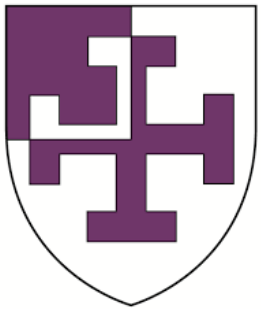
The Application of Mass Spectrometry in the Study of Protein-Lipid Interactions

Jingwen Li

University of Oxford

St Cross College



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

Michaelmas 2019

Abstract

Interactions between membrane proteins and their surrounding lipids play an integral role in coordinating membrane function. However, capturing and differentiating key lipid binding to membrane proteins within the complex and the dynamic membrane environment remains challenging for existing methods. Recent developments in native mass spectrometry have enabled non-covalent interactions of protein complexes to be maintained in the gas phase allowing critical membrane protein-lipid interactions in biological systems to be probed. In this thesis, native mass spectrometry is used throughout, sometimes coupled with liquid chromatography-mass spectrometry based lipidomics, to interrogate membrane proteins and their natural lipid environments.

A protocol to control the degree of delipidation from membrane proteins was developed (Chapter 2). The resulting mass spectra revealed the target membrane proteins, MsbA and MdtK, in different lipid bound states, enabling the relationship between membrane proteins and neighbouring lipids to be observed. Next MS was used to characterise a novel detergent-free particle known as a lipodisq (Chapter 3). Lipodisqs encapsulate membrane proteins in their original lipid environment. By optimising lipodisq preparation and MS methodology, well-resolved MS spectra of rhodopsin and MdtK with bound lipids from their lipodisq preparations were obtained. Finally, MS was used in conjunction with cryo-EM to investigate a respiratory protein supercomplex (Chapter 4). MS not only provided mass information which assisted in the identification of the protein subunits and binding lipids, but also detected the non-covalent interactions between protein and other ligand molecules in the supercomplex system. Overall, the research reported in this thesis has contributed to

broadening the applications of mass spectrometry in the emerging field of protein-lipid interactions.

Table of contents

1 Introduction

1.1 Membrane proteins, lipids and their interactions

1.1.1 The importance of membrane proteins

Biological membranes separate cell components and control the exchange of substances into and out of the cell. They are primarily composed of fluid lipid molecules and embedded proteins which give them their selective permeability. It is the arrangement of lipid and protein molecules which allows the plasma membrane to fulfil its normal function[1].

There are two main categories of membrane proteins: peripheral and integral membrane proteins (Figure 1-1 A). Peripheral membrane proteins reside on membranes transiently and are easily removed from the membrane via changes in ionic strength or pH. Conversely, integral membrane proteins reside permanently in membranes with removal requiring a structural disruption of the lipid environment using chemical reagents for molecule isolation[2,3]. These membrane proteins have evolved architectures to perform diverse functions essential for life, including molecular transport, signal transduction, intercellular adhesion, cell-cell communication, attachment to the cytoskeleton and extracellular matrix, and catalysis (Figure 1-1 B)[1]. Therefore, membrane proteins are an important target for biological research and therapeutic design, and currently over 50% of modern drugs target membrane proteins[4-6].

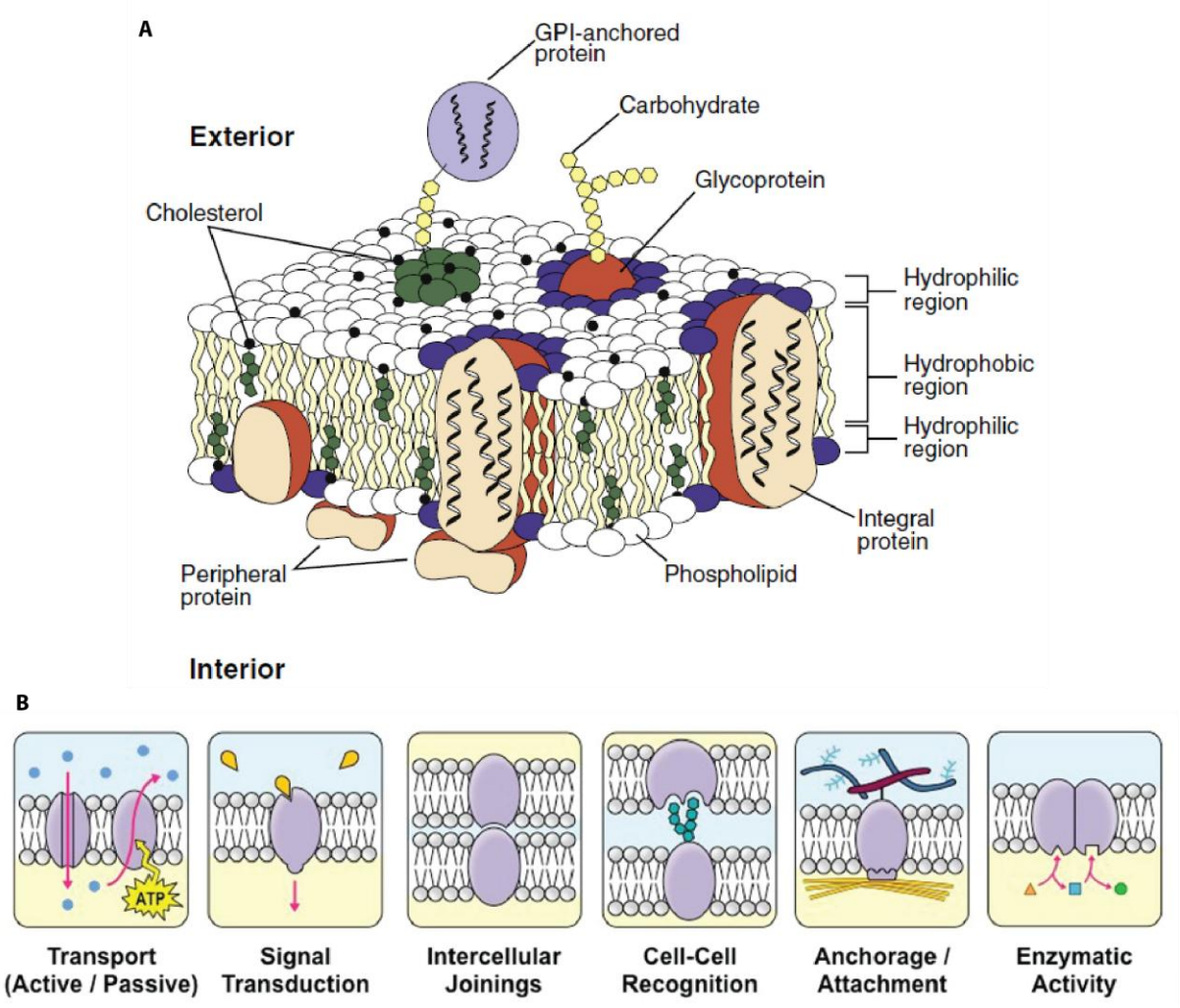


Figure 1-1 Proteins in biological membranes. A, representative model for the plasma membrane structure of a eukaryotic cell. The membrane is formed by two layers of phospholipids with hydrophilic heads on the exterior and hydrophobic tails on the interior, containing a variety of proteins and lipids. Peripheral membrane proteins attach to the hydrophilic surface of lipids or other membrane proteins, while integral membrane proteins span the membrane with their hydrophobic transmembrane domains. Lipids are dynamic in the membrane, but their diffusion and distribution are influenced by protein-protein, protein-lipid and lipid-lipid interactions. Figure is from Dawhan *et al.*[7]. **B,** Some functions of membrane proteins. In many cases, a single protein performs multiple tasks. Figure is from Campbell Biology[1].

Although membrane proteins are encoded by approximately 30% of the human genome and 20% of the *Escherichia coli* (*E. coli*) genome[8-11], in contrast to soluble proteins, there are relatively few solved structures for membrane proteins[9], which

limits the development of therapeutics. There are many factors which contribute to the difficulties in elucidating membrane proteins: firstly, the low level of natural abundance is insufficient for characterisation[3,12-14]; secondly, the inherent plasticity of membrane proteins may interfere with the structure determination using high-resolution techniques[15]; lastly, the diverse structures and functions of membrane proteins require a range of proper experimental conditions to maintain their stability and activity during isolation and analysis[2,16-18]. To date, there is still no universal protocol for extracting all types of membrane proteins from the native membrane whilst preserving their structure and function in their lipid environment.

1.1.2 The key role of lipids in biological membranes

The other major components of membranes are lipids, occupying around 50% of the bio-mass of the system[2]. In cells, membrane lipids are categorised into eight main classes: fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterols, prenols, saccharolipids, and polyketides[16]; of these phospholipids, sphingolipids and sterols are the predominant types[1,3]. With the variation in head groups and acyl chains (Figure 1-2), more than 30,000 different lipid structures derived from these basic families above have been reported[19], resulting in the complexity and heterogeneity of lipid membranes.

Due to this lipid diversity, membranes from different cells, organelles and even from two leaflets of the same bilayer possess a varied lipid composition and show distinct physical properties[2]. For instance, phosphatidylcholine (PC) constitutes approximately 40% of phospholipids in mammal membranes and 10% in yeast[20], but does not exist in the majority of bacteria, including *E. coli*[21]. The difference in lipid

composition and distribution in membranes results in variations in membrane permeability, fluidity, curvature, lateral pressure and transition temperature[22].

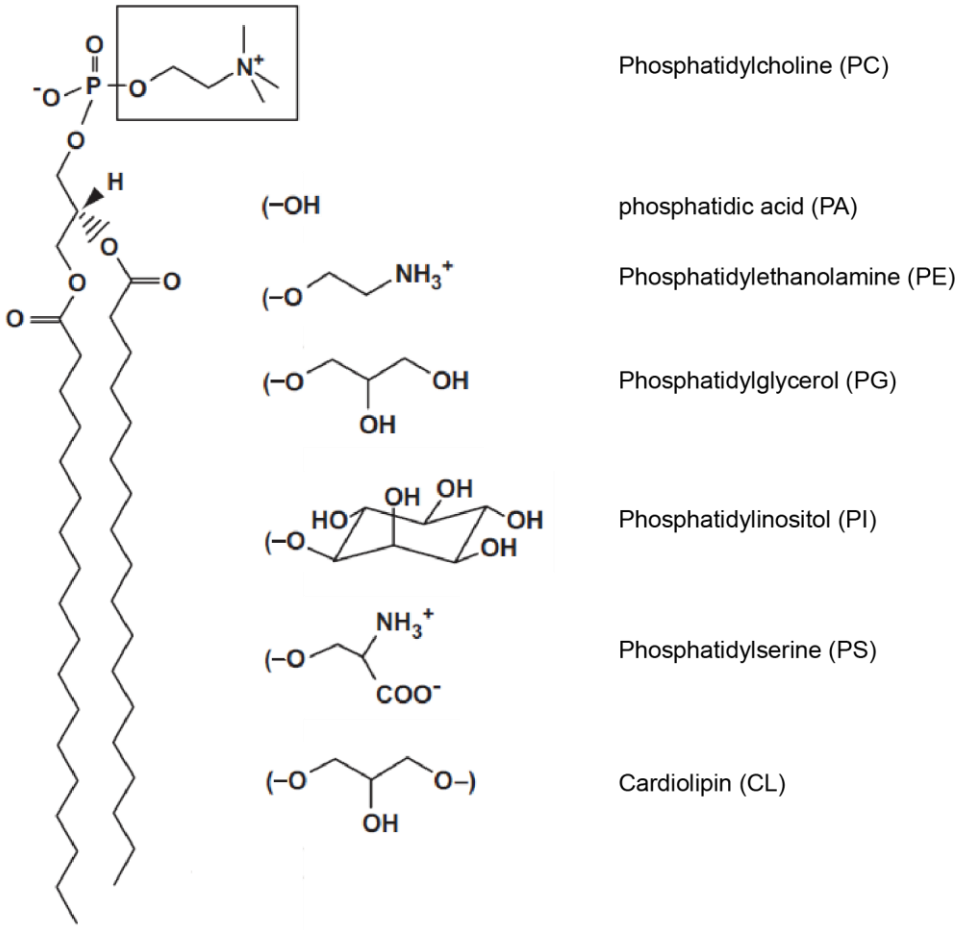


Figure 1-2 Chemical structures of glycerophosphate-based lipids. It was observed that the glycerophospholipids have the greatest diversity of lipid structures[16]. The structure of PC (18:0) above is shown as a representative phospholipid. On the right it shows the substitutions of the head group choline in box, which forms the other subclasses of phospholipids. The lipid acyl chains in membranes commonly have 12-24 carbons and 0-2 double bonds[2]. For illustration, a simple lipid model with two saturated 18-carbon tails is shown here. Figure is modified from Dowhan *et al.*[7].

Located in the biological bilayer formed by amphiphilic lipids, membrane proteins affect the lipid distribution in the neighbouring membrane area, and are also mediated by their lipid environment (Figure 1-1 A)[3,23]. Researchers have reported that interactions between membrane proteins and surrounding lipids may affect

membrane proteins in a series of ways: structural stability, conformational dynamics, oligomerisation, ligand/lipid affinity, functional activity, etc.[18,24-26]. These findings demonstrate that the lipid environment is a critical factor in modulating the structure and function of membrane proteins, as important as their amino acid sequence and post-translational modification (PTM)[2,18,27,28].

Depending on their position with membrane proteins, membrane lipids can be classified into three categories (Figure 1-3): bulk lipids, annular lipids and non-annular lipids, and each type affects proteins differently. Bulk lipids, the major component in membranes, have a low residence time at the protein-lipid interface through random collisions, and diffuse rapidly in the lipid bilayer plane[29]. These lipids determine the general characteristics of the membrane[23,30]. In direct contact with membrane proteins, the first shell of surrounding lipids are referred as annular lipids or boundary lipids[30-32]. Compared to bulk lipids, the diffusion and exchange rates of annular lipids are significantly lower. The lipid composition of the annulus is determined by the hydrophobic architecture of the attached proteins, which results in binding preference towards selected lipid classes[29]. Besides, these lipids are reported to affect the structure and function of the proteins they encapsulate[3,24,33]. Non-annular lipids, also known as lipid cofactors, interact with membrane proteins at specific sites with high affinity[29]. These lipids often reside within membrane proteins and are directly involved in the configuration and function of proteins, for example, gating control in ion channels and stimulation of ATPase activity[34,35].

Interactions between membrane proteins and lipids, therefore, are critical for the functioning of an organism. However, due to their diversity and complexity, many questions in this area remain unanswered. A better understanding of protein-lipid interactions will not only provide further valuable insights into this complicated area but also has the potential to aid in the development of therapeutics.

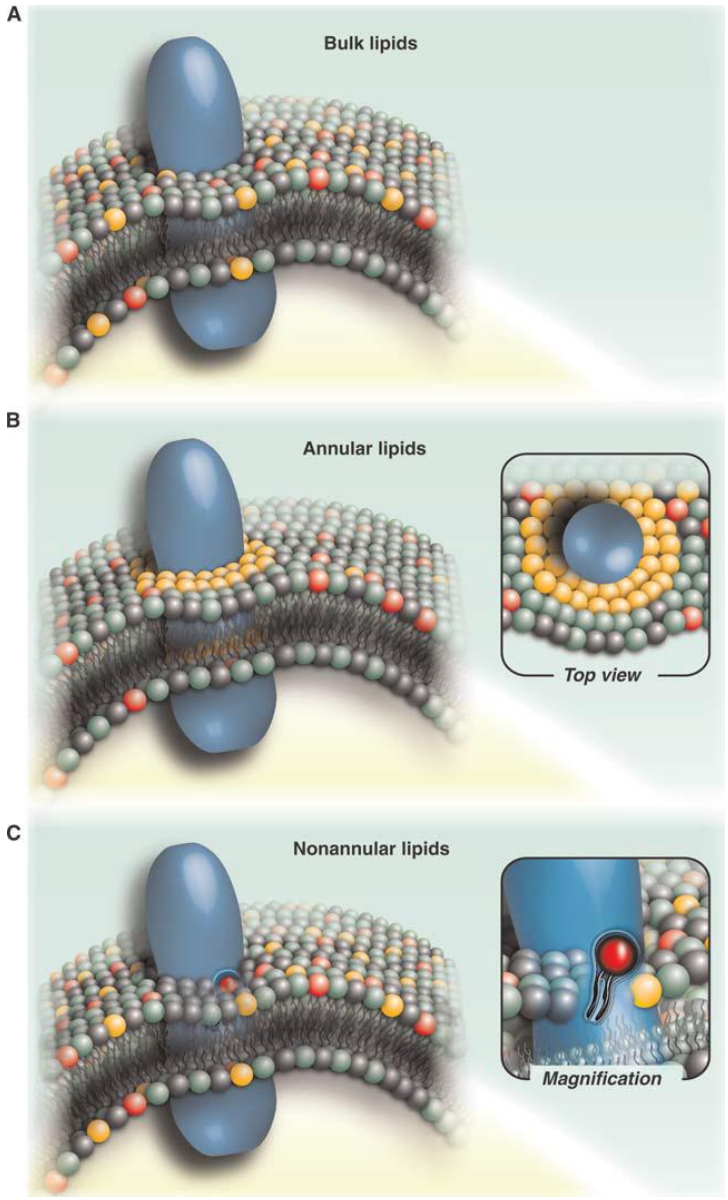


Figure 1-3 Membrane protein–lipid interactions within membranes. Transmembrane protein interacts with bulk lipids (**A**), annular lipids (**B**) and non-annular lipids (**C**). Figure is from Contreras *et al.*[29].

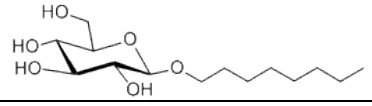
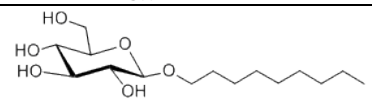
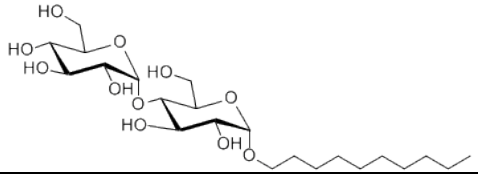
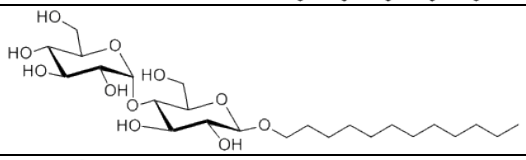
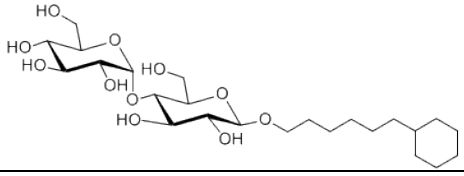
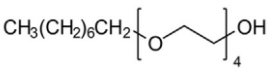
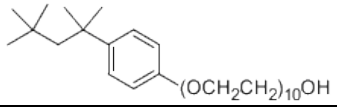
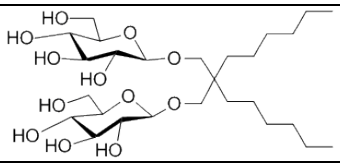
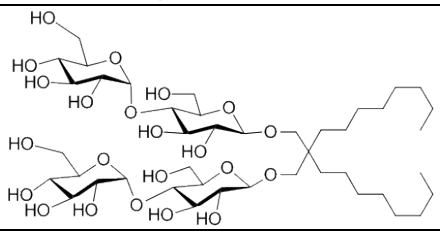
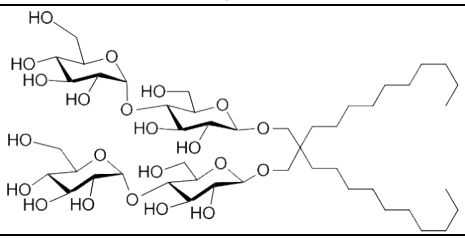
1.1.3 Methods for membrane proteins solubilisation

One of the greatest challenges in studying membrane proteins is the need to isolate them from other proteins in the same membrane whilst preserving their structure and function[36,37]. A number of approaches have been developed to stabilise membrane proteins when extracting them from their native lipid bilayer: the use of detergents, amphipols, bicelles, for example, allow interactions between lipids and membrane proteins to be investigated.

Detergents are the most widely used strategy for membrane protein extraction[12,38]. The amphiphilic detergent molecule contains a hydrophilic head group and hydrophobic tail moiety. The head groups tend to occupy more molecular space than the linear alkyl chains, forming spherical micelles with an apolar interior and polar exterior[39]. At concentrations over critical micelle concentration (CMC), detergents disrupt the membrane and envelop membrane proteins and remaining lipids into water soluble micelles[40,41].

Depending on the properties of hydrophilic head group, detergents can be categorised as non-ionic, zwitterionic or ionic, from mild to harsh, respectively. To preserve the inter- and intra-molecular interactions of membrane proteins, nonionic and zwitterionic detergents are usually used in the purification and structural determination of membrane proteins[42,43]. As no single detergent has been proven to be suitable for all membrane proteins, to achieve successful extraction of membrane proteins with different properties, a wide range of detergents needs to be considered[36,44]. Detergents utilised in the studies reported in this thesis are listed in

Table 1-1. We tested these detergents and their combinations, at different concentrations for the solubilisation of distinct target proteins (Section 2.7).

