. *et al.* Crystal structures of leukotriene C4 synthase in complex with product analogs: implications for the enzyme mechanism. *J Biol Chem* **289**, 5199-5207, doi:10.1074/jbc.M113.534628 (2014).

21 Liu, Y., Cong, X., Liu, W. & Laganowsky, A. Characterization of Membrane Protein-Lipid Interactions by Mass Spectrometry Ion Mobility Mass Spectrometry. *J Am Soc Mass Spectrom* **28**, 579-586, doi:10.1007/s13361-016-1555-1 (2017).

22 Urner, L. H. *et al.* Modular detergents tailor the purification and structural analysis of membrane proteins including G-protein coupled receptors. *Nat Commun* **11**, 564, doi:10.1038/s41467-020-14424-8 (2020).

23 Ahmad, S. *et al.* Catalytic characterization of human microsomal glutathione S-transferase 2: identification of rate-limiting steps. *Biochemistry-Us* **52**, 1755-1764, doi:10.1021/bi3014104 (2013).

24 Rinaldo-Matthis, A. *et al.* Arginine 104 Is a Key Catalytic Residue in Leukotriene C-4 Synthase. *J Biol Chem* **285**, 40771-40776, doi:10.1074/jbc.M110.105940 (2010).

25 Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* **2**, 715-726, doi:10.1038/nprot.2007.73 (2007).

Concluding Remarks

In this thesis, I have applied native mass spectrometry (MS) to obtain information about membrane protein (MP)-lipid interactions. I have started with describing method developments to improve the current toolkit of this gas-phase technique (**Chapters 2** and **3)**. I have then moved on to apply more traditional native MS techniques to characterise biologically relevant protein systems (**Chapters 4** and **5**).

In **Chapter 2**, I have presented for the first time the ability of native MS combined with surface-induced dissociation (SID) to identify interfacial lipids that stabilise MP oligomers. For the example of SemiSWEET (SS), I was able to show the difference between cardiolipin (CDL), which significantly stabilises the dimeric form of the protein, and other phospholipids (PLs). Moreover, I demonstrated the sensitivity of this method to observe very minor differences in interfacial stabilisation energies of SS dimers when bound to CDLs containing different chain lengths. The data that I have acquired are in agreement with previous observations, indicating biological relevancy of the obtained results. In addition, I introduce a new way to use statistical analysis, specifically for the experiments where subtle differences are expected to occur. Finally, I describe several potential practical problems that can arise during the native SID experiments, as well as a troubleshooting guide on how to fix them.

In **Chapter 3**, I develop a novel procedure for a controllable delipidation of MPs for subsequent native MS characterisation. I demonstrate a set of seven detergents, four of which are introduced for the first time in this work, which have gradual changes in the delipidation propensities. I describe a standardised method,

applicable to example MPs of various molecular weights and oligomeric states, where one of the detergents from the set is used during the last purification stage to predictably control the amount of remaining MP-PL interactions. I also show that a different procedure can be used to remove lipopolysaccharide (LPS) from MPs, in order to produce well-resolved and easily interpretable mass spectra.

I also investigated more generally the properties of detergents that control their delipidation abilities. I discovered that the delipidation strengths of detergents are not dependent on the exact chemical structure, but rather on physical parameters, such as hydrophobicity and shape. The charge reducing properties of detergents, were, however, dependent on the chemical identity of the polar head group. These results can be used as guidelines for rational design of novel detergents in the future.

In **Chapter 4**, I investigated a member of Na^+/H^+ exchangers (NHE) family, NHE9. NHEs are important ion transporters, with dysfunctions associated with numerous diseases in humans. I have employed native MS to complement the high resolution structure obtained by cryogenic electron microscopy (cryo-EM), acquiring information about the effect of lipid binding on the NHE9 dimerisation. No *apo* dimers were observed in the mass spectra, implying that lipids are essential for formation of the functionally relevant oligomeric state of NHE9. By combining native MS data with mutations and thermal stabilisation in solution, both the nature of the specific lipids (PIP_2 and PIP_3) and the location of the binding site at the dimer interface were identified.

Finally, in **Chapter 5**, I apply native MS analysis to investigate lipid interaction for the members of membrane-associated proteins in eicosanoid and glutathione

metabolism (MAPEG) superfamily, namely leukotriene C_4 synthase (LTC4S) and microsomal glutathione S-transferase 2 (MGST2). MAPEG proteins are potential drug targets, due to their role in synthesis of cysteinyl-LTs that are known inflammation mediators in asthma. I have applied collision-induced unfolding (CIU) activation to LTC4S in the presence of lipid to identify specific stabilisation of that protein by specific PLs. The results obtained by this gas-phase method were in excellent agreement with a follow-up in-solution assay. I also demonstrate the instability of MGST2 when solubilised in detergent micelles in the absence of endogenous lipids.