Detergent		**CMC/%**
Nonionic		
OG		0.53
NG		0.2
DM		0.08
DDM		0.0087
CYMAL-6		0.028
C_8E_4		0.25
Triton X-100		0.01-0.016
OGNG		0.058
DMNG		0.0034
LMNG		0.001

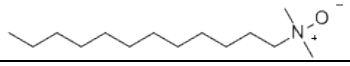
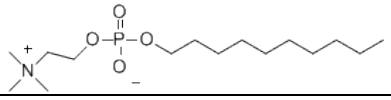
Zwitterionic		
LDAO		0.023
Fos-Choline-10		0.35

Table 1-1 Typical detergents for solubilisation of membrane proteins. The CMC values of different detergents in aqueous solution are from the reagent supplier Anatrace[19].

In contrast to the detergent method, using bicelles, membrane proteins are reconstituted into synthetic lipid bilayer discs formed by a combination of long-chain and short-chain phospholipids[45], providing an environment in which the structure and function of membrane proteins may be kept as *in vivo*[27]. For instance, research has shown that the effect of lipids on the biophysical properties of membrane proteins can be preserved in bicelles[46]. However, the specific lipid species suitable for bicelle formation limits its applications[47].

Nanodiscs are a relatively new approach for stabilising membrane proteins. Proteins are incorporated into a small section of the lipid bilayer and enclosed by membrane scaffold proteins (MSPs)[48]. As opposed to bicelles, various lipid species are amenable to the formation of nanodiscs[49], allowing the study of membrane proteins in a native-like lipid environment. Furthermore, it was reported that the presence of MSPs improves the stability of the embedded protein[50]. MSP nanodiscs are therefore a promising method for membrane protein research.

Although membrane proteins are incorporated in lipid bilayers in the two systems described above, both methods require detergents to extract membrane proteins

from their native lipid environment prior to particle formation. a new type of nanodiscs, saposin-lipoprotein particles (Salipro), was reported recently[51]. Under acidic condition, the open state saposin A (SapA) proteins bind to lipids with low specificity and form disc-shaped complexes for membrane protein incorporation[51-53]. Different with the common scaffold proteins, salipro provide a tunable range of size by manipulating the lipid/SapA ratio, rather than a rigid scaffold system with defined diametre[52,54]. More importantly, this protein can function on crude membrane to extract membrane proteins into native Salipro particles directly[55]. However, in this method, detergents are normally added to improve the accessibility towards SapA for better solubilisation[55].

To mitigate the effect of detergents, a novel detergent-free approach has been developed in recent years[56]. Lipodisqs, also known as styrene-maleic acid lipid particles (SMALPs) and native nanodiscs[56-58], are small discs of lipid bilayer encapsulated with a styrene-maleic acid (SMA) co-polymer. In contrast to the methods mentioned previously, lipodisqs are formed by the direct insertion of SMA co-polymer into the biological membrane of proteins[59]. A wide range of membrane proteins such as G protein-coupled receptors (GPCRs), transporters and ion channels have been successfully studied in lipodisqs[57,60,61] utilising common biophysical techniques: for example, circular dichroism spectroscopy[56], cryo-electron microscopy (cryo-EM)[62], dynamic light scattering[63], solid and solution state nuclear magnetic resonance (NMR) spectroscopy [64,65] and X-ray crystallography[66].

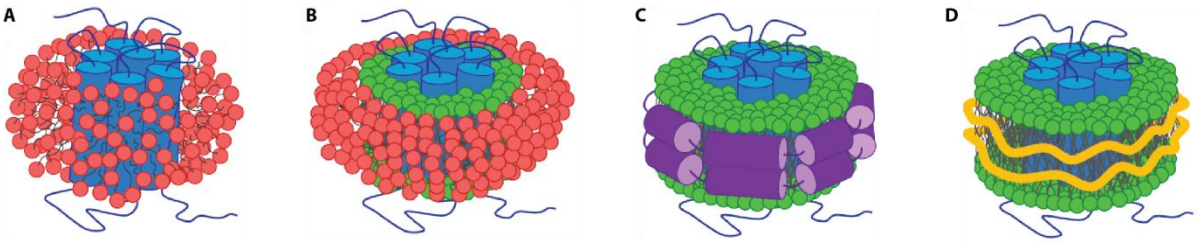


Figure 1-4 Membrane-mimetic systems for membrane protein solubilisation and stabilisation. The membrane protein (blue cylinders) is isolated in the detergent (red) micelle **(A)**, lipid (green) bicelle **(B)**, nanodisc stabilised by MSP (purple) **(C)**, and lipodisq stabilised by SMA co-polymer (yellow) **(D)**. Figure is modified from Dorr *et al.*[37]

1.2 Mass spectrometry: the tool for the study of protein-lipid interactions

As stated above, there are a number of analytical methods currently used to probe protein-lipid interactions. These include cryo-EM, differential scanning calorimetry, electron spin resonance (also known as electron paramagnetic resonance), Fourier transform-infrared spectroscopy, solid state NMR and X-ray crystallography[29]. In addition, chemical biology approaches, based on radioactive labelling of lipids and/or proteins, have also been developed to assess the interactions of the intramembranes[29,67].

Cryo-EM, solid state NMR and X-ray crystallography are the most common techniques used in structural biology, providing high-resolution structural information for membrane proteins[16,29]. However, these methods not only require large quantities of sample, but more importantly, their application is restricted by the complex nature of biological membrane systems, for example factors like intrinsic disorder, low abundance, heterogeneity and transiency[67]. Furthermore, treatments involved with

these methods, such as freezing or crystallisation, often hinder the detection of ligand bound states and conformations of membrane proteins[67,68].

The research reported in this thesis employs mass spectrometry (MS) as the major approach to interrogate protein-lipid complexes. As the conventional technique for mass determination, MS has already provided much information in the field of proteomics, lipidomics and metabolomics[69]. In recent decades, with the development of instrumentation and methodology, MS is emerging as a method of choice in the analysis of intact protein complexes, including soluble and membrane proteins[70,71]. This technique is often known as 'native MS' or 'non-covalent MS', since intra- and intermolecular non-covalent interactions from complexes are maintained in the gas phase (Figure 1-5)[72]. This, in turn, allows composition and structural information from proteins to be obtained[73-75]. In contrast to other methods, native MS reveals subunit stoichiometry, ligand or lipid binding and thermodynamic information in a rapid analysis procedure, with limited requirement of protein size and extra preparation steps[67,76]. Moreover, the lack of ensemble-averaging enables the detection of co-populated ligand bound states and conformations[77], which may be important when studying the mechanism of proteins and relevant dynamics and kinetics with functional ligands[78,79]. A brief introduction about native MS approaches used in this study is provided in the following sections.

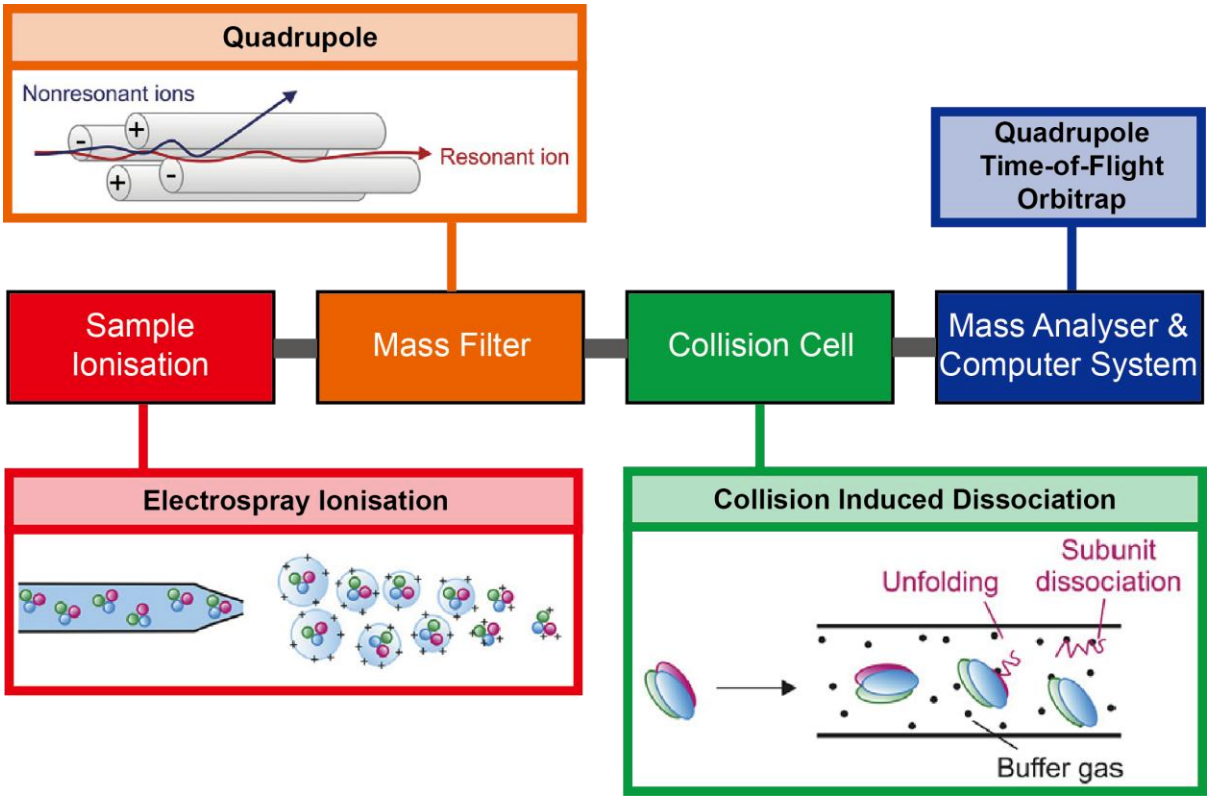


Figure 1-5 Components of a representative mass spectrometer for the analysis of non-covalent protein complexes. In our research, native MS analysis starts from sample ionisation by ESI. With the application of electrical potential on the tip of the capillary, sample ions are released from the charged droplets into the gas phase and then guided for mass filtration. Sample ions with specific *m/z* are selected through the quadrupole chamber and delivered into the collision cell for activation and dissociation. The final mass of ions is analysed and detected, and their signals are transformed into readable data via computational devices. Figure is modified from Calabrese *et al.*[67].

1.2.1 Electrospray ionisation

Proteins are introduced into the mass spectrometer by electrospray ionisation (ESI) at atmospheric pressure[73]. As the most widely-used soft ionisation method, samples are initiated with low residual energy to maintain the integrity of complexes[80,81]. This is in contrast to other common ionisation methods like matrix-assisted laser desorption ionisation (MALDI)[73].

In native MS, a miniaturised ESI, a nano-ESI (nESI) is utilised for protein ionisation. Compared with conventional ESI, nESI requires lower sample flow rates (nL/min) so less sample is consumed. Moreover, nESI is more tolerant to salt contamination in buffer solution, which is beneficial for prevention of non-specific oligomers[73,82]. Additionally, due to the high desolvation efficiency of the small size droplets generated by nESI, a gentle condition such as low source temperature and collision energy could be applied, enabling the labile interactions in protein complexes to be retained in the gas phase[83]. These advantages make nESI a more appropriate method than ESI to interrogate membrane proteins in native MS.

Protein sample in solution is ionised by the application of high electrical potential (0.5-1.8 kV) between the gold-coated capillary and the sample cone at the entrance of the instrument. Under the strong electrical field, positive charges are accelerated at the tip of the capillary, forming a 'Taylor cone'[84], and then emitted into nebulised droplets from the cone tip. As the droplets move forward by the electrical potential and pressure gradient, the solvent of droplets starts evaporate, leading to a reduction in size and coulombic repulsion between charges increases. The initial droplets fissure into smaller droplets at the point when the coulombic repulsion is strong enough to overcome the surface tension of the liquid. The newly formed droplets carry roughly 2% of the mass and 15% of the charge of their parent droplets[85,86]. With further solvent evaporation of offspring droplets, the sample ions are finally generated in the gas phase (Figure 1-5 red box)[73].

1.2.2 Collision-induced dissociation

In the presence of neutral gas molecules or atoms such as N_2 or Argon in the electrical field, the gas-phase ions made by nESI are focused into an accelerated beam via a radio frequency (RF) ion guide; they are then selected and transmitted through the mass filter (a quadrupole with low frequency) to the collision cell (Figure 1-5 orange box)[73]. The details of the quadrupole will be introduced as a mass analyser in the next section (Section 1.2.3).

At the stage of collision-induced dissociation (CID), the collision between the sample ions and inert buffer gas converts the kinetic energy of ions into their internal energy, resulting in the dissociation and fragmentation of samples if the energy is too high (Figure 1-5 green box). In native MS of non-covalent complexes, membrane proteins are encapsulated with the addition of a solubilisation reagent such as detergents and polymers (Section 1.1.3), which maintain the complex integrity during sample ionisation. With the activation from CID, these molecules are stripped from protein complexes carrying some internal energy to protect their structure from denaturing[87], and release the protein into the gas phase, producing well resolved mass spectra[88]. At the same time, the interactions between the binding lipids and membrane proteins are relatively strong, allowing them to remain intact during removal by detergents or polymers. By manipulating the accelerating potentials and gas pressures in the cell, the intensity of CID can be controlled[73]. Therefore, this technique also achieves the dissociation of lipid/ligand binding to proteins. This has already been applied to good effect in the interrogation of protein-lipid interactions[24,89].

1.2.3 Mass analysis

To determine ion mass, a wide range of mass analysers has been developed but few are able to be utilised for the study of non-covalent complexes.

In the early stage of protein MS, the quadrupole is used as the mass analyser, both because of the low cost and the tolerance of the ion beam from ESI[73]. A single quandrupole mass analyser comprises four parallel cylindrical electrodes equidistant to the central axis, with the same voltage but opposite polarity on the neighbouring rods. The potential on electrodes is composed of a direct current voltage superimposing an alternating current voltage with RF, generating a periodically changed electrical field to repel and attract injected ions onto the cross section of this vacuum chamber. In defined alternating current voltage and RF, only ions with the selected mass-to-charge ratio (*m/z*) range can pass through the quadrupole successfully in an oscillated trajectories, while the rest ions exit through the side of quadrupole or collide with the poles to be discharged (Figure 1-5 orange box). In the analysis of large protein complexes, quadrupole frequencies have to be reduced to allow larger ions through (Mathieu Equation), resulting in poor resolution[90]. Therefore, quadrupole analysers are commonly employed as mass filters and ion guides (Section 1.2.2).

More recently, the time-of-flight (ToF) analyser has been the preferential tool in the study of non-covalent complexes[91]. The ToF contains an accelerator, a field-free drift region, a reflectron and a detector (normally a microchannel plate) in the flight tube with a high vacuum (Figure 1-6 A). In contrast to the quadrupole, the ToF is a pulsed

technique, in which a constant electrical field between electrodes pushes ions from the previous chamber into the drift tube, converting electrical potential energy into kinetic energy of ions. Samples fly through a fixed distance to reach the ion detector in different times. As the time of flight is proportional to the square root of *m/z*, the *m/z* values of detected ions can be measured, so theoretically there is no limit in the mass range for detection[73]. To eliminate the effect of the initial distribution of kinetic energy on ions, the drift tube is positioned orthogonally to the incoming direction of samples. Additionally, a reflectron method is designed to synchronise the ions with the same *m/z*, but different kinetic energies arriving at the detector[75], improving the resolution of the instrument.

In the projects reported in this thesis, a hybrid Quadrupole Time-of-Flight (Q-ToF) instrument was used. In Q-ToF, a quadrupole, a collision cell and a ToF are arranged in series. Target ions are filtered in the quadrupole based on their *m/z* values, then activated in the collision cell and subsequently their masses are analysed in the ToF. This hybrid system enables the selective dissociation and mass measurement of sample ions in the gas phase with high sensitivity and resolution[92], as well as high accuracy quantification analysis of detected species[93]. Furthermore, Q-ToF MS provides the full spectrum of ion transmission in a single run of fast detection, aiding the identification of unknown binding molecules in the investigation of protein-lipid interactions; such molecules may be essential for the structure and function of the membrane protein.

Another analyser applied in our research is the Orbitrap analyser. Orbitraps are ion-trapping devices which utilise a static electrical field and act as both mass analyser and detector[94]. The Orbitrap consists of a spindle-shaped central electrode and two halves of outer electrodes separated by an insulating ring encapsulating the central electrode (Figure 1-6 B). In the Orbitrap, ions are injected into the cavity between the two electrodes, orthogonally to the *z*-axis of the Orbitrap. The electrostatic field ensures ions are trapped in the chamber and circulated around the central electrode in a spiral trajectory, describing an orbit motion and an harmonic oscillated motion along the z-axis, which is related to the *m/z* of individual ions only[94]. According to the image current detected, the oscillation frequencies can be calculated via Fourier transform and the mass of trapped ions can be measured. Compared to the orthogonal reflectron ToF analyser, the Orbitrap possesses a higher accuracy, sensitivity and resolution because of the signal superposition of the trapped ions[95]. Therefore, Orbitraps have a widespread application in MS methodology including tandem MS (MS/MS) fragmentation, bottom-up and top-down proteomics[96], contributing to the identification, quantification and characterisation of components in biological systems[94].

The high *m/z* values of non-covalent ions require condition modifications to the mass spectrometer in order to analyse the mass of membrane proteins. High pressure in the initial chamber and the ion guide, low-frequency in the quadrupoles and low sampling frequencies in the ToF analysers can contribute to the selection, transmission and detection of high *m/z* ions[73,90]. With the development of native MS, further methodology are being optimised for the investigation of protein complexes.

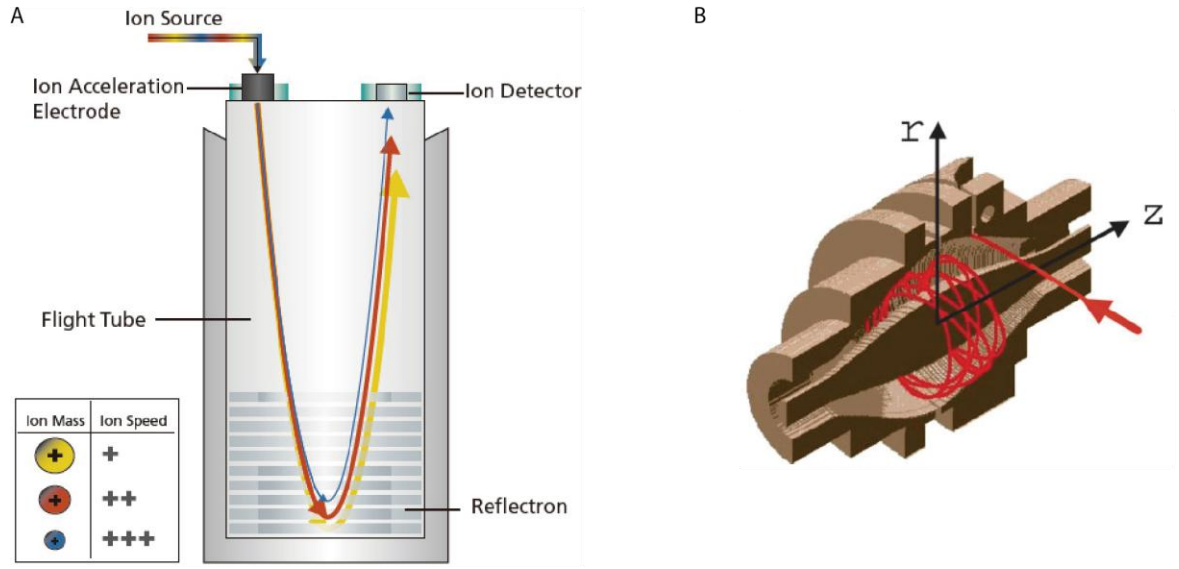


Figure 1-6 Schematic of the typical mass analysers of the mass spectrometer. A, a cross-sectional view of the ToF mass analyser. Ions are accelerated and deflected into the ToF chamber orthogonally to the original direction, and reach the detector after a period of flight. The arrival time of ions is used for their *m/z* calculation. Reflectron is involved for the compensation of the initial kinetic energy difference between ions. Figure is from Shimadzu Corporation publication[93]. **B,** a cross-sectional view of the Orbitrap mass analyser. Ions are injected and orbit around the central electrode in *z*-direction, generating image current. The *m/z* of trapped ions can be determined from their oscillation frequencies by Fourier transform. Figure is from Hu *et al.*[94]

Due to its unique capability, native MS is now employed in many studies of protein-lipid interactions in biophysical complexes, including transporters, ion channels, receptors and respiratory supercomplexes[89,97-99]. Using native MS, the selectivity and specificity of various lipid species towards membrane proteins can be revealed[24,26,100]. Equally, using native MS it is also possible to evaluate the effects of lipid binding on the structure and function of membrane proteins[26,98,101], as well as interactions with other lipids on binding sites[102]. Besides, native MS can also be used in combination with technologies such as X-ray crystallography and cryo-EM to provide additional information and to assist in the assignment of ambiguous conformations[99,103], furthering our understanding of protein-lipid interactions.

1.2.4 Lipidomic analysis

As well as protein-lipid interactions, MS is also used for the identification and quantification of the phospholipids co-purified with the membrane proteins. Various mass spectrometric techniques have been used for the characterisation of phospholipids, such as fast-atom bombardment and MALDI in combination with MS[104,105].

In our study, liquid chromatography/electrospray ionisation with tandem mass spectrometry using reversed-phase liquid chromatography (RP-LC/ESI-MS/MS) was utilised for the analysis of phospholipids extracted from membranes. The extracted phospholipid mixtures were separated through a C18 reversed-phase column in the LC system, and their species and individual acyl chain substituents were identified by the fragmentation patterns provided by MS/MS[106]. By means of an internal standard, the relative quantification of phospholipid species present could be determined.

1.3 Objectives

Of the many factors governing the stability and function of biological membranes, the non-covalent interaction between membrane proteins and surrounding lipids is emerging as one of the most important. The brief overview above shows that MS is a powerful technique to investigate intact protein assemblies and co-purified lipid molecules. With recent developments in instrumentation and methodology, native MS is capable to preserve the non-covalent interactions in the gas phase, enabling the direct interrogation of membrane protein-lipid interactions.

In this thesis, the use of MS is explored in the study of protein-lipid complexes. Starting from the traditional detergent purification method, we established a protocol to control the degree of delipidation of membrane proteins, thus capturing the membrane proteins in various lipid bound states by native MS. This enabled the investigation of effects from different membrane lipids on bound proteins in the native environment. As this study required significant time and samples, we introduced, a novel detergent-free protein extraction method in the following study. As an alternative membrane mimic, the polymer-based membrane protein complexes were evaluated in native MS, and compared with the detergent approach, particularly the preservation of protein-lipid interactions. The third project is a collaborative work. MS was applied in combination with cryo-EM and other experimental assays, supporting a comprehensive understanding of the mechanism of a respiratory supercomplex in mycobacteria.

Finally, we hope this research will aid the elucidation of new protein-lipid interactions and their modulation on organisms, as well as expanding the use of MS in the study of biological molecules.

References

(1) Campbell, N. A.; Reece, J. B.: *Campbell biology*; Pearson: Boston, Mass, 2014.

(2) Lodish, H. F., A. Berk, P. Matsudaira, C. Kaiser, M. Krieger, M. Scott, L. Zipursky, and J. E. Darnell: *Molecular Cell Biology*; Scientific American Press, 2004.

(3) Luckey, M.: *Membrane Structural Biology: With Biochemical and Biophysical Foundations*; Cambridge University Press, 2008.

(4) Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* **2017**, *16*, 19-34.

(5) Filmore, D. It's a GPCR world. *Modern Drug Discovery* **2004**, *7*, 24-28.

(6) Uhlen, M.; Fagerberg, L.; Hallstrom, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, A.; Kampf, C.; Sjostedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P. H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.; von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Ponten, F. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419.

(7) Dowhan, W.; Bogdanov, M.; Mileykovskaya, E.: Functional Roles of Lipids in Membranes. In *Biochemistry of Lipids, Lipoproteins and Membranes*, 2016; pp 1-40.

(8) Wallin, E.; von Heijne, G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein science : a publication of the Protein Society* **1998**, *7*, 1029-1038.

(9) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res* **2000**, *28*, 235-242.

(10) Okuno, Y.; Tamon, A.; Yabuuchi, H.; Niijima, S.; Minowa, Y.; Tonomura, K.; Kunimoto, R.; Feng, C. GLIDA: GPCR--ligand database for chemical genomics drug discovery--database and tools update. *Nucleic Acids Res* **2008**, *36*, D907-912.

(11) Almen, M. S.; Nordstrom, K. J.; Fredriksson, R.; Schioth, H. B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol* **2009**, *7*, 50.

(12) Hunte;, C.; Jagow;, G. v.; Schagger, H.: *Membrane Protein Purification and Crystallization: A Practical Guide*; 2nd ed.; Academic Press, 2003. pp. 316.

(13) Lacapere, J. J.; Pebay-Peyroula, E.; Neumann, J. M.; Etchebest, C. Determining membrane protein structures: still a challenge! *Trends in biochemical sciences* **2007**, *32*, 259-270.

(14) Zorman, S.; Botte, M.; Jiang, Q.; Collinson, I.; Schaffitzel, C. Advances and challenges of membrane-protein complex production. *Curr Opin Struct Biol* **2015**, *32*, 123-130.

(15) Rawson, S.; Davies, S.; Lippiat, J. D.; Muench, S. P. The changing landscape of membrane protein structural biology through developments in electron microscopy. *Mol Membr Biol* **2016**, *33*, 12-22.

(16) Barrera, N. P.; Zhou, M.; Robinson, C. V. The role of lipids in defining membrane protein interactions: insights from mass spectrometry. *Trends in cell biology* **2013**, *23*, 1-8.

(17) Marsh, D. Protein modulation of lipids, and vice-versa, in membranes. *Biochimica et biophysica acta* **2008**, *1778*, 1545-1575.

(18) Phillips, R.; Ursell, T.; Wiggins, P.; Sens, P. Emerging roles for lipids in shaping membrane-protein function. *Nature* **2009**, *459*, 379-385.

(19) https://www.anatrace.com/Products/Detergents.

(20) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* **2008**, *9*, 112-124.

(21) Chen, F.; Zhao, Q.; Cai, X.; Lv, L.; Lin, W.; Yu, X.; Li, C.; Li, Y.; Xiong, M.; Wang, X. G. Phosphatidylcholine in membrane of Escherichia coli changes bacterial antigenicity. *Can J Microbiol* **2009**, *55*, 1328-1334.

(22) Pomorski, T. G.; Nylander, T.; Cardenas, M. Model cell membranes: discerning lipid and protein contributions in shaping the cell. *Adv Colloid Interface Sci* **2014**, *205*, 207-220.

(23) Dowhan, W. Molecular basis for membrane phospholipid diversity: Why are there so many lipids? *Annual review of biochemistry* **1997**, *66*, 199-232.

(24) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **2014**, *510*, 172-175.

(25) Gupta, K.; Donlan, J. A.; Hopper, J. T.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **2017**, *541*, 421-424.

(26) Bechara, C.; Noll, A.; Morgner, N.; Degiacomi, M. T.; Tampe, R.; Robinson, C. V. A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nature chemistry* **2015**, *7*, 255-262.

(27) Zhou, H. X.; Cross, T. A. Influences of membrane mimetic environments on membrane protein structures. *Annu Rev Biophys* **2013**, *42*, 361-392.

(28) Palsdottir, H.; Hunte, C. Lipids in membrane protein structures. *Biochimica et biophysica acta* **2004**, *1666*, 2-18.

(29) Contreras, F. X.; Ernst, A. M.; Wieland, F.; Brugger, B. Specificity of intramembrane protein-lipid interactions. *Cold Spring Harbor perspectives in biology* **2011**, *3*.

(30) Lee, A. G. Lipid–protein interactions in biological membranes: a structural perspective. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2003**, *1612*, 1-40.

(31) Lee, A. G. Biological membranes: the importance of molecular detail. *Trends in biochemical sciences* **2011**, *36*, 493-500.

(32) Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochimica et biophysica acta* **2014**, *1838*, 1548-1559.

(33) Dowhan, W.; Bogdanov, M. Lipid-protein interactions as determinants of membrane protein structure and function. *Biochemical Society transactions* **2011**, *39*, 767-774.

(34) Hansen, S. B.; Tao, X.; MacKinnon, R. Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2. *Nature* **2011**, *477*, 495-498.

(35) Starling, A. P.; East, J. M.; Lee, A. G. Phosphatidylinositol 4-Phosphate Increases the Rate of Dephosphorylation of the Phosphorylated Ca2+-Atpase. *Journal of Biological Chemistry* **1995**, *270*, 14467-14470.

(36) Seddon, A. M.; Curnow, P.; Booth, P. J. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et biophysica acta* **2004**, *1666*, 105-117.

(37) Dorr, J. M.; Scheidelaar, S.; Koorengevel, M. C.; Dominguez, J. J.; Schafer, M.; van Walree, C. A.; Killian, J. A. The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J* **2016**, *45*, 3-21.

(38) Garavito, R. M.; Ferguson-Miller, S. Detergents as tools in membrane biochemistry. *The Journal of biological chemistry* **2001**, *276*, 32403-32406.

(39) detergents and their uses in membrane protein science. https://www.anatrace.com/Technical-Documentation/Catalogs/Anatrace-Detergents-Booklet-FINAL.

(40) le Maire, M.; Champeil, P.; Moller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et biophysica acta* **2000**, *1508*, 86-111.

(41) Lichtenberg, D.; Ahyayauch, H.; Goni, F. M. The mechanism of detergent solubilization of lipid bilayers. *Biophys J* **2013**, *105*, 289-299.

(42) VanAken, T.; Foxall-VanAken, S.; Castleman, S.; Ferguson-Miller, S. Alkyl glycoside detergents: synthesis and applications to the study of membrane proteins. *Methods Enzymol* **1986**, *125*, 27-35.

(43) Oxenoid, K.; Chou, J. J. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 10870-10875.

(44) Prive, G. G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **2007**, *41*, 388-397.

(45) Kiessling, V.; Domanska, M. K.; Murray, D.; Wan, C.; Tamm, L.: *Supported Lipid Bilayers: Development and Applications in Chemical Biology*, 2008; Vol. 4.

(46) Mohanty, P.; Lee, J.; Glover, K. J.; Landskron, K. Discoid Bicelles as Efficient Templates for Pillared Lamellar Periodic Mesoporous Silicas at pH 7 and Ultrafast Reaction Times. *Nanoscale Res Lett* **2011**, *6*.

(47) Durr, U. H.; Gildenberg, M.; Ramamoorthy, A. The magic of bicelles lights up membrane protein structure. *Chemical reviews* **2012**, *112*, 6054-6074.

(48) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett* **2002**, *2*, 853-856.

(49) Hagn, F.; Etzkorn, M.; Raschle, T.; Wagner, G. Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *Journal of the American Chemical Society* **2013**, *135*, 1919-1925.

(50) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *Journal of the American Chemical Society* **2004**, *126*, 3477-3487.

(51) Frauenfeld, J.; Loving, R.; Armache, J. P.; Sonnen, A. F.; Guettou, F.; Moberg, P.; Zhu, L.; Jegerschold, C.; Flayhan, A.; Briggs, J. A.; Garoff, H.; Low, C.; Cheng, Y.; Nordlund, P. A saposin-lipoprotein nanoparticle system for membrane proteins. *Nature methods* **2016**, *13*, 345-351.

(52) Flayhan, A.; Mertens, H. D. T.; Ural-Blimke, Y.; Martinez Molledo, M.; Svergun, D. I.; Low, C. Saposin Lipid Nanoparticles: A Highly Versatile and Modular Tool for Membrane Protein Research. *Structure* **2018**, *26*, 345-355 e345.

(53) Popovic, K.; Holyoake, J.; Pomes, R.; Prive, G. G. Structure of saposin A lipoprotein discs. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 2908-2912.

(54) Kehlenbeck, D. M.; Josts, I.; Nitsche, J.; Busch, S.; Forsyth, V. T.; Tidow, H. Comparison of lipidic carrier systems for integral membrane proteins - MsbA as case study. *Biol Chem* **2019**, *400*, 1509-1518.

(55) Salipro® goes DirectMX™: One-step reconstitution of membrane proteins from crude membranes. https://www.salipro.com/application-note-directmx.html.

(56) Knowles, T. J.; Finka, R.; Smith, C.; Lin, Y. P.; Dafforn, T.; Overduin, M. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *Journal of the American Chemical Society* **2009**, *131*, 7484-7485.

(57) Dorr, J. M.; Koorengevel, M. C.; Schafer, M.; Prokofyev, A. V.; Scheidelaar, S.; van der Cruijsen, E. A.; Dafforn, T. R.; Baldus, M.; Killian, J. A. Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K+ channel: the power of native nanodiscs. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 18607-18612.

(58) Orwick, M. C.; Judge, P. J.; Procek, J.; Lindholm, L.; Graziadei, A.; Engel, A.; Grobner, G.; Watts, A. Detergent-free formation and physicochemical characterization of nanosized lipid-polymer complexes: Lipodisq. *Angewandte Chemie* **2012**, *51*, 4653-4657.

(59) Jamshad, M.; Grimard, V.; Idini, I.; Knowles, T. J.; Dowle, M. R.; Schofield, N.; Sridhar, P.; Lin, Y.; Finka, R.; Wheatley, M.; Thomas, O. R. T.; Palmer, R. E.; Overduin, M.; Govaerts, C.; Ruysschaert, J.-M.; Edler, K. J.; Dafforn, T. R. Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins. *Nano Research* **2014**, *8*, 774-789.

(60) Wheatley, M.; Charlton, J.; Jamshad, M.; Routledge, S. J.; Bailey, S.; La-Borde, P. J.; Azam, M. T.; Logan, R. T.; Bill, R. M.; Dafforn, T. R.; Poyner, D. R. GPCR-styrene maleic acid lipid particles (GPCR-SMALPs): their nature and potential. *Biochemical Society transactions* **2016**, *44*, 619-623.

(61) Postis, V.; Rawson, S.; Mitchell, J. K.; Lee, S. C.; Parslow, R. A.; Dafforn, T. R.; Baldwin, S. A.; Muench, S. P. The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochimica et biophysica acta* **2015**, *1848*, 496-501.

(62) Parmar, M.; Rawson, S.; Scarff, C. A.; Goldman, A.; Dafforn, T. R.; Muench, S. P.; Postis, V. L. G. Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure. *Biochim Biophys Acta Biomembr* **2018**, *1860*, 378-383.

(63) Zhang, R.; Sahu, I. D.; Liu, L.; Osatuke, A.; Comer, R. G.; Dabney-Smith, C.; Lorigan, G. A. Characterizing the structure of lipodisq nanoparticles for membrane protein spectroscopic studies. *Biochimica et biophysica acta* **2015**, *1848*, 329-333.

(64) Bersch, B.; Dorr, J. M.; Hessel, A.; Killian, J. A.; Schanda, P. Proton-Detected Solid-State NMR Spectroscopy of a Zinc Diffusion Facilitator Protein in Native Nanodiscs. *Angewandte Chemie* **2017**, *56*, 2508-2512.

(65) Puthenveetil, R.; Nguyen, K.; Vinogradova, O. Nanodiscs and Solution NMR: preparation, application and challenges. *Nanotechnol Rev* **2017**, *6*, 111-126.

(66) Broecker, J.; Eger, B. T.; Ernst, O. P. Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs. *Structure* **2017**, *25*, 384-392.

(67) Calabrese, A. N.; Radford, S. E. Mass spectrometry-enabled structural biology of membrane proteins. *Methods* **2018**, *147*, 187-205.

(68) Laganowsky, A.; Reading, E.; Hopper, J. T.; Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nature protocols* **2013**, *8*, 639-651.

(69) Smith, R.; Mathis, A. D.; Ventura, D.; Prince, J. T. Proteomics, lipidomics, metabolomics: a mass spectrometry tutorial from a computer scientist's point of view. *BMC Bioinformatics* **2014**, *15 Suppl 7*, S9.

(70) Katta, V.; Chait, B. T. Observation of the Heme Globin Complex in Native Myoglobin by Electrospray-Ionization Mass-Spectrometry. *Journal of the American Chemical Society* **1991**, *113*, 8534-8535.

(71) Ganem, B.; Li, Y. T.; Henion, J. D. Detection of Noncovalent Receptor Ligand Complexes by Mass-Spectrometry. *Journal of the American Chemical Society* **1991**, *113*, 6294-6296.

(72) Leney, A. C.; Heck, A. J. Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom* **2017**, *28*, 5-13.

(73) Benesch, J. L.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chemical reviews* **2007**, *107*, 3544-3567.

(74) Sharon, M. How far can we go with structural mass spectrometry of protein complexes? *J Am Soc Mass Spectrom* **2010**, *21*, 487-500.

(75) Sharon, M.; Robinson, C. V. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annual review of biochemistry* **2007**, *76*, 167-193.

(76) Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. Studying 18 MDa virus assemblies with native mass spectrometry. *Angewandte Chemie* **2013**, *52*, 4020-4023.

(77) Smith, D. P.; Giles, K.; Bateman, R. H.; Radford, S. E.; Ashcroft, A. E. Monitoring copopulated conformational states during protein folding events using

electrospray ionization-ion mobility spectrometry-mass spectrometry. *J Am Soc Mass Spectrom* **2007**, *18*, 2180-2190.

(78) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature chemistry* **2014**, *6*, 281-294.

(79) Lossl, P.; van de Waterbeemd, M.; Heck, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO J* **2016**, *35*, 2634-2657.

(80) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass-Spectrometry of Large Biomolecules. *Science* **1989**, *246*, 64-71.

(81) Hilton, G. R.; Benesch, J. L. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *J R Soc Interface* **2012**, *9*, 801-816.

(82) Wilm, M.; Mann, M. Analytical properties of the nanoelectrospray ion source. *Analytical chemistry* **1996**, *68*, 1-8.

(83) Hernandez, H.; Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols* **2007**, *2*, 715-726.

(84) Fernández de la Mora, J. The Fluid Dynamics of Taylor Cones. *Annual Review of Fluid Mechanics* **2007**, *39*, 217-243.

(85) Gomez, A.; Tang, K. Charge and fission of droplets in electrostatic sprays. *Physics of Fluids* **1994**, *6*, 404-414.

(86) Kebarle, P.; Peschke, M. On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions. *Anal Chim Acta* **2000**, *406*, 11-35.

(87) Borysik, A. J.; Hewitt, D. J.; Robinson, C. V. Detergent release prolongs the lifetime of native-like membrane protein conformations in the gas-phase. *Journal of the American Chemical Society* **2013**, *135*, 6078-6083.

(88) Barrera, N. P.; Di Bartolo, N.; Booth, P. J.; Robinson, C. V. Micelles protect membrane complexes from solution to vacuum. *Science* **2008**, *321*, 243-246.

(89) Yen, H. Y.; Hoi, K. K.; Liko, I.; Hedger, G.; Horrell, M. R.; Song, W.; Wu, D.; Heine, P.; Warne, T.; Lee, Y.; Carpenter, B.; Pluckthun, A.; Tate, C. G.; Sansom, M. S. P.; Robinson, C. V. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **2018**, *559*, 423-427.

(90) 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. *Analytical chemistry* **2002**, *74*, 1402-1407.

(91) Cotter, R. J. Time-of-flight mass spectrometry for the structural analysis of biological molecules. *Analytical chemistry* **1992**, *64*, 1027A-1039A.

(92) Benesch, J. L.; Ruotolo, B. T.; Sobott, F.; Wildgoose, J.; Gilbert, A.; Bateman, R.; Robinson, C. V. Quadrupole-time-of-flight mass spectrometer modified for higher-energy dissociation reduces protein assemblies to peptide fragments. *Analytical chemistry* **2009**, *81*, 1270-1274.

(93) Fundamental Guide to Liquid Chromatography Mass Spectrometry (LCMS). https://www.shimadzu.com/an/lcms/support/fundamental/index.html.

(94) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. The Orbitrap: a new mass spectrometer. *Journal of mass spectrometry : JMS* **2005**, *40*, 430-443.

(95) Eliuk, S.; Makarov, A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif)* **2015**, *8*, 61-80.

(96) Scigelova, M.; Makarov, A. Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics* **2006**, *6 Suppl 2*, 16-21.

(97) Marcoux, J.; Wang, S. C.; Politis, A.; Reading, E.; Ma, J.; Biggin, P. C.; Zhou, M.; Tao, H.; Zhang, Q.; Chang, G.; Morgner, N.; Robinson, C. V. Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 9704-9709.

(98) Reading, E.; Walton, T. A.; Liko, I.; Marty, M. T.; Laganowsky, A.; Rees, D. C.; Robinson, C. V. The Effect of Detergent, Temperature, and Lipid on the Oligomeric State of MscL Constructs: Insights from Mass Spectrometry. *Chemistry & biology* **2015**, *22*, 593-603.

(99) Gong, H.; Li, J.; Xu, A.; Tang, Y.; Ji, W.; Gao, R.; Wang, S.; Yu, L.; Tian, C.; Li, J.; Yen, H. Y.; Man Lam, S.; Shui, G.; Yang, X.; Sun, Y.; Li, X.; Jia, M.; Yang, C.; Jiang, B.; Lou, Z.; Robinson, C. V.; Wong, L. L.; Guddat, L. W.; Sun, F.; Wang, Q.; Rao, Z. An electron transfer path connects subunits of a mycobacterial respiratory supercomplex. *Science* **2018**, *362*.

(100) Gault, J.; Donlan, J. A.; Liko, I.; Hopper, J. T.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.; Belov, M.; Damoc, E.; Makarov, A.; Robinson, C. V. High-resolution mass

spectrometry of small molecules bound to membrane proteins. *Nature methods* **2016**, *13*, 333-336.

(101) Mehmood, S.; Corradi, V.; Choudhury, H. G.; Hussain, R.; Becker, P.; Axford, D.; Zirah, S.; Rebuffat, S.; Tieleman, D. P.; Robinson, C. V.; Beis, K. Structural and Functional Basis for Lipid Synergy on the Activity of the Antibacterial Peptide ABC Transporter McjD. *The Journal of biological chemistry* **2016**, *291*, 21656-21668.

(102) Patrick, J. W.; Boone, C. D.; Liu, W.; Conover, G. M.; Liu, Y.; Cong, X.; Laganowsky, A. Allostery revealed within lipid binding events to membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115*, 2976-2981.

(103) Zhou, M.; Morgner, N.; Barrera, N. P.; Politis, A.; Isaacson, S. C.; Matak-Vinkovic, D.; Murata, T.; Bernal, R. A.; Stock, D.; Robinson, C. V. Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. *Science* **2011**, *334*, 380-385.

(104) Jensen, N. J.; Tomer, K. B.; Gross, M. L. FAB MS/MS for phosphatidylinostitol,-glycerol,-ethanolamine and other complex phospholipids. **1987**, *22*, 480-489.

(105) Ishida, Y.; Madonna, A. J.; Rees, J. C.; Meetani, M. A.; Voorhees, K. J. Rapid analysis of intact phospholipids from whole bacterial cells by matrix-assisted laser desorption/ionization mass spectrometry combined with on-probe sample pretreatment. *Rapid communications in mass spectrometry : RCM* **2002**, *16*, 1877-1882.

(106) Cabrera, G. M.; Murga, M. L.; de Valdez, G. F.; Seldes, A. M. Direct analysis by electrospray ionization tandem mass spectrometry of mixtures of

phosphatidyldiacylglycerols from Lactobacillus. *Journal of mass spectrometry : JMS* **2000**, *35*, 1452-1459.

2 Identifying key lipids in membrane proteins using progressive delipidation coupled with native mass spectrometry and quantitative lipidomics

2.1 Introduction

The local lipid environment is emerging as one of the most important factors modulating the stability and activity of membrane proteins[1,2]. Not only the strongly bound specific lipids, but also the weakly bound annular lipids could impact the character of proteins[3]. However, the complexity of the lipid environment makes it difficult to distinguish the different kinds of bound lipids and their corresponding functions[4]. Native MS has been successfully applied in a number of membrane protein studies in which key lipids and their role in protein structure and function have been identified[5-7]. In the majority of these studies the proteins were almost 'naked' following a thorough detergent delipidation process. As often only lipids with extremely high binding preference survived, a full understanding of the influence of surrounding lipids on the membrane protein was not possible in these cases.

Controlled delipidation of the adenosine triphosphate (ATP)-binding cassette (ABC) transporter, TmrAB, achieved successive lipid removal using native MS[8]. Based on this study, our aim was to develop a universal progressive delipidation protocol to control the extent of membrane protein delipidation, whilst preserving the important protein-lipid interactions in their native environment. This would allow further analysis of their role in protein structure and function. Native MS would then be employed to examine the results of the progressive delipidation, followed by quantitative lipidomics

to provide details of bound lipids at different lipidation stages. Whilst, ion mobility-mass spectrometry (IM-MS), would then be used to evaluate the stability of diverse lipid bound proteins.

In IM spectroscopy, samples are passed through a chamber filled with an inert neutral gas under the influence of an electric field and separated according to their mobility[9-11]. The probability of extended and compact ions colliding with the buffer gas varies due to the difference in the collisional cross section (CCS), resulting in different travel speed and then drift times, so that proteins can be separated in this cell[12]. In this project, IM-MS was introduced to measure the time taken for different lipid bound proteins to pass through the IM cell. With increasing activation energy, the CCS of proteins increased, so a longer drift time was observed for the particles to traverse the IM cell. A comparison of protein unfolding voltage revealed the stability difference between these protein-lipid complexes and demonstrated the effect of binding lipids. This method has been successfully employed previously to achieve measurements that cannot be accomplished in the solution phase[13].

In this project, membrane proteins from two different families were selected to examine the feasibility of the protocol. We first tested *E. coli* MsbA, an analogous ABC transporter of TmrAB used in the previous study[8]. MsbA is a homodimeric exporter located in the inner membrane where it flips lipid A from the cytoplasmic to the periplasmic leaflet, contributing in the outer membrane construction of bacteria (Figure 2-1 A)[14]. Another transporter, MdtK, from the multidrug and toxic compound extrusion (MATE) family was also used to expand the range of the progressive delipidation protocol. MdtK functions as multidrug efflux to export xenobiotics and

endogenous hazardous substances such as fluoroquinolones, trimethoquinolones and chloramphenicol[15], via an electro-chemical gradient of H^+/Na^+ across the inner membrane (Figure 2-1 B)[16]. Therefore, both proteins are attractive targets for pharmaceutical development[17,18].

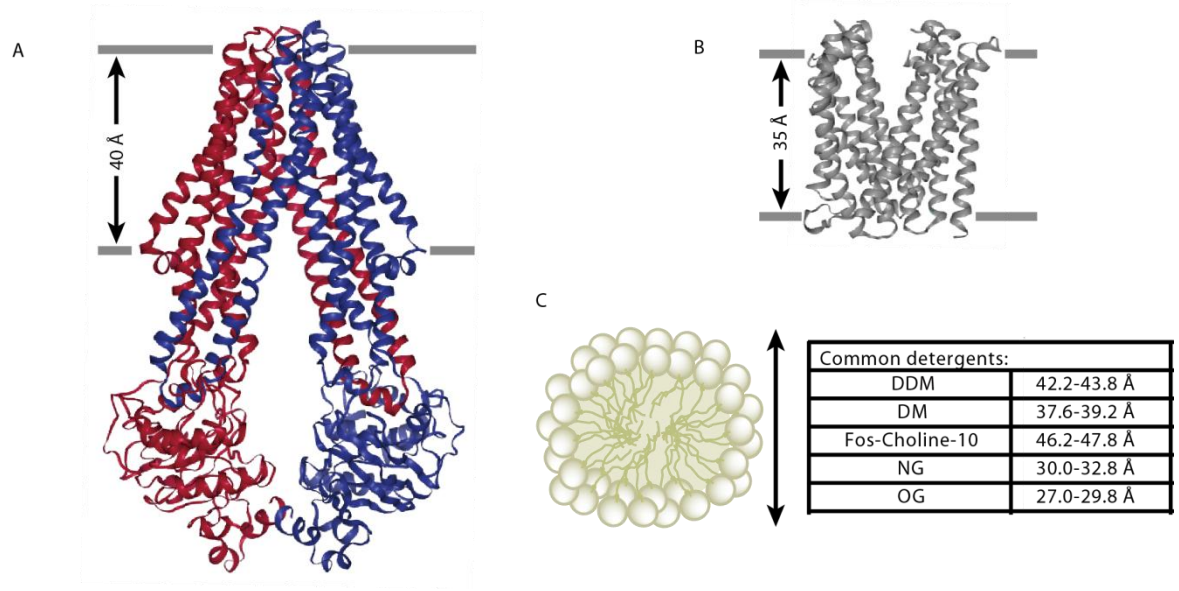


Figure 2-1 Overall structures of MsbA from *E. coli* and MATE from *P. furiosus*. A, Each subunit of homodimeric MsbA contains one TMD forming substrate translocation pathway, and one nucleotide-binding domain for ATP binding and hydrolysis (Protein Data Bank (PDB) ID: 5TV4)[19]. **B,** MATE transporter consists of 12 transmembrane helices forming a pseudo-two-fold symmetric structure with a V-shaped cavity for cation binding[16]. As an analogue of *E. coli* MATE transporter MdtK, PfMATE is presented here in a outward-open conformation.(PDB ID: 3VVN).**C,** the detergent micelle shape and size of common detergent species tested in our experiment[20].

2.2 Delipidation of an IMAC purified membrane protein

In line with previous studies on MsbA purification[19,21-24], recombinant MsbA was extracted from *E. coli* membrane with 1% DDM followed by immobilised metal affinity chromatography (IMAC) and size-exclusion chromatography (SEC) purifications (Figure 2-2 A and B). MsbA was enriched after IMAC purification and isolated from other impurities after SEC purification (Figure 2-2 C). However, the initial native MS

analysis of purified MsbA showed an unresolved broad peak, indicating overlap of heterogeneous masses (Figure 2-2 D). Although the highest possible gas-phase collision voltage of the instrument (240 V) was applied, we could not record a native mass spectrum with resolved charge states[25]. This indicates that the unresolved broad peak in the spectrum was due to multiple lipids attaching to MsbA.

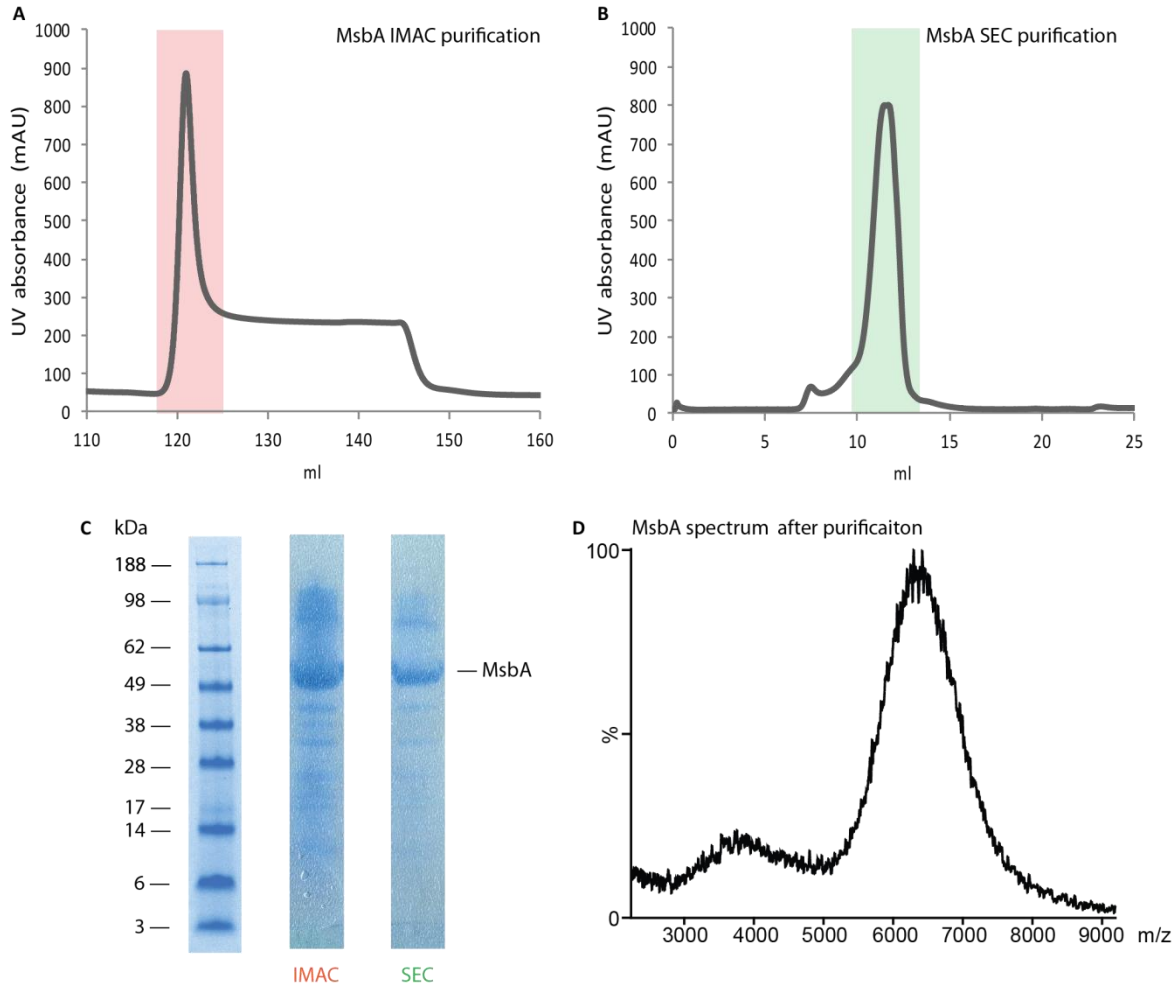


Figure 2-2 The purification of *E. coli* MsbA. A and **B,** The representative chromatograms of IMAC and SEC purification of MsbA. The fractions containing MsbA are highlighted. **C,** SDS-PAGE analysis of MsbA containing fractions in IMAC and SEC. The 50 kDa bands (indicated) were subjected to LC-MS/MS based proteomics identification. **D,** Native MS analysis of purified MsbA in MS buffer containing 0.02% DDM.

To reduce the number of the bulk lipids surrounding MsbA, we introduced a delipidation protocol which incubates purified MsbA in an excess of DDM micelles[8]. During this procedure, lipomicelles arising from the collision between proteomicelles and free DDM micelles are removed by microfiltration (Amicon microcentrifugal filter NMWL 100 kDa)[26] with the empty detergent micelles replenished in the solution at regular intervals. Delipidation of TmrAB was successfully achieved using this method[8]. In that study, all but the most tightly bound lipids were removed during a 48 hours delipidation, and the process was recorded using native mass spectrometry.

We performed delipidation on purified MsbA over different durations. Given that MsbA is less thermostable than TmrAB, the delipidation proceeded at 4°C with rotation. However, MsbA precipitated after 30 hours of incubation, indicating MsbA was denatured at this point. Native MS was then used to analyse delipidated MsbA at shorter incubation times of 1 hour, 4 hours and 14 hours. In line with the control spectrum, the spectrum of MsbA was still unresolved after 1 hour of delipidation (Figure 2-3 A, 1h). As the delipidation time increased, the charge states slightly resolved after 4 hours incubation. Although the peaks were not baseline resolved, two peak series with masses around 67 kDa and 135 kDa were assigned to the MsbA monomer and dimer, respectively (Figure 2-3 A, 4h). The MsbA monomer was visible at 4 hours incubation, and became dominant in the spectra after 14 hours delipidation (Figure 2-3 A). After 14 hours, however, the spectrum was not fully resolved, implying that many lipids were still associated with MsbA. As MsbA precipitated after 30 hours delipidation, this suggests that the MsbA dimer is unstable in solution over a prolonged period, gradually dissociates into monomeric subunits during incubation and is finally denatured. This is quite different to our previous delipidation experiment

using TmrAB, and therefore, this progressive delipidation protocol is not suitable for removing surrounding lipids from MsbA.

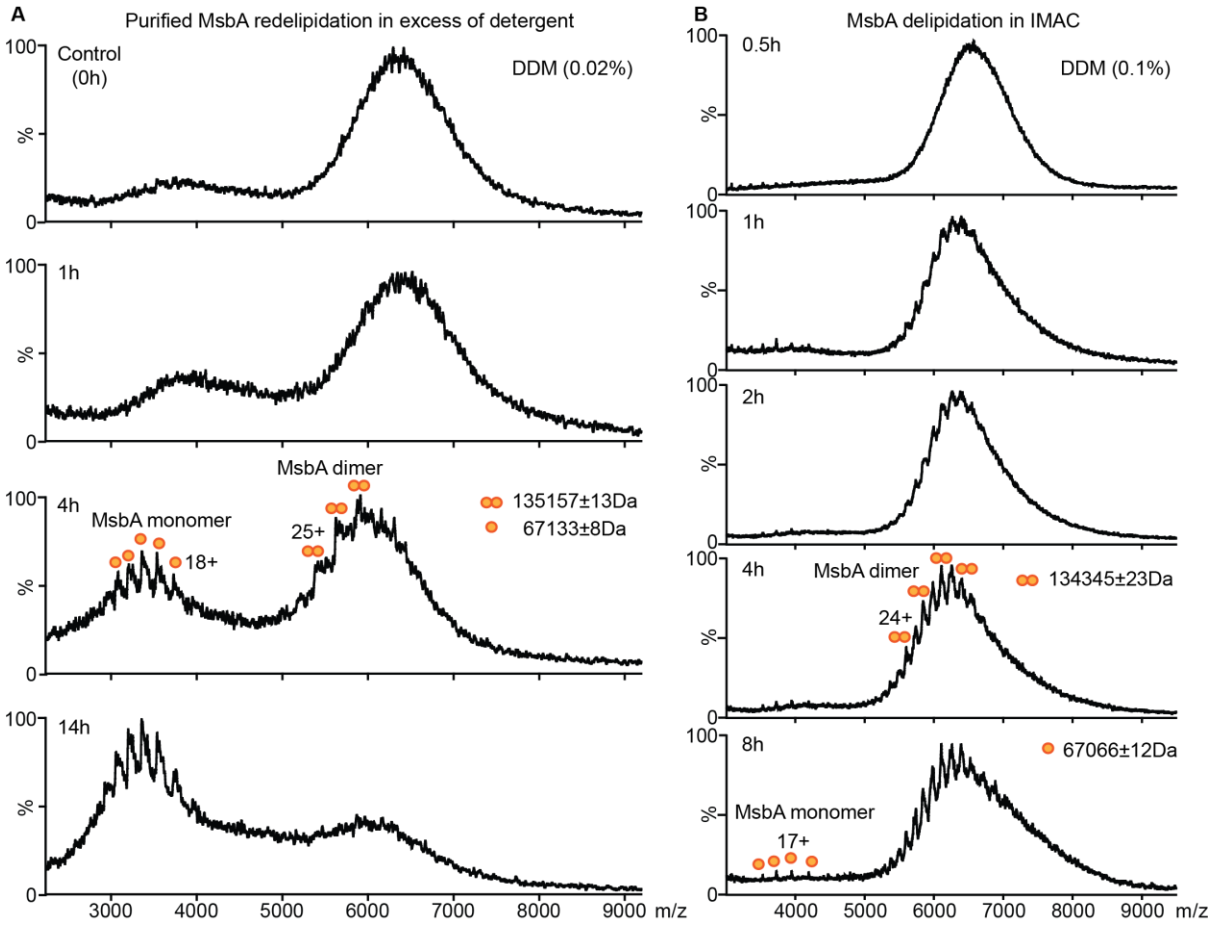


Figure 2-3 Comparisons of two delipdiation approaches for MsbA by native MS analaysis. A, spectra of MsbA delipidated by prolonged incubation (following the established protocol[8]). MsbA sample was incubated with DDM buffer at 4 °C for 0 hours (control), 1 hours, 4 hours and 14 hours with rotation. **B,** spectra of MsbA on-column delipidation in IMAC. The purified MsbA was delipidated on an IMAC column for 0.5 hours, 1 hour, 2 hours, 4 hours and 8 hours followed by SEC re-purification. MsbA monomer (orange filled circle) and dimer (orange filled double circles) were assigned in spectra.

To avoid MsbA oligomer dissociation, we tested different delipidation approaches and recorded the outcomes using native MS. We found that the stoichiometry of dimeric MsbA was maintained when MsbA was immobilised on an IMAC column and delipidated by detergent in an IMAC washing buffer for a prolonged period. Figure 2-3

B shows MsbA delipidated using IMAC over different times from 0.5 hours to 8 hours. The spectra show that the prolonged washing time enabled the recording of a well-resolved dimeric MsbA charge state series. These results indicate that multiple lipids bound to MsbA form the heterogeneous broad peak in the initial spectrum and that these bulk lipids can be removed by detergents. In contrast to the prolonged incubation protocol, the dimeric MsbA was discernible and remained after 8 hours of incubation with DDM with only a low intensity of dissociated MsbA monomer appearing in the spectrum (Figure 2-3 B, 8h). Compared with the theoretical molecular weight of MsbA (67274 Da for the monomer), the masses of MsbA assigned from the partially resolved spectra were roughly 200 Da less, indicating truncation at the terminal of the MsbA sequence. Notably, another charge state series with a larger mass of 3.3 kDa was located between two neighbouring charge states of dimeric MsbA (Figure 2-3 B, 4h and 8h). This tightly bound lipid species survived delipidation, implying a higher preference toward MsbA than those already eluted. This suggests that delipidation by IMAC may be a feasible method of preserving membrane protein bound lipid species.

As shown in Figure 2-3 B, even with an increased DDM concentration from 0.02% to 0.1% used in IMAC, the spectrum of MsbA was still not fully resolved and it was difficult to assign the co-purified lipid species. For a better understanding of the reserved lipids from delipidation, it is crucial to control the extent of delipidation to isolate and identify only tightly binding lipids.

2.3 Condition selection to preserve key lipids in delipidation

We performed a screening to select proper conditions to achieve a controlled delipidation, removing co-purified lipids from the membrane protein successively. Considering the character difference of the two target membrane proteins, we optimised the IMAC washing buffer for MsbA and MdtK individually before starting delipidation condition screening. On-column delipidation for MsbA was performed with 50 mM imidazole in the washing buffer to reduce weakly bound impurities[22,27], and imidazole-free washing buffer was used for MdtK delipidation, since the presence of imidazole led to serious protein loss in IMAC.

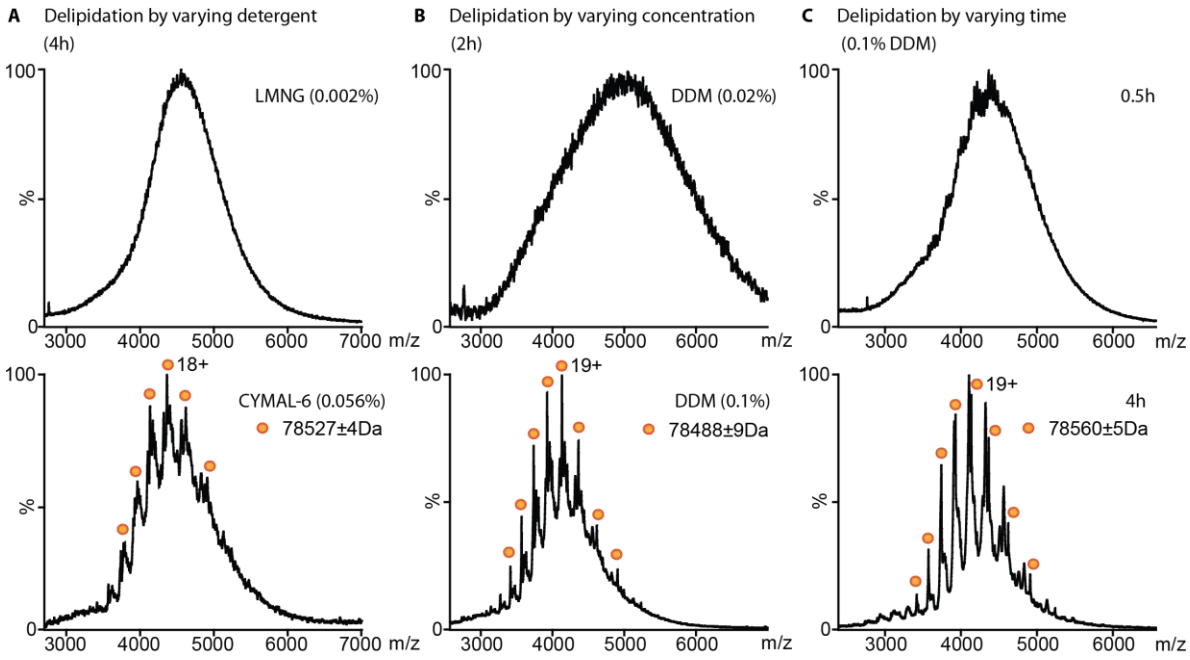


Figure 2-4 Delipidation condition screenings for MdtK delipidation. Representative native mass spectra of MdtK delipidation condition screening by changing detergent type (**A**), detergent concentration (**B**) and delipidation time (**C**). **A**, the spectra of MdtK delipidated for 1 hour by 2 × CMC of LMNG and CYMAL-6. **B,** the spectra of MdtK delipidated for 2 hours by 0.02% and 0.1% DDM. **C,** the spectra of MdtK delipidated by 0.1% DDM for 0.5 hour and 4 hours. The *apo* formed MdtK (orange filled circle) were assigned in discernable spectra with the similar masses to the theoretical mass of MdtK (78511 Da).

A series of delipidation conditions were screened for the delipidation of MsbA and MdtK (Section 2.7). To preserve the interactions between endogenous lipids and membrane proteins, screening usually started with a low concentration of non-ionic mild detergent such as DDM. Here we used 2 × CMC DDM as the initial detergent condition tested for delipidation.

The representative spectra of MdtK delipidated in different conditions are shown in Figure 2-4. The obvious variation between top and bottom panels shows that detergent type, concentration and incubation time are three critical factors which affect the extent of delipidation. Correct detergent type, high concentration and long delipidation time all enabled lipid removal from MdtK and resolution of the spectra of MdtK from broad peak to distinct charge states. With the bound adduct reducing, the remaining lipids became discernible and assignable. Comparing the two bottom figures in B and C, we found that the intensive lipid bound species changed with time in the two spectra, indicating the varied lipid preferences of MdtK, which will be discussed in detail in the following sections.

As shown in Figure 2-4 A, a range of detergents have a different effect on the solubility of the protein and the disruption of the connection between the membrane protein and its endogenous lipids. We screened a variety of different detergents including non-ionic detergents and zwitterionic detergents. From the details listed in methods in Section 2.7, zwitterionic detergents, Fos-Choline-10 and LDAO, led to a serious protein precipitation in delipidation, and the same results were observed using polyoxyethylene glycol detergents C_8E_4 and Triton X-100. The zwitterionic

detergent tends to act as an ionic detergent because of the charges on head groups, and polyoxyethylene units tend to decrease the hydrophobicity of the detergent[28]. Therefore, these two classes of detergent were considered too harsh for membrane protein solubilisation and preservation of co-purified lipids. We also observed that MsbA and MdtK were more stable in a solution of maltoside detergents (DDM, CYMAL-6) than glucoside detergents (OG, NG); this was due to the higher hydrophilicity of the two sugar moieties in the head group of the maltoside detergent[29]. Additionally, the solubility and stability of target proteins were affected by the hydrocarbon chain length of the detergent. Comparing the MsbA delipidation results of DM (C10) and DDM (C12), short chain detergents are typically more deactivating than longer chain species for membrane proteins[30-32]. Delipidation condition screening was also undertaken with new neopentyl glycol class detergents (LMNG, OGNG). The central quaternary carbon with two branched hydrophilic heads and two hydrophobic tails generates subtle constraints on overall conformational flexibility, causing close packing of molecules in micelles, which supplies excessive thermal stability and solubility for membrane proteins[33-36]. However, in our experiment, the protein-LMNG interaction was too tight to be disrupted by the activation, so we were unable to detect the lipid-protein interactions and only obtained unresolved broad peak spectra (Figure 2-4 A, top). Although the glucosyl head groups of OGNG perturb its efficiency in delipidation, we were able to obtain a resolved delipidation result for MsbA from a detergent mixture of DDM and OGNG, which was not achieved using DDM alone (details discussed in the next section).

Determining the influence of varying detergent concentration and incubation time was more straightforward. From Figure 2-4 B and C, a high concentration of detergents

creates more free detergent micelles, accelerating the process of lipid collisional transfer from proteolipomicelles to free micelles, and a long delipidation time enables a better delipidation, demonstrating that lipid transfer from proteolipomicelles to free micelles is a slow process, particularly at 4 °C. In both approaches, a reduction in *m/z* values was observed during delipidation. As the extent of delipidation increased, the binding adducts were gradually displaced by the detergent DDM and dissociated from MdtK. The solvent accessible surface of the protein became larger, and thus more amino acid side chains were exposed in the subsequent activation in the mass spectrometer[37], giving a higher charge state distribution in the spectra.

By adjusting detergent type, detergent concentration and incubation time, therefore, we were able to control the extent of delipidation allowing a progressive delipidation and enabling the capture of membrane proteins in various lipid bound states by native MS. To investigate the co-purified lipid cohort, lipid extraction was coupled with native MS to isolate the remained lipids from each stage of delipidation; these species were identified and quantified in the following lipidomic analysis using RP-LC/ESI-MS/MS[4,8].

2.4 Identification and quantification of membrane co-purified lipids using progressive delipidation-mass spectrometry

2.4.1 The progressive delipidation of MsbA by varying detergent conditions

In detergent condition screening, we noticed that incubation using IMAC over long periods resulted in some sample loss. For MsbA, therefore we performed progressive delipidation by varying detergent composition in a fixed washing time of 4 hours. The progressive delipidation of MsbA was recorded using native MS (Figure 2-5 A). When

MsbA was purified with 0.02% DDM, the weakest concentration, the spectra showed a broad peak due to multiple lipids binding to the protein. When the concentration of DDM was increased to 1%, a proportion of lipid was removed, revealing a series of charge states at the top of the broad peak. The following quantitative lipidomics (Figure 2-5

Figure 2-5 B), showed that the average number of PE and PG bound only decreased from 8.4 to 6.7 per MsbA, while the detergent concentration increased 50 times, indicating that the surrounding phospholipids cannot be removed in a high concentration of DDM. Consequently, the harsher detergent OGNG was introduced. With the increase of OGNG concentration, the number of phospholipids reduced sharply and were almost eluted when OGNG reached 0.5% (Figure 2-5 B), implying that the bound endogenous phospholipids were removed by OGNG. At the same stage, we assigned the *apo* MsbA charge states and two following lipid bound species with masses of 2.7 and 3.5 kDa, respectively in the resolved spectrum (Figure 2-6 A, top), which cannot be identified by lipidomics. Those lipid-bound species disappeared when the concentration of OGNG increased to 1%, and fully resolved MsbA charge states without lipid binding were recorded in the central panel of Figure 2-6 A. In this condition, MsbA was considered to be fully delipidated.

From the relatively large masses, we inferred that those lipids could be assigned to lipopolysaccharide (LPS). Therefore, a silver stain method was applied to identify the remaining LPS from the sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of MsbA delipidated by 1% DDM + 0.5% OGNG and 1% DDM + 1% OGNG, respectively[38]. Comparing the silver stained gel from different lipidation states,

there were two intensive bands near the 3 kDa mark in the MsbA sample from delipidation with 1% DDM + 0.5% OGNG; however this was not observed in the other band (Figure 2-5 C). The gel result is consistent with the native spectrum, proving that MsbA has high preference in binding with LPS.

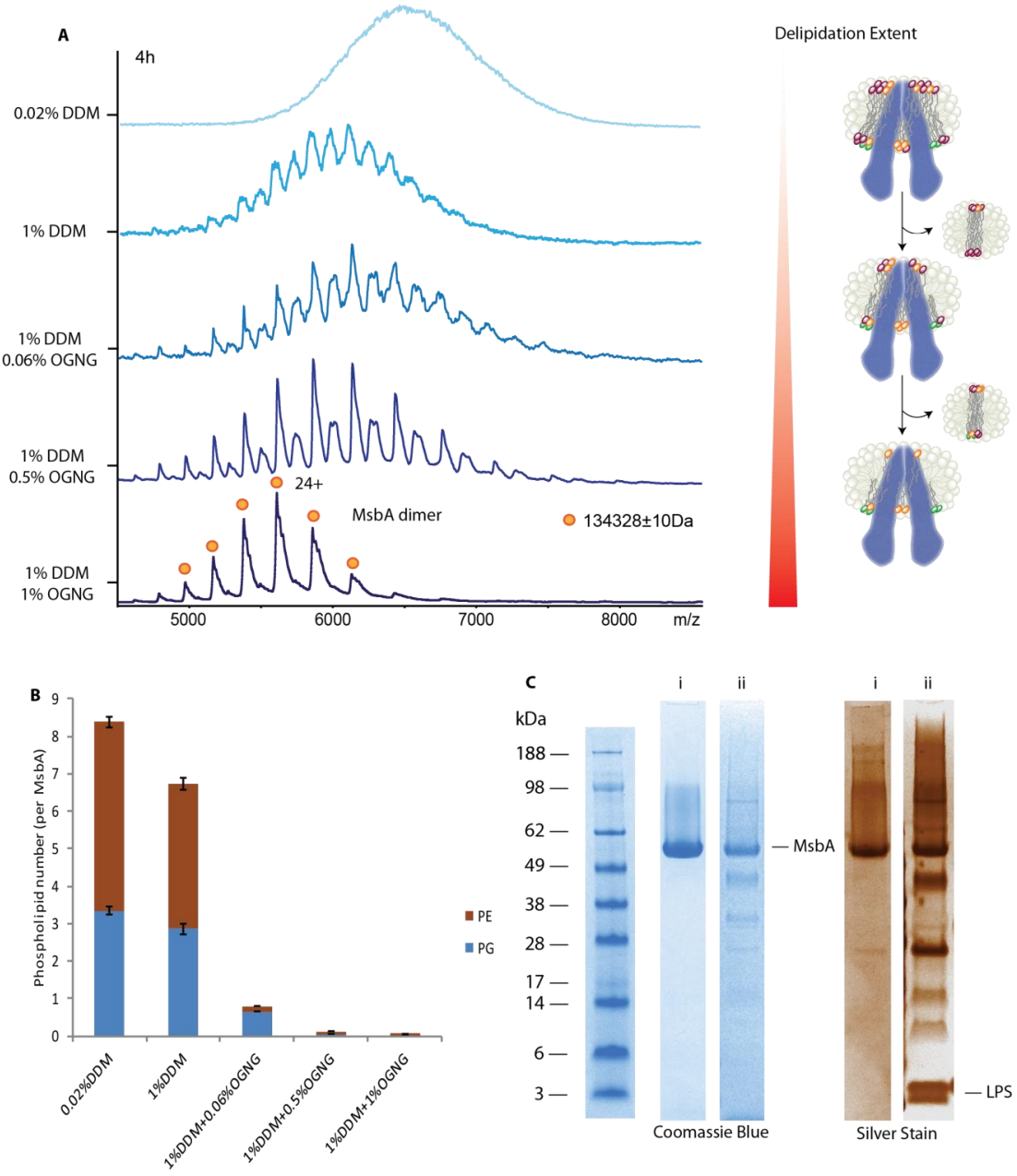


Figure 2-5 The progressive delipidation of MsbA by changing detergent and concentration. A, the native MS spectra of MsbA after progressive delipidation. The native MS finally yielded well-resolved spectra with the detection of *apo* formed MsbA dimer (orange filled circle) after incubation in 1% DDM + 1% OGNG at 4 °C for 4 hours. Here the obtained mass (134328 Da) was close to the theoretical value. **B,** the average number of phospholipids in different conditions of MsbA progressive delipidation quantified by LC-MS with three replications as standard deviation. **C,** the SDS-PAGE 4-12% Bis-Tris protein gel of MsbA delipidated by 1% DDM + 1% OGNG (i) and 1% DDM + 0.5% OGNG (ii) stained with Coomassie Blue and silver.The two bound LPS species were identified in silver stained gel.

The main function of MsbA is to 'flip' the LPS from the inner to the outer side of the inner membrane of bacteria. It has been reported that a hydrophobic pocket and a hydrophilic cavity forms between the two transmembrane domains (TMDs) to accommodate the substrate LPS[19] (Figure 2-6 C), so LPS binds in MsbA with high affinity and specificity. Exogenous monophosphoryl Lipid A (MPLA, 1763.47 Da) was added back into delipidated MsbA (from 1% DDM + 1% OGNG delipidation). In the bottom spectrum of Figure 2-6 A, we can see that lipid A can bind with MsbA in the absence of ATP; this accords with the mechanism of MsbA-mediated LPS flipping proposed in previous research. In a 'trap and flip' model, LPS firstly enters and binds with MsbA, and then the stably bound LPS molecule restricts the conformation for ATP binding[19].

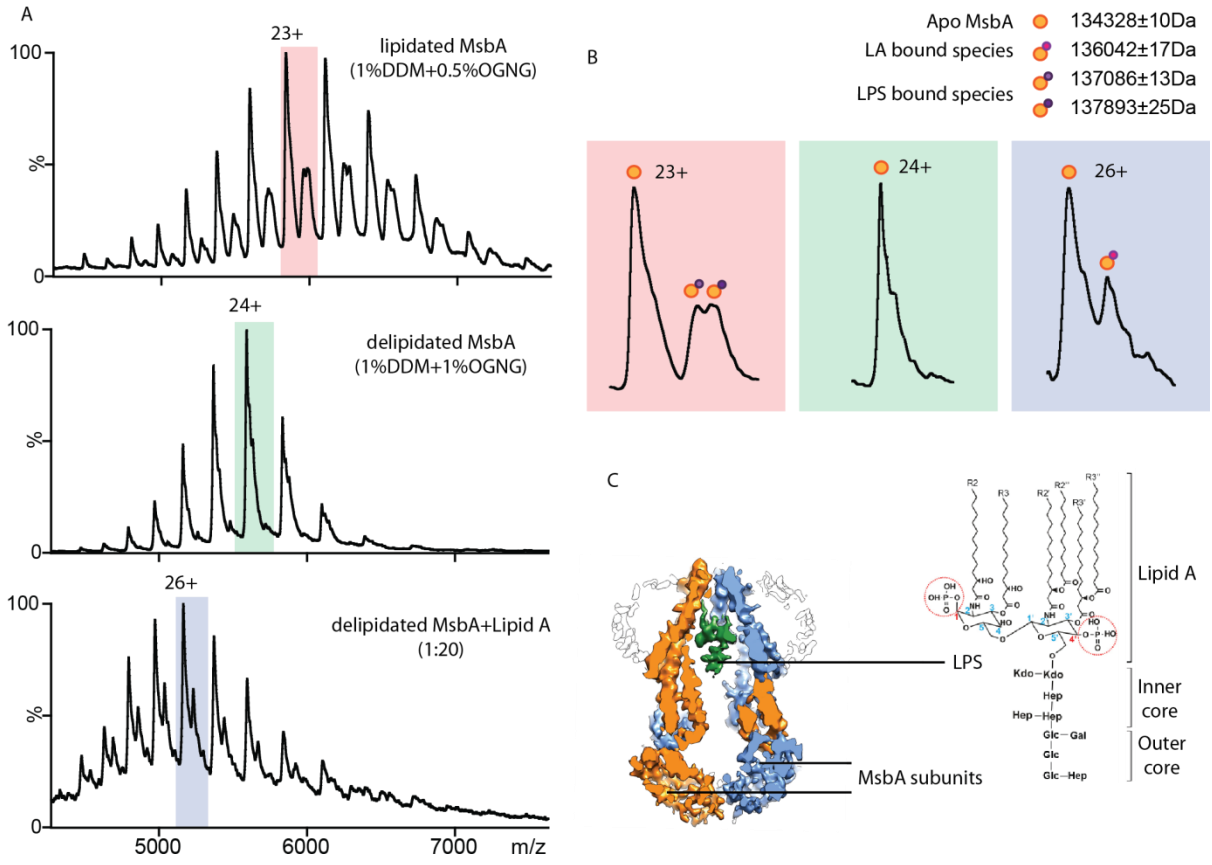


Figure 2-6 MsbA binding with LPS monitored by native MS. A, the native MS spectra of lipidated MsbA (top), delipidated MsbA (central) and MsbA binding with substrate MPLA (bottom). The exogenous MPLA was added into delipidated MsbA at a ratio of 20:1 and incubated at 37°C for 15 min. **B,** The amplified spectra of the highlighted charge states in the three spectra of **A**. *Apo* formed MsbA dimer (orange filled circle), species bound with 1.7 kDa MPLA (magenta filled circle), and species bound with 2.7 kDa and 3.5 kDa LPS (violet and purple filled circle) were assigned. **C,** the cross-sectional view of the MsbA 3D reconstruction from Cryo-EM. LPS (green) was embedded in the cavity formed by MsbA monomeric subunits (blue and orange). The chemical structure of *E. coli* rough LPS was shown on the right. Figure is modified from Gong *et al*[19].

2.4.2 The progressive delipidation of MdtK by controlling delipidation time

Progressive delipidation was then performed using different delipidation times. From the results of the condition screening above, 0.1% DDM was used in IMAC washing buffer. As with the delipidation of MsbA, MdtK was gradually resolved from a multi-lipid bound broad peak to discernible charge states with prolonged incubation

time (Figure 2-7 A). Consistently, the quantitative lipidomics showed a stepwise decrease of average bound phospholipid number from 11 to 0.6 per MdtK (Figure 2-7 B). MdtK was observed to simultaneously carry up to two phospholipids (680 Da) after 2 hours delipidation and two LPS (3.5 kDa) after 4 hours delipidation (Figure 2-7 A). The presence of LPS was also confirmed by the silver stained gel in Figure 2-7 C.

Unlike MsbA, LPS is not a potential substrate of MdtK, and to our knowledge, there are no studies to show that LPS binds to MdtK with high specificity. On the contrary, several researchers report a strong interaction between LPS and membrane proteins overexpressed in *E. coli* cells[39,40]. In these studies, the outer membrane LPS, ranging from 3.2 to 3.7 kDa, was co-extracted with the membrane protein and survived after purification with DDM. It was also reported that OGNG is effective in removing LPS binding. In the condition screening of MdtK, we had already tested OGNG and mixed micelles of DDM and OGNG; we found that the presence of OGNG led to the precipitation of MdtK during IMAC washing, so we did not obtain a clean, well-resolved *apo* MdtK spectrum as we did with MsbA. Therefore, LPS was the membrane adduct co-purified with MdtK, as with other membrane proteins overexpressed in *E. coli.*

A comparison of the spectra of progressive delipidation in MsbA and MdtK shows that, with the same delipidation time of 4 hours, MsbA was slightly resolved at the top of the multi-lipid bound broad peak in 1% DDM (Figure 2-5 A, 1% DDM); MdtK, however, was well-resolved and only LPS still bound in 0.1% DDM (Figure 2-7 A, 4h). Furthermore, MsbA can be delipidated by a high concentration of DDM and OGNG;

MdtK, however, cannot survive in the buffer containing OGNG for long incubation periods (Section 2.7). Moreover, the high concentration of DDM (0.5%) resulted in the auto-cleavage between recombinant MdtK with the green fluorescent protein (GFP) tag.

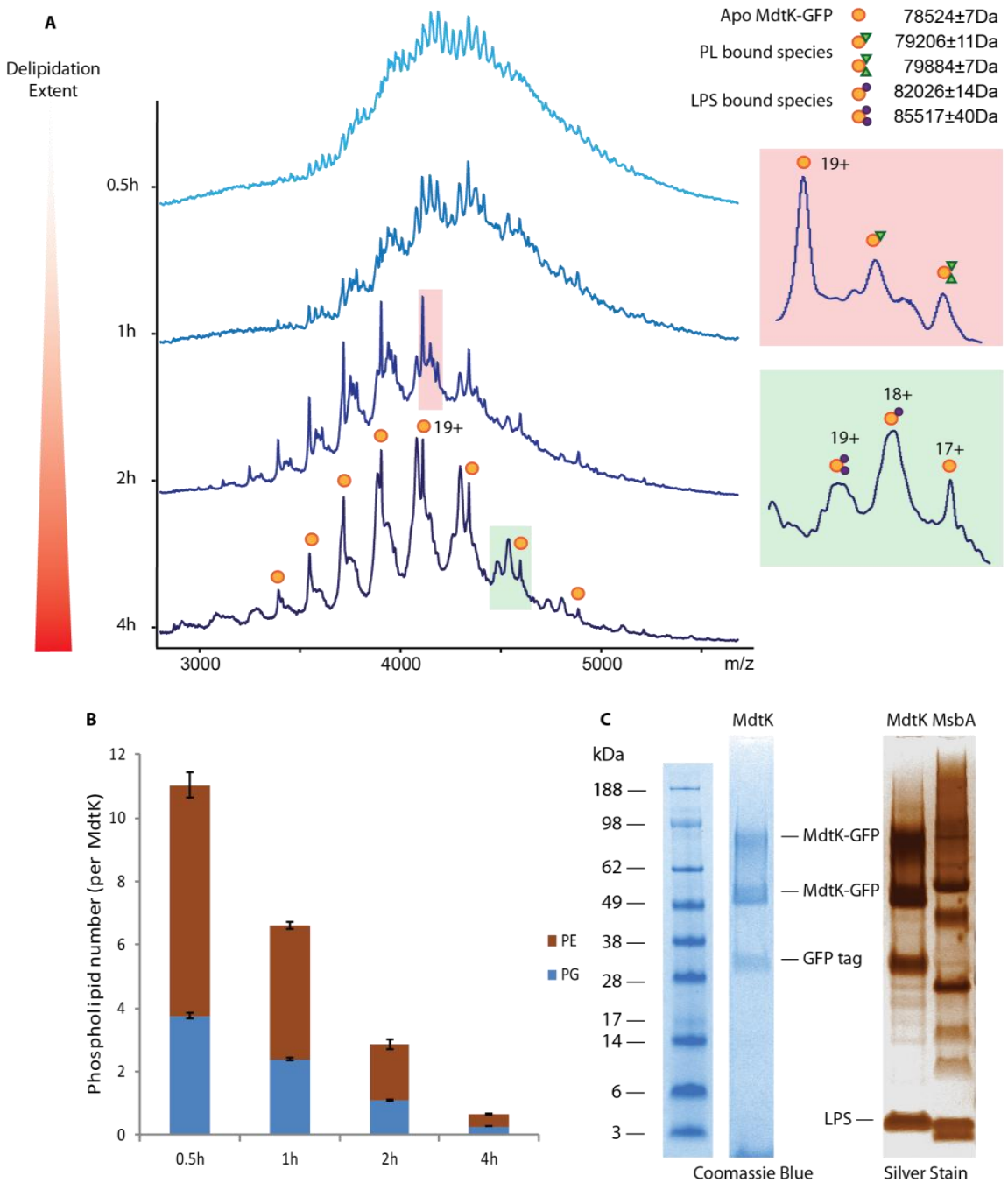


Figure 2-7 The progressive delipidation of MdtK by changing incubation time. A, Native MS spectra of MdtK after progressive delipidation. The native MS yielded well-resolved spectra with the detection of *apo* formed MdtK (orange filled circle) after incubation in 0.1% DDM for 4 hours. Here the mass obtained (78524 Da) was close to the theorectical value. Lipid bound species were from incubating for 2 hours (red) and 4 hours (green), the binding of 680 Da (green triangle) and 3.5 kDa (purple filled circle) lipids should belong to phospholipid and LPS species, respectively. **B,** the average number of phospholipids in different conditions of MdtK progressive delipidation quantified by LC-MS with three replications as standard deviation. **C,** the SDS-PAGE 4-12% Bis-Tris protein gel of MdtK delipidated for 4 hours stained with Coomassie Blue and silver. The protein band was identified by proteomic analysis by LC-MS. The bound LPS species were identified in silver stained gel.

To summarise: we developed a generic progressive delipidation approach to control the process of membrane protein delipidation. Detergent type, detergent concentration and delipidation time are three key conditions for removing excessive lipids from the membrane protein. The optimisation of delipidation conditions for individual membrane proteins is critical for maintaining solubility, oligomerisation and stability.

2.5 Lipidomic analysis of lipid cohort from different stages of delipidation

In the previous section we showed that a number of lipids remain tightly bound to MsbA and MdtK during the process of progressive delipidation. To ascertain the lipid distribution in the surrounding membrane, we identified and quantified the component of co-purified lipids from different lipidated states in progressive delipidation proteins[41]. phosphatidylglycerol (PG) and phosphatidylethanolamine (PE) species with shorter acyl chain length 12:0/13:0 were introduced as an internal standard to quantify the co-purified lipids (Figure 2-8 A). In parallel with the lipidomics, the intact protein

concentration was estimated from absorbance measurements at 280 nm, to calculate the lipid-protein ratio.

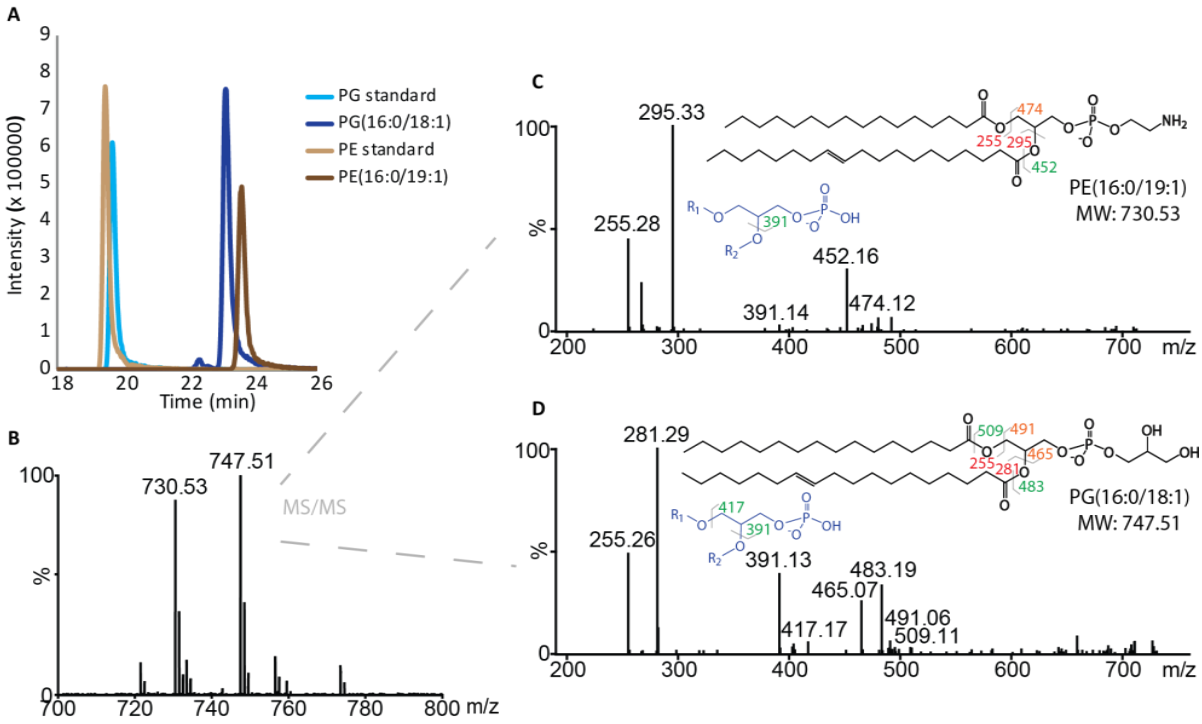


Figure 2-8 The lipidomic analysis of phospholipids bound to MdtK. A, quantification of representative phospholipid species (PG 16:0/18:1 and PE 16:0/19:1) as well as internal standards (PG 12:0/13:0 and PE 12:0/13:0) from progressive delipidation of MdtK. The extracted ion chromatograms of the representative phospholipid species were observed in the MS spectrum **B**. **C** and **D,** MS/MS spectra of the precursor ions of PG 16:0/18:1 and PE 16:0/19:1 respectively from MS spectrum **B.** sn1 and sn2 positions and the double bound position were not characterised. The masses of the resulting fragments were labelled respectively (red, sn $RCOO^-$; orange, neutral loss of RCOOH; green, loss of acyl chains as ketene; blue, loss of glycerol).

MsbA and MdtK from different lipidated stages in progressive delipidation were subjected to LC-MS/MS analysis in the negative ion mode after independent digestion. Due to the limit of detection, we only detected the two major components of phospholipids in *E. coli* membrane: negatively charged PG and zwitterionic PE. In the case of MsbA (Figure 2-9 A and B), a total of 9 PG species and 11 PE species were

identified respectively, with similar acyl chain lengths distributed from 14 to 19 carbon atoms, with the mono-unsaturated double chain 16:0/17:1, the most enriched species in all PGs and PEs. Similarly in MdtK, as shown in Figure 2-9 C and D, a total of 10 PG and 12 PE species were identified respectively, with the same acyl chain lengths ranging from 14 to 19 carbon atoms and the same predominant mono-unsaturated chain 16:0/17:1. In both MsbA and MdtK, all lipid species were reduced as delipidation increased, and no species showed highly preferential binding to target proteins.

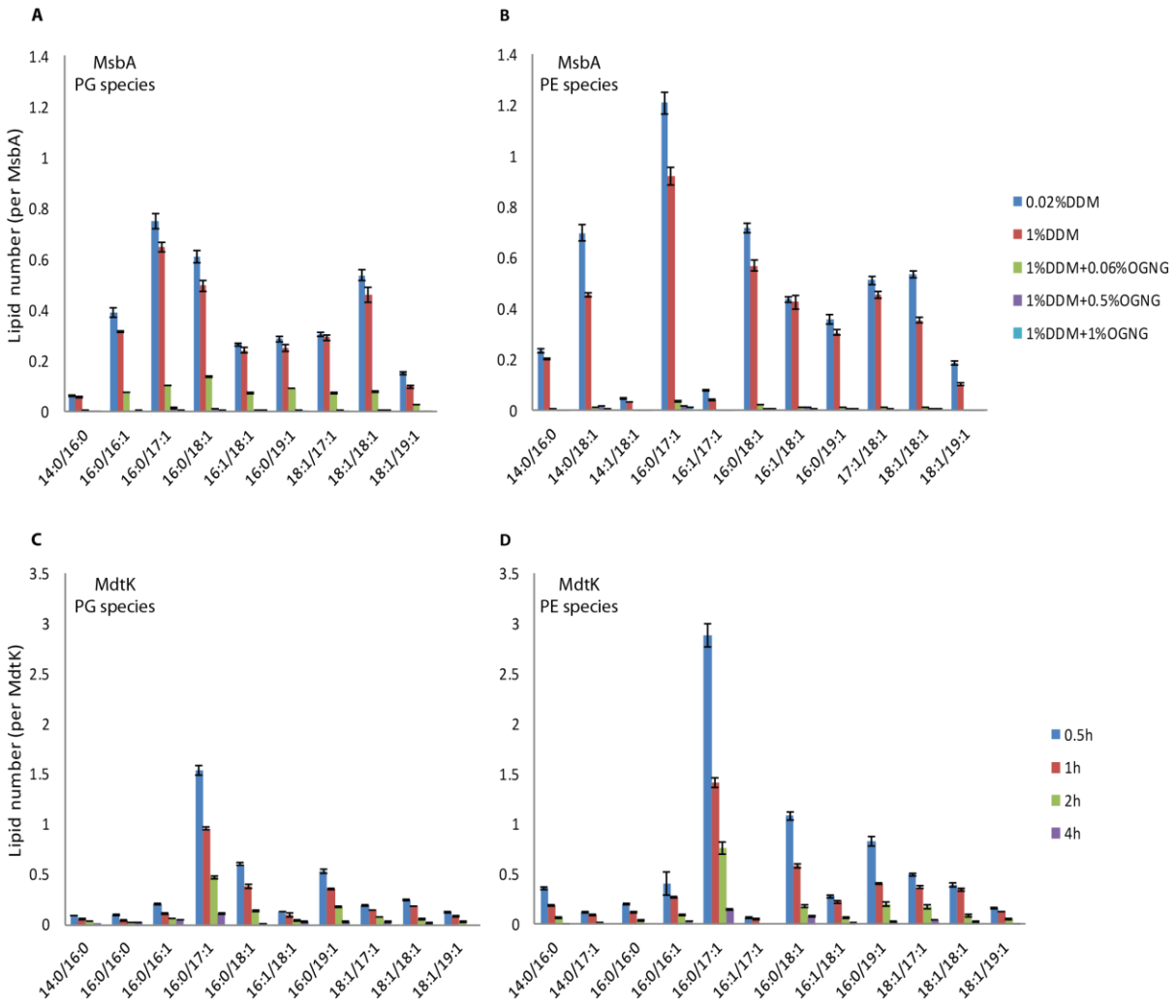


Figure 2-9 Identified phospholipid species bound to membrane protein in the process of progressive delipidation. A and **B,** the phospholipids of MsbA progressive delipidation by varying detergents were listed respectively as PG (**A**) and PE (**B**) species. **C** and **D**, the phospholipids of MdtK progressive delipidation by varying incubation time were listed respectively as PG (**C**) and PE (**D**) species. Data shown are the average number of lipid molecules from each species bound to one protein. Error bars represented the standard deviation from three repeated experiments.

2.5.1 MsbA and MdtK prefer to bind anionic phospholipids

As shown in Figure 2-5 B and Figure 2-7 B, we quantified the phospholipid number per protein in different lipidated states of progressive delipidation. *E. coli* membrane usually has 75–80% PE as the most common lipid, and also contains 15–25% of negatively charged lipids such as PG and cardiolipin (CL)[42-45]. To compare the preference between PG and PE toward MsbA, we normalised the ratio of PG to PE in different conditions during MsbA delipidation. Figure 2-10 A shows that the value of PG/PE increased from 0.27 (the reference value[42] of *E. coli* membrane) to 0.67 after MsbA was extracted from the membrane and delipidated by 0.02% DDM. In this condition, MsbA was still lipidated with over 8 phospholipids (Figure 2-5 B). A comparison with the phospholipid distribution in the membrane indicates there were more PGs allocated in the lipid belt of MsbA. When DDM concentration was added to 1%, PG/PE increased slightly to 0.74. With the introduction of the harsher delipidation detergent OGNG, the PG/PE ratio sharply increased to nearly 5, and the majority of PE species present were removed, leaving only a small number of tightly bound PG species. This trend of ascending PG/PE ratio during progressive delipidation demonstrates that MsbA possesses a greater preference for negatively charged PG than zwitterionic PE. This is consistent with previous studies reporting a general preference of membrane proteins for anionic lipids[46-48].

To confirm the preference of MsbA for PG, two typical phospholipids, exogenous POPG and POPE (PG(18:1/16:0) and PE(18:1/16:0)) were added back into delipidated MsbA (from 1% DDM + 1% OGNG delipidation) in the same protein/lipid ratio (1:10) and incubated for lipid binding. From the spectra acquired by native MS (Figure 2-10 B), dimeric MsbA can bind with multiple PGs and PEs; the binding was quite strong and could not be disrupted even at a high collision energy of 240 V. To quantify the binding intensity difference between PG and PE bound species, we normalised the peak intensity of each lipid bound species, and found that the relative binding intensity of PG was always higher than that of PE (Figure 2-10 C), and the relative difference increased from 26% to 57% with the bound lipid number rising from 1 to 4. This result supported our hypothesis that PG binds preferentially to MsbA.

It is well-established that the local lipid environment is an important factor governing the stability of membrane proteins[2,8]. To assess the effect of lipid binding on maintaining the folded state of membrane proteins in the gas phase, next, IM-MS was applied. The following collision-induced unfolding approaches used in previous studies were employed[2,49]. We plotted the arrival time against voltage to show that the protein is possibly unfolding gradually. Unfolding plots of MsbA binding with POPG were generated from the analysis of the 17+ charge state spectra obtained by collision voltage from 120V to 240V (Figure 2-10 D). *Apo* MsbA, 1-POPG and 2-POPG bound MsbA unfolded around 160V, 170V and 180V, respectively. The more POPG bound, the higher the activation voltage needed for unfolding. Therefore, POPG binding improved the stability of MsbA in the gas phase. Additionally, when we prepared the MsbA sample from progressive delipidation for the MS acquisition, we noticed that the delipidated MsbA precipitated more quickly than the less delipidated

sample in MS buffer, showing that the lipid bound MsbA was more stable in solution. This confirms the importance of bound lipids for the stability of the MsbA in both gas and aqueous phases.

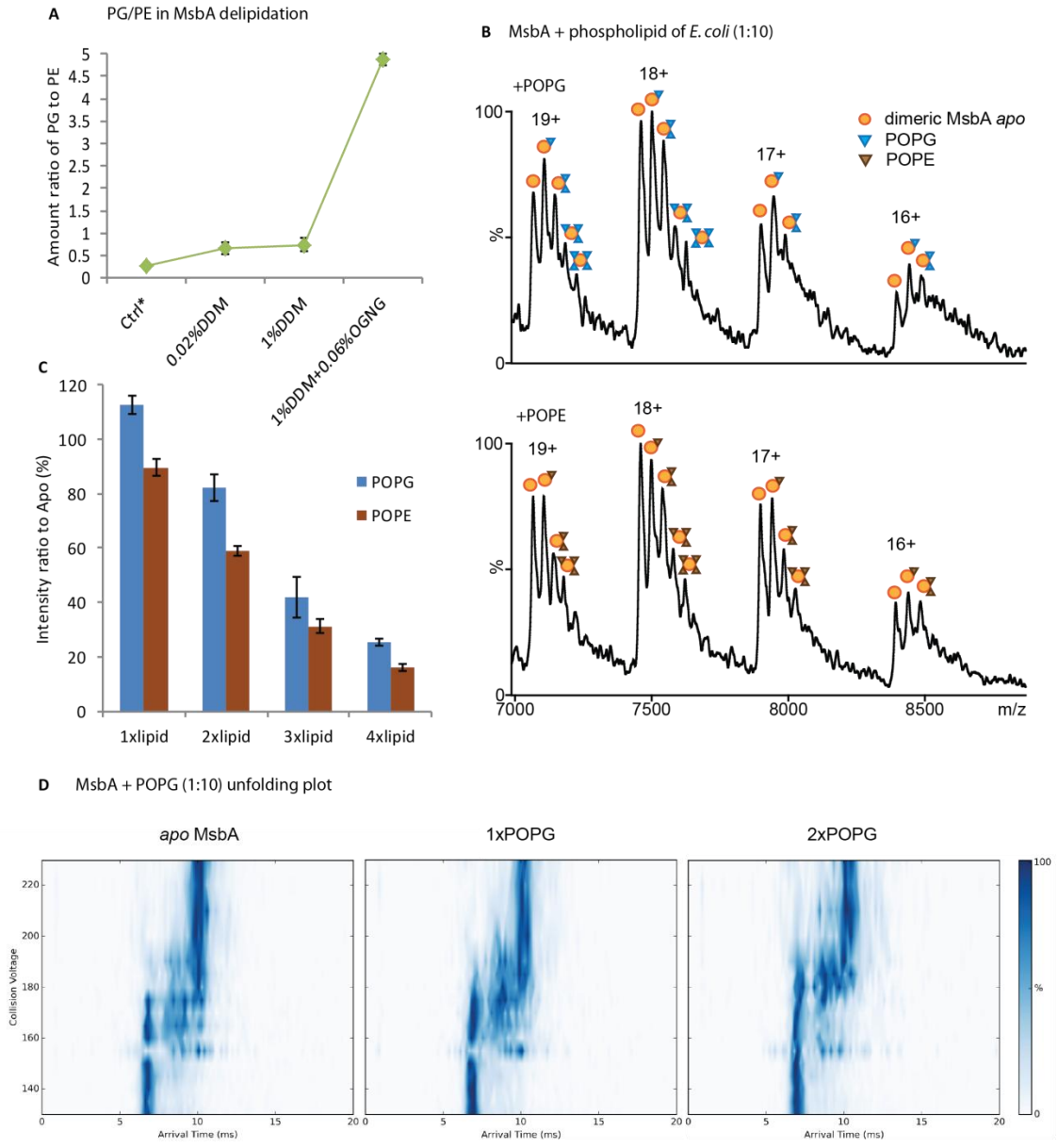


Figure 2-10 MsbA binding with phospholipids. A, the ratio of PG to PE species in different conditions from progressive delipidation of MsbA. The control ratio of PG to PE amount (Ctrl*) was 20%/75% from the reference[42]. Error bars represented the standard deviation from three repeated experiments. **B,** the spectra of delipidated MsbA (orange filled single circle) bound with exogenous POPE (brown triangle) and POPG (blue triangle) respectively as the ratio of 1:10. MsbA was mixed with lipids in SEC buffer and incubated at 4°C for 15 mins and then buffer exchanged into MS buffer for acquisition. **C,** the relative intensity ratio of lipid bound species to *apo* MsbA in **B.** The ratio was analysed using in-house software UniDec (Universal Deconvolution)[50]. Error bars represented the standard deviation of three repeated experiments. **D,** Gas-phase unfolding plots of collision voltage against arrival time for 17+ ions of *apo* to 2 lipid bound MsbA in the spectra of **B.** The ion mobility data were analysed using in-house software PULSAR (Protein Unfolding for Ligand Stability and Ranking)[13].

In the same way that MsbA was used to compare the MdtK preference between PG and PE, we normalised the ratio of PG to PE with different delipidating times. As shown in Figure 2-11 C, the value of PG/PE increased from 0.27 to 0.51 after MsbA was extracted from membrane using 0.1% DDM. With the incubation time extended, the ratio slowly ascended to 0.74 when MdtK was delipidated for 4 hours. Although the rising trend of PG/PE in MdtK from delipidation was relatively weak, we propose that MdtK preferentially attracts PG rather than PE.

To confirm the PG preference of MdtK discovered from lipidomic analysis, as with MsbA, we added exogenous lipids back to purified MdtK to evaluate the binding preference. POPE, POPG and 16:0/18:1 CL were selected as representative lipids from major components of *E. coli* membrane to mix with MdtK individually at the ratio of 5:1. *Apo* MdtK and the three lipid bound species were assigned in the native spectra shown in Figure 2-11 A. In contrast to the other two lipids, the relative

intensity of POPG bound species was higher than PE and CL bound species, and the signal of 2-POPG bound MdtK was detected as well. Comparing 1-lipid bound species in the most intensive charge state (12+), we found that the relative binding peak intensities of POPE and CL were close to each other, but that of POPG was almost twice as high as those of POPE and CL (Figure 2-11 D), indicating, as with MsbA, a strong binding preference of POPG to MdtK.

It is known that PE is a zwitterionic lipid and PG and CL are negatively charged lipids (Figure 1-2). CL possesses two negative charge centres but PG has only one, so the preference of MdtK binding to PG is not due to the negative charge alone. Conversely, CL is larger than PG and PE, and the low binding peak intensity in Figure 2-11 B indicates that the larger the lipid the less favourable it is to approach MdtK. Similar results have been reported in early research, which assumed that the high binding affinity was not just due to the net charge from positive charged residues on the interface of protein and lipid, but also dictated by steric effects[46,47].

Furthermore, a phospholipid competitive binding test was performed to explore the relationship of MdtK binding with PG and PE. Purified MdtK was mixed with POPG/POPE mixture at a ratio of 1:5:5 and incubated in the same way as the single type lipid binding experiment above. In Figure 2-11 E, it was observed that POPG and POPE could associate with MdtK simultaneously, but the binding intensities were varied. We then normalised the peak intensity of POPG and POPE lipid bound species (Figure 2-11 F). The relative intensities of the two species both reduced (POPG bound species: 38.03% to 33.77%, POPE bound species: 17.52% to 10%);

compared to Figure 2-11 D, the intensity of POPG bound species was over three times higher than that of POPE bound species. This result suggests that PG and PE share binding sites on MdtK, and PG is more competitive than PE to occupy the common binding site.

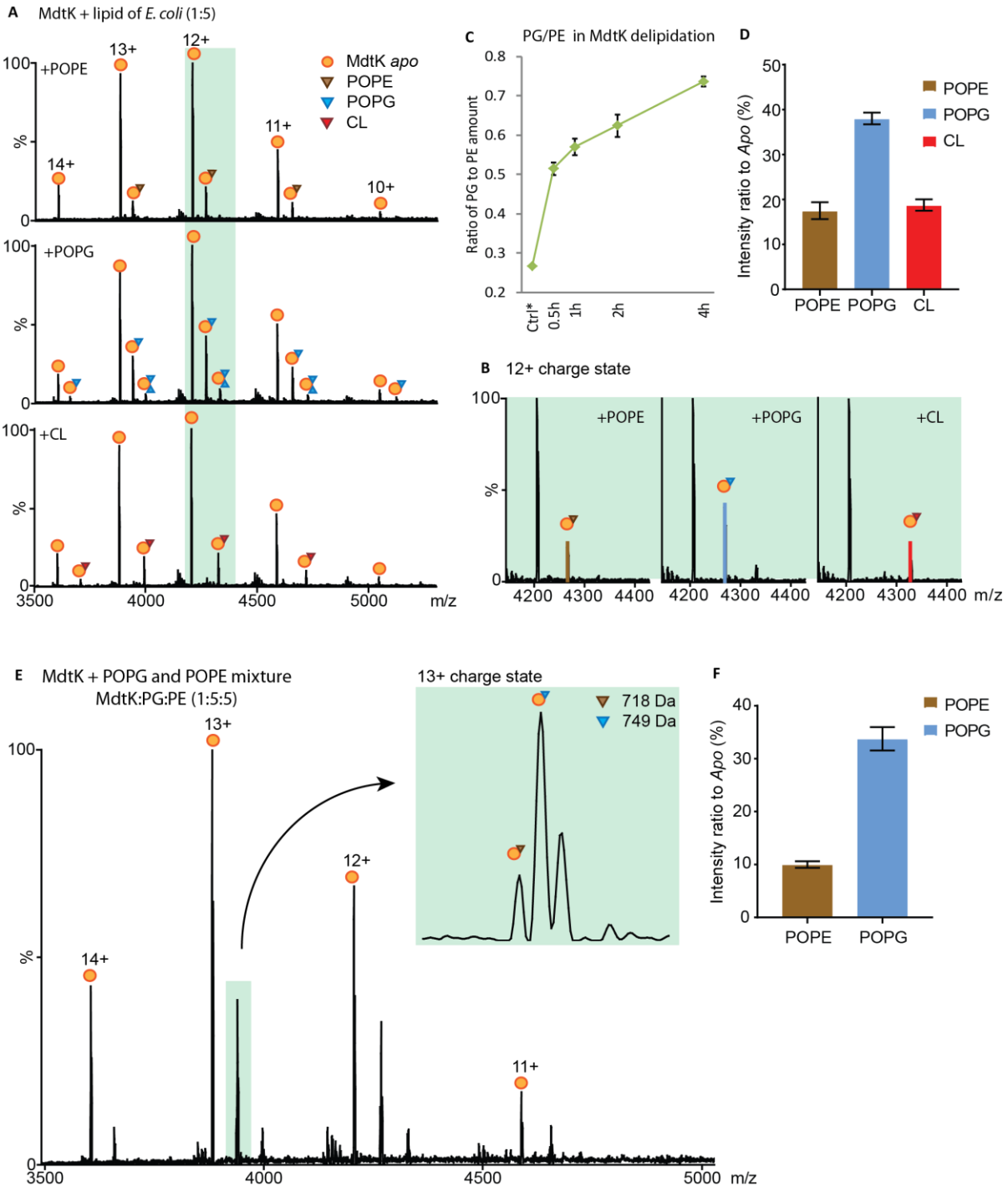


Figure 2-11 MdtK binding with phospholipids. A, the spectra of delipidated MdtK (orange filled single circle) bound with POPE (brown triangle), POPG (blue triangle) and CL 18:1 (red triangle) respectively at the ratio of 1:5. MdtK was mixed with lipids in SEC buffer and incubated at 4°C for 15 mins and then buffer exchanged into MS buffer for acquisition. **B,** the amplified spectra of 12+ charge state highlighted in **A**. **C,** the ratio of PG to PE in different conditions from progressive delipidation of MdtK. The control ratio of PG to PE (Ctrl*) was 20%/75% from the reference[42]. Error bars represented the standard deviation from three repeated experiments. **D,** the relative intensity ratio of lipid bound species to *apo* MdtK in **A** was analysed using UniDec. Error bars represented the standard deviation of three repeated experiments. **E,** the spectra of MdtK bound with the mixture of POPE and POPG at the ratio of 1:5:5 incubated as in **A**. The inset spectrum showed POPE and POPG bound species and the masses of assigned lipids. **F,** the relative intensity ratio of lipid bound species to *apo* MdtK in **E** analysed using UniDec. Error bars represented the standard deviation of three repeated experiments.

2.5.2 The character of acyl chains impact the lipid selectivity of MsbA and MdtK

Next, we investigated the acyl chain length composition of co-purified lipids in different states of progressive delipidation. We calculated and normalised the amount of lipids with each chain length in the corresponding delipidation condition, and compared them with the lipid chain length composition in *E. coli* membrane using the heatmap in Figure 2-12. In general, the phospholipid chain length composition from delipidation was similar to that in the membrane. The dominant phospholipid chain length centred from C32 to C35, and no C29 and C31 lipids were detected. Conversely, a shorter chain length C28 was identified in the membrane implying a slight binding preference of the two proteins toward longer chain lipid species.

A comparison of the lipid chain length between PG and PE from MsbA (Figure 2-12 A and B), shows they share a relatively close chain length, indicating that the chain

lengths of bound lipid were not affected by their different head groups. We also noticed that the chain lengths of both PG and PE did not change obviously during the process of delipidation, suggesting that the lipids with different chain length were allocated randomly around MsbA, and MsbA has no preference towards lipids with specific chain lengths. Similar results were observed in MdtK: the chain length composition of PG was consistent with that of PE, and these compositions did not vary during prolonged delipidation time (Figure 2-12 C and D). However, the percentage of C33 lipids in MdtK was much higher than in other species, which revealed the binding preference of MdtK toward those species. This assumption correlated with the lipidome above (Figure 2-9 C and D), showing that phospholipids with a chain length of 16:0/17:1 were mostly preferred by MdtK.

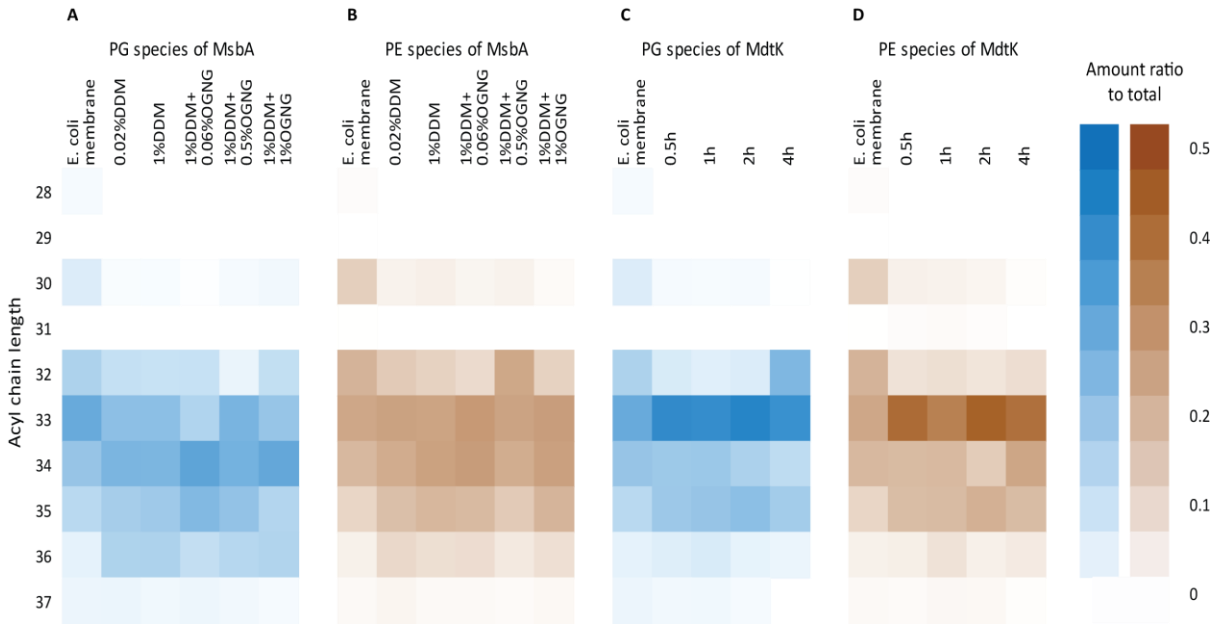


Figure 2-12 The heatmap analysis of the acyl chain length composition of phospholipids from progressive delipidation. A and **B,** the acyl chain length compositions of PG (**A**) and PE (**B**) species bound to MsbA from different conditions of delipidation compared with those in *E. coli* membrane. **C** and **D,** the acyl chain length compositions of PG (**C**) and PE (**D**) species bound to MdtK from different conditions of delipidation compared with those in *E. coli* membrane. The control value of *E. coli* membrane was from the same strain used for membrane protein overexpression. The amount of lipid with different chain lengths was normalised to the total lipid amount in the corresponding lipidating state.

We also inspected the unsaturation degree of the co-purified phospholipid from the progressive delipidation of MsbA and MdtK and compared the differences with *E. coli* membrane. From lipidomic analysis (Figure 2-9), we found that the degree of unsaturation of phospholipids was only 0, 1 or 2, and the two unsaturated bonds were separated into two acyl chains. After normalisation, in Figure 2-13, the monounsaturated phospholipids constituted the majority of the lipid cohort; the amount of saturated species constituted the least and were always less than 10% of the total, which was identical in membrane and co-purified phospholipids of MsbA and MdtK. Conversely, both MsbA and MdtK associate preferentially with unsaturated lipids. Since the co-purified lipids were extracted with MsbA by delipidation, the percentage of saturated PG and PE both reduced to half of that in the membrane (FIgure 2-13 A and B), implying that the lipid belt around MsbA is composed primarily of unsaturated lipids. In MdtK, the ratio of monounsaturated phospholipids is obviously higher than that in the membrane, and the ratio of diunsaturated species reduced (Figure 2-13 C and D). The results indicate that MdtK preferentially binds with monounsaturated phospholipids, and the richest three species of both PG and PE identified by lipidomics were all monounsaturated lipids (Figure 2-9 C and D).

Notably, as in chain length composition above, the composition of co-purified lipids with different unsaturation numbers in MsbA and MdtK was largely invariant in the most and least lipidated membrane proteins. In addition, we also did not observe a distinct difference between PG and PE. Therefore, we propose that the binding preference of MsbA and MdtK toward unsaturated phospholipids was not impacted by the process of progressive delipidation and head groups of phospholipids.

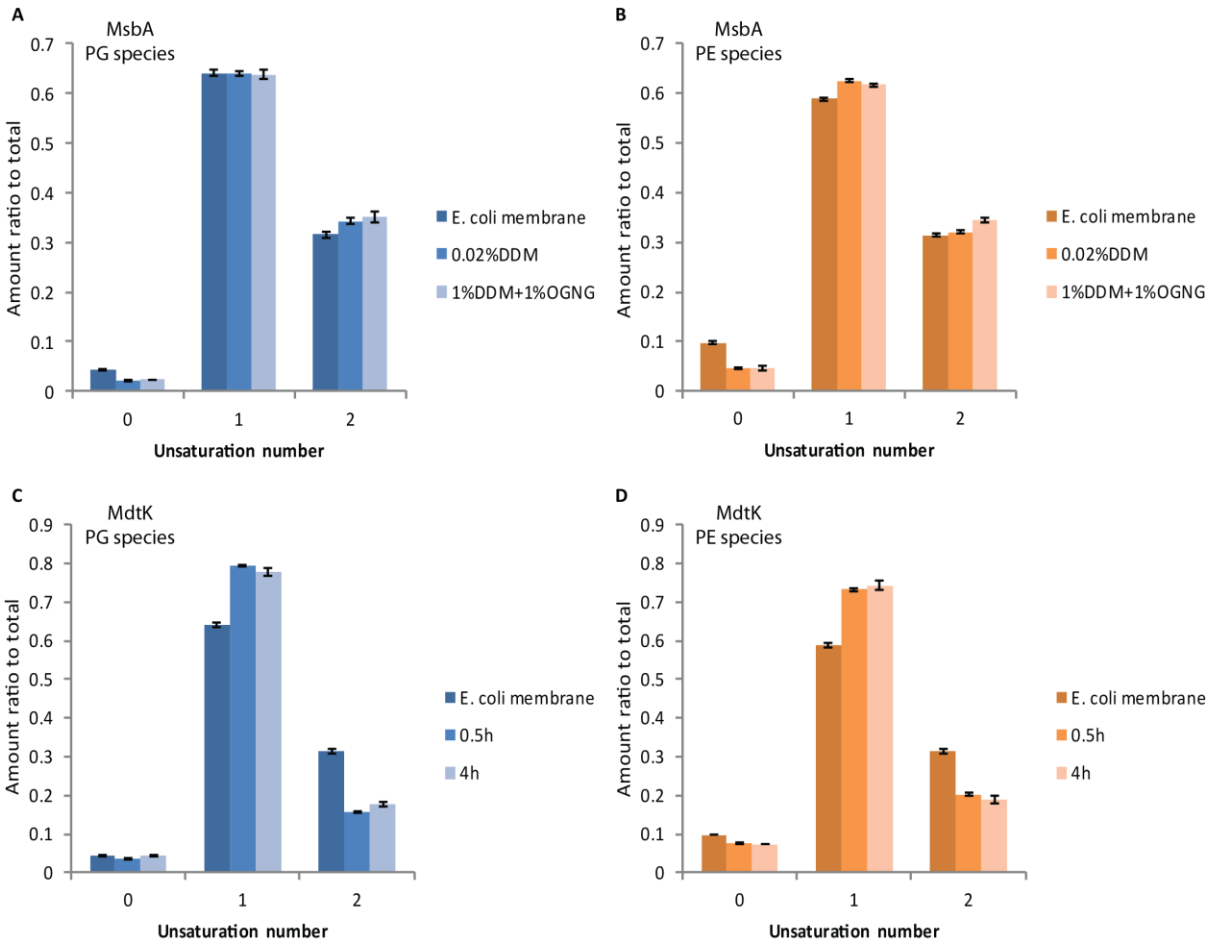


Figure 2-13 The unsaturation degree of the acyl chain of phospholipids from progressive delipidation. A and **B,** the phospholipid unsaturation composition of PG (**A**) and PE (**B**) species bound to MsbA from representative progressive delipidation conditions compared with those in *E. coli* membrane. **C** and **D,** the phospholipid unsaturation composition of PG (**C**) and PE (**D**) species bound to MdtK from representative progressive delipidation conditions compared with those in *E. coli* membrane. The amount of lipid with different unsaturation numbers was normalised to the total lipid amount in the corresponding lipidating state. Error bars represented the standard deviation of three repeated experiments.

2.6 Conclusion

In conclusion, we built up a platform of progressive delipidation coupled with native MS and LC-MS/MS analysis to identify and quantify the endogenous lipid belt of membrane proteins. By controlling the detergent concentration and class, or incubation time, we successfully achieved a stepwise lipid isolation to capture MsbA and MdtK in different lipid bound states at each stage of delipidation, giving a complete view of the entire lipidome of phospholipids co-purified with the two proteins. Using native MS and quantitative lipidomics, we observed the changes of the lipid environment during the process of delipidation, revealing the selectivity and specificity of the target protein toward lipids. we then identified these lipids.

Based on this delipidation protocol, we highlight opportunities for further applications to study the effect of bound lipids on the structure and function of membrane proteins. Recently, a high-energy source platform combined with native MS was developed following progressive delipidation, using multistep tandem MS to visualise, select and identify the lipids directly binding to the protein. This method has already been successfully used with three proteins: LeuT, SemiSWEET and MdtK[4]. Additionally,

IM-MS could be employed after progressive delipidation. By measuring the collision CCS and plotting unfolding trajectories of lipid membrane proteins from different stages of delipidation, we are able to evaluate the modulation of lipids on the conformation and stability of membrane proteins[2]. Biophysical assays such as ATPase activity analysis could be applied with progressive delipidation to explore the impact of binding lipids on the function of membrane proteins. For instance, previous research has shown that the ATPase activity of TmrAB was reduced by 60-70% as a result of delipidation[8], demonstrating the importance of non-specific bound phospholipid on membrane protein.

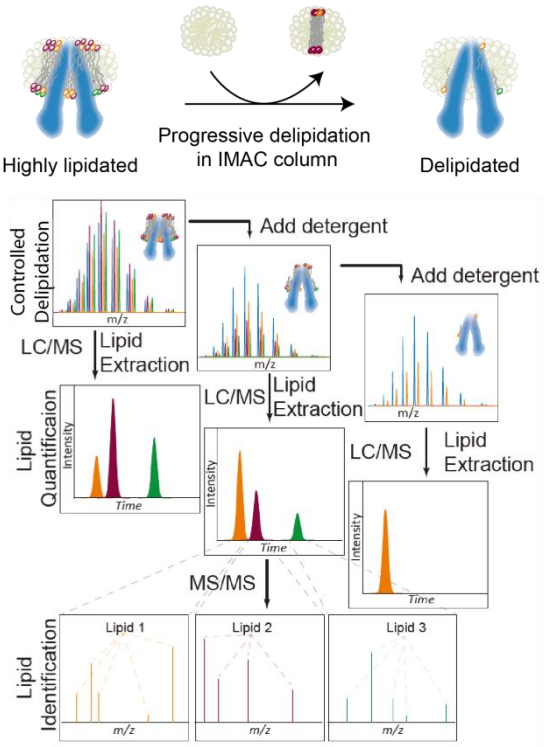


Figure 2-14 Schematic diagram of progressive delipidation of membrane proteins. A stepwise procedure of progressive delipidation in solution, monitored by native MS and coupled with LC-MS/MS to identify co-purified phospholipids[4].

Whilst this contributes to a better understanding of the interactions between lipids and membrane proteins, we should note, however, that the progressive delipidation

protocol was not ideal for all membrane proteins. As different membrane proteins are stable in various detergent conditions, it is imperative to select appropriate detergent conditions for progressive delipidation. This process may be costly in terms of time and consumables and it is possible that more than one condition is feasible for progressive delipidation.

2.7 Methods

Membrane protein expression and purification. Based on established protocols[27], **the** recombinant MsbA gene was cloned into pET-19b vector (Ampicillin resistance; contains an N-terminal decahistidine tag) and the MdtK gene was cloned into pET-28a vector (Ampicillin resistance; contains a C-terminal GFP tag followed by decahistidine tag). Membrane proteins were overexpressed in *E. coli* strain C43 (DE3). A tip of cell stock was pre-cultured in 500 ml of Terrific Broth (TB) supplemented with 0.1 mg/ml sodium salt ampicillin at 37 °C with 180 rotations per minute (RPM) overnight, and 5 ml of precultured cells was transferred into 1 L TB supplemented with 0.1 mg/ml sodium salt ampicillin at 37 °C until the optical density (OD) (D600) reached 0.7-0.8. Protein expression was induced by addition of 0.5 mM isopropyl β-d-1-thiogalacto-pyranoside (IPTG), and continued incubation for 3.5 h at 37 °C. Cells were harvested by centrifugation at 5000g for 10 min at 4 °C, then pellets were re-suspended in lysis buffer (20 mM Tris-Cl, 150 mM NaCl, pH 7.4, with protease inhibitor Complete EDTA free) 25 ml. The re-suspension was passed through a microfluidiser, and centrifuged at 20 kg for 20 min at 4 °C to obtain a soluble fraction. Next, the cell lysates were ultracentrifuged at 27 kg for 2 h at 4°C. Final membrane pellets were re-suspended in solubilisation buffer (20 mM Tris-Cl, 50 mM NaCl, pH 8.0). Membranes were stored at -80 °C. All purification steps were carried out at 4 °C. Membrane proteins were extracted with extraction buffer (20 mM

Tris-Cl, 300mM NaCl, 1% (w/v) DDM, protease inhibitor, pH 8.0) with rotation overnight. The mixture was centrifuged at 13.5 kg for 15 min at 4 °C and filtered through 0.22 μm filter paper to remove unsolubilised residues. Then the membrane protein was purified by IMAC and SEC chromatography in fast protein liquid chromatography (FPLC) at 4 °C. The sample was loaded in a pre-equilibrated IMAC column (HisTrap^{TM} column, 5 ml) and purified with IMAC buffer A (20 mM Tris-Cl, 50 mM NaCl, 50 mM imidazole, 0.02% DDM, pH 8.0), and eluted by IMAC buffer B (20 mM Tris-Cl, 50 mM NaCl, 500 mM imidazole, 0.02% DDM, pH 8.0). The collected protein sample was concentrated at 5000g in concentrator (MWCO 100 kDa) and filtered for purification in SEC column (Superdex 200 10/300 gel-filtration column, 24 ml) with SEC buffer (20 mM Tris-Cl, 50 mM NaCl, 0.02% DDM, pH 8.0). Purified membrane protein was snap frozen by liquid nitrogen and stored at -80 °C.

The progressive delipidation of membrane proteins. The extracted protein sample was divided into aliquots containing protein harvested from 0.5-1.5 L cell culture. In the IMAC purification stage, IMAC buffer A was replaced by delipidation buffers (20 mM Tris-Cl, 300 mM NaCl, 50 mM imidazole, with detergents listed below, pH 8.0) for delipidation. Progressive delipidation was achieved by changing the delipidation conditions: detergent species, concentration and delipidation time (see

Table 2-1 and Table 2-2). The IMAC elusion and SEC purification steps followed the protocols described above.

Table 2-1 Conditions tested for MsbA progressive delipidation.

Detergents	Detergent Con./%	Time/h	Delipidation Result	MS Spectra
C_8E_4	0.5	0.25	precipitation	
DDM	0.02	0.25		unresolved
DDM	0.02	4		unresolved
DDM	0.1	0.25		unresolved
DDM	0.1	0.5		unresolved
DDM	0.1	1		unresolved
DDM	0.1	2		unresolved
DDM	0.1	4		unresolved
DDM	0.1	8		partially resolved
DDM	1	0.25		unresolved
DDM	1	4		partially resolved
DDM+OGNG	0.5+1	0.25		partially resolved
DDM+OGNG	0.5+1	4		denatured
DDM+OGNG	1+0.06	0.25		unresolved
DDM+OGNG	1+0.06	4		partially resolved
DDM+OGNG	1+0.5	0.25		partially resolved
DDM+OGNG	1+0.5	4		resolved
DDM+OGNG	1+1	0.25		resolved
DDM+OGNG	1+1	4		resolved
DM	0.174	0.25	precipitation	

LDAO	0.046	0.25	precipitation	
LMNG	0.002	0.25		unresolved
NG	0.4	0.25	precipitation	
OG	1.06	0.25	precipitation	
OGNG	0.116	0.25	precipitation	
Triton X-100	0.03	0.25	precipitation	

Table 2-2 Conditions tested for MdtK progressive delipidation.

Detergents	Detergent Con./%	Imidazole Con./mM	Time/h	Delipidation Result	MS Spectra
C_8E_4	0.5	50	0.5	precipitation	
CYMAL-6	0.056	50	0.5		unresolved
CYMAL-6	0.056	50	1	sample loss	
CYMAL-6	0.056	0	1		partially resolved
CYMAL-6	0.056	0	2		partially resolved
DDM	0.02	50	0.25		unresolved
DDM	0.02	50	0.5		unresolved
DDM	0.02	50	1	sample loss	
DDM	0.02	0	1		unresolved
DDM	0.1	0	0.5		unresolved
DDM	0.1	0	1		partially resolved
DDM	0.1	0	2		partially resolved
DDM	0.1	0	4		partially resolved

DDM	0.5	0	2		denatured
DDM	1	0	2		denatured
Fos-Choline-10	0.7	50	0.5		denatured
LDAO	0.046	50	0.5	precipitation	
LMNG	0.002	50	0.5		unresolved
LMNG	0.002	50	1		unresolved
LMNG	0.01	0	1		unresolved
OG	1.06	50	0.5	precipitation	
OGNG	0.116	50	0.5	precipitation	
DDM+LDAO	0.015+0.015	50	0.5	precipitation	
DDM+LMNG	0.02+0.02	50	0.5		unresolved
DDM+LMNG	0.1+0.01	0	1		unresolved
DDM+OG	0.015+0.26	50	0.5	precipitation	
DDM+OGNG	0.02+0.02	50	0.5	precipitation	
DDM+OGNG	0.1+0.1	0	0.5	precipitation	
DDM+OGNG	0.1+0.05	0	0.5		partially resolved
DDM+OGNG	0.1+0.05	0	1		partially resolved
DDM+OGNG	0.1+0.05	0	2	precipitation	

Phospholipid extraction for identification and quantification. Following the published method[4,41], the protein sample after progressive delipidation was digested by a 1:50 molar ratio of trypsin overnight at 37 °C, and then the digest dried down with a vacuum concentrator (SpeedVac). To extract lipid from whole cells, cells were washed with solubilisation buffer twice and pelleted by 1000 g centrifugation for 10 min at 4 °C, and re-suspend to a final volume of 1 ml and transferred to a glass vial. 2 ml

methanol and 0.8 ml chloroform were added and sonicated for 10 min in a bath at room temperature (RT). 1 ml chloroform and 1 ml solubilisation buffer were added into the cell mixture, and chloroform phase was collected and dried under a gentle stream of nitrogen in fume hood. The dried lipids were re-suspended with 70% LC buffer A (acenitrile (ACN): H_2O as ratio of 6:4, 10 mM ammonium formate, 0.1% (v/v) formic acid) and 30% LC buffer B (isopropanol (IPA): ACN as ratio of 9:1, 10 mM ammonium formate, 0.1% formic acid) before LC analysis. The sample was sonicated for at least 10 min in a bath at RT. The sample was centrifuged at 10 kg for 1min at 4 °C and the supernatant collected for LC-MS. For lipid quantification, internal lipid standards PG 12:0/13:0 and PE 12:0/13:0 (Avanti Polar Lipids) were added to evaluate the concentration of PG and PE species. The sample was diluted to a final concentration of 10 pg/µl to 1 ng/µl with 70% LC buffer A and 30% LC buffer B.

LC-MS/MS lipidomic analysis of co-purified phospholipid. The prepared lipids were loaded into a 5-µl sample loop using an autosampler using a full-loop method with an overfill factor of 1.4, and then transferred to a 1-µl injection loop using a loading pump at a flow rate of 5 µl/min with 70% LC buffer A and 30% LC buffer B. The sample in the injection loop was injected onto a C18 column (Acclaim PepMap 100, C18, 75 µm × 15 cm, Thermo Fisher Scientific) using a nanopump at a flow rate of 300 nl/min. The lipids were separated on the C18 column at 40 °C by a gradient starting from 70% LC buffer A and 30% LC buffer B. After 10 min, LC buffer B was ramped to 65% over 1 min, 80% over 6 min, and ramped to 99% over 6 min and held for 7 min, then held at 80% for 10 min. The column eluent was delivered via a dynamic nanospray source to a hybrid LTQ Orbitrap mass spectrometer (Thermo Scientific). The conditions used for acquisition were: spray voltage 1.8 kV and

capillary temperature 175 °C. The LTQ-Orbitrap XL was operated in negative ion mode and in data-dependent acquisition with one MS scan followed by five MS/MS scans. Survey full-scan MS spectra were acquired in the orbitrap (m/z 350 − 2,000) with a resolution of 60,000. CID fragmentation in the linear ion trap was performed for the five most intense ions at an automatic gain control target of 30,000 and a normalised collision energy of 38%[4,51].

Native mass spectrometry. For MS analysis, samples fractioned from SEC purification were buffer exchanged into MS spray buffer using biospin column (Micro Bio-Spin 6 chromatography columns, MW exclusion limit 6 kDa). Spray buffer was 1 M ammonium acetate (pH 8.0) with 2 × CMC of DDM (0.02%). 1.5 µL of the solution containing sample was introduced into the ion-source using a gold coated nanoflow capillary needle into the instrument[52]. Spectra were recorded in the positive ion mode in native MS Q-TOF2 (Quadrupole-Ion Mobility-Time-of-Flight) (Waters)[53], and the conditions used for acquisition were: capillary voltage 1.6 kV, cone voltage 180 V, extractor 4 V, collision voltage 150 V, ion energy 7 V, backing pressure 5.5 x 10^{-3} mbar, gas cell 8 psi[25,46]. In ion mobility mode of the IM-MS Synapt3 (G1) (Quadrupole-Ion Mobility-Time-of-Flight) HDMS (Waters), N_2 gas was introduced. IM parameters were set as: trap 6 ml/min, IMS 22 ml/min, trap DC bias 20, IMS wave velocity 300 m/s, IMS wave height 15 V[54]. For the modified Q-Exactive mass spectrometer (Thermo) as described previously[55], the sample was introduced into the instrument and ions were transferred into the higher-energy collisional dissociation (HCD) cell following a gentle voltage gradient: injection flatapole 7.9 V, interflatapole lens 6.94 V, bent flatapole 5.9 V, and transfer multipole 4 V, respectively. The spray voltage was maintained at 1.4 kV at temperature of 200 °C. The optimised

acceleration voltage was 100 V for in-source trapping and HCD cell, and 50 V for source fragmentation. The backing pressure was maintained around 1.00×10^{-9} mbar. Data were analysed using Xcalibur 2.2 SP1.48.

References

(1) Dowhan, W.; Bogdanov, M. Lipid-protein interactions as determinants of membrane protein structure and function. *Biochemical Society transactions* **2011**, *39*, 767-774.

(2) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **2014**, *510*, 172-175.

(3) Lee, A. G. Lipid–protein interactions in biological membranes: a structural perspective. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2003**, *1612*, 1-40.

(4) Gupta, K.; Li, J.; Liko, I.; Gault, J.; Bechara, C.; Wu, D.; Hopper, J. T. S.; Giles, K.; Benesch, J. L. P.; Robinson, C. V. Identifying key membrane protein lipid interactions using mass spectrometry. *Nature protocols* **2018**, *13*, 1106-1120.

(5) Zhang, H.; Harrington, L. B.; Lu, Y.; Prado, M.; Saer, R.; Rempel, D.; Blankenship, R. E.; Gross, M. L. Native Mass Spectrometry Characterizes the Photosynthetic Reaction Center Complex from the Purple Bacterium Rhodobacter sphaeroides. *J Am Soc Mass Spectrom* **2017**, *28*, 87-95.

(6) Landreh, M.; Marklund, E. G.; Uzdavinys, P.; Degiacomi, M. T.; Coincon, M.; Gault, J.; Gupta, K.; Liko, I.; Benesch, J. L.; Drew, D.; Robinson, C. V. Integrating mass spectrometry with MD simulations reveals the role of lipids in Na(+)/H(+) antiporters. *Nature communications* **2017**, *8*, 13993.

(7) Gupta, K.; Donlan, J. A.; Hopper, J. T.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **2017**, *541*, 421-424.

(8) Bechara, C.; Noll, A.; Morgner, N.; Degiacomi, M. T.; Tampe, R.; Robinson, C. V. A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nature chemistry* **2015**, *7*, 255-262.

(9) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature chemistry* **2014**, *6*, 281-294.

(10) Zhong, Y.; Hyung, S. J.; Ruotolo, B. T. Ion mobility-mass spectrometry for structural proteomics. *Expert Rev Proteomics* **2012**, *9*, 47-58.

(11) Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R. Ion mobility mass spectrometry of proteins and proteinassemblies. *Chem. Soc. Rev.* **2010**, *39*, 1633-1655.

(12) Shvartsburg, A. A.; Smith, R. D. Fundamentals of traveling wave ion mobility spectrometry. *Analytical chemistry* **2008**, *80*, 9689-9699.

(13) Allison, T. M.; Reading, E.; Liko, I.; Baldwin, A. J.; Laganowsky, A.; Robinson, C. V. Quantifying the stabilizing effects of protein-ligand interactions in the gas phase. *Nature communications* **2015**, *6*, 8551.

(14) Eckford, P. D.; Sharom, F. J. The reconstituted Escherichia coli MsbA protein displays lipid flippase activity. *The Biochemical journal* **2010**, *429*, 195-203.

(15) Yilmaz, C.; Ozcengiz, G. Antibiotics: Pharmacokinetics, toxicity, resistance and multidrug efflux pumps. *Biochem Pharmacol* **2017**, *133*, 43-62.

(16) Tanaka, Y.; Hipolito, C. J.; Maturana, A. D.; Ito, K.; Kuroda, T.; Higuchi, T.; Katoh, T.; Kato, H. E.; Hattori, M.; Kumazaki, K.; Tsukazaki, T.; Ishitani, R.; Suga, H.; Nureki, O. Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature* **2013**, *496*, 247-251.

(17) Reuter, G.; Janvilisri, T.; Venter, H.; Shahi, S.; Balakrishnan, L.; van Veen, H. W. The ATP binding cassette multidrug transporter LmrA and lipid transporter MsbA have overlapping substrate specificities. *The Journal of biological chemistry* **2003**, *278*, 35193-35198.

(18) Kusuhara, H.; Ito, S.; Kumagai, Y.; Jiang, M.; Shiroshita, T.; Moriyama, Y.; Inoue, K.; Yuasa, H.; Sugiyama, Y. Effects of a MATE protein inhibitor, pyrimethamine, on the renal elimination of metformin at oral microdose and at therapeutic dose in healthy subjects. *Clin Pharmacol Ther* **2011**, *89*, 837-844.

(19) Mi, W.; Li, Y.; Yoon, S. H.; Ernst, R. K.; Walz, T.; Liao, M. Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature* **2017**, *549*, 233-237.

(20) Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Doniach, S.; Columbus, L. Dependence of micelle size and shape on detergent alkyl chain length and head group. *PLoS One* **2013**, *8*, e62488.

(21) Ward, A.; Reyes, C. L.; Yu, J.; Roth, C. B.; Chang, G. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 19005-19010.

(22) Doerrler, W. T.; Raetz, C. R. ATPase activity of the MsbA lipid flippase of Escherichia coli. *The Journal of biological chemistry* **2002**, *277*, 36697-36705.

(23) Buchaklian, A. H.; Funk, A. L.; Klug, C. S. Resting state conformation of the MsbA homodimer as studied by site-directed spin labeling. *Biochemistry* **2004**, *43*, 8600-8606.

(24) Siarheyeva, A.; Sharom, F. J. The ABC transporter MsbA interacts with lipid A and amphipathic drugs at different sites. *The Biochemical journal* **2009**, *419*, 317-328.

(25) Laganowsky, A.; Reading, E.; Hopper, J. T.; Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nature protocols* **2013**, *8*, 639-651.

(26) Strop, P.; Brunger, A. T. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein science : a publication of the Protein Society* **2005**, *14*, 2207-2211.

(27) Newby, Z. E.; O'Connell, J. D., 3rd; Gruswitz, F.; Hays, F. A.; Harries, W. E.; Harwood, I. M.; Ho, J. D.; Lee, J. K.; Savage, D. F.; Miercke, L. J.; Stroud, R. M. A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nature protocols* **2009**, *4*, 619-637.

(28) Rosen, M. J. Surfactants and Interfacial Phenomena. **2004**.

(29) Stetsenko, A.; Guskov, A. An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. *Crystals* **2017**, *7*, 197.

(30) Lund S, O. S., de Foresta B, Champeil P, le Maire M, Møller JV. Detergent structure and associated lipid as determinants in the stabilization of solubilized Ca2+-ATPase from sarcoplasmic reticulum. *J. Biol. Chem* **1989**, *264*, 4907-4915.

(31) Seddon, A. M.; Curnow, P.; Booth, P. J. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et biophysica acta* **2004**, *1666*, 105-117.

(32) le Maire, M.; Champeil, P.; Moller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et biophysica acta* **2000**, *1508*, 86-111.

(33) Chae, P. S.; Wander, M. J.; Bowling, A. P.; Laible, P. D.; Gellman, S. H. Glycotripod amphiphiles for solubilization and stabilization of a membrane-protein

superassembly: importance of branching in the hydrophilic portion. *Chembiochem : a European journal of chemical biology* **2008**, *9*, 1706-1709.

(34) Hoffmann, R. W. Flexible Molecules with Defined Shape?Conformational Design. *Angewandte Chemie International Edition in English* **1992**, *31*, 1124-1134.

(35) McQuade, D. T.; Quinn, M. A.; Yu, S. M.; Polans, A. S.; Krebs, M. P.; Gellman, S. H. Rigid amphiphiles for membrane protein manipulation. *Angew Chem Int Edit* **2000**, *39*, 758-+.

(36) Chae, P. S.; Rasmussen, S. G.; Rana, R. R.; Gotfryd, K.; Chandra, R.; Goren, M. A.; Kruse, A. C.; Nurva, S.; Loland, C. J.; Pierre, Y.; Drew, D.; Popot, J. L.; Picot, D.; Fox, B. G.; Guan, L.; Gether, U.; Byrne, B.; Kobilka, B.; Gellman, S. H. Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nature methods* **2010**, *7*, 1003-1008.

(37) Kaltashov, I. A.; Mohimen, A. Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Analytical chemistry* **2005**, *77*, 5370-5379.

(38) Tsai, C. M.; Frasch, C. E. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal Biochem* **1982**, *119*, 115-119.

(39) Reading, E.; Walton, T. A.; Liko, I.; Marty, M. T.; Laganowsky, A.; Rees, D. C.; Robinson, C. V. The Effect of Detergent, Temperature, and Lipid on the Oligomeric State of MscL Constructs: Insights from Mass Spectrometry. *Chemistry & biology* **2015**, *22*, 593-603.

(40) Mehmood, S.; Corradi, V.; Choudhury, H. G.; Hussain, R.; Becker, P.; Axford, D.; Zirah, S.; Rebuffat, S.; Tieleman, D. P.; Robinson, C. V.; Beis, K. Structural and Functional Basis for Lipid Synergy on the Activity of the Antibacterial

Peptide ABC Transporter McjD. *The Journal of biological chemistry* **2016**, *291*, 21656-21668.

(41) Oursel, D.; Loutelier-Bourhis, C.; Orange, N.; Chevalier, S.; Norris, V.; Lange, C. M. Lipid composition of membranes of Escherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization. *Rapid communications in mass spectrometry : RCM* **2007**, *21*, 1721-1728.

(42) Cronan, J. E. Bacterial membrane lipids: where do we stand? *Annual review of microbiology* **2003**, *57*, 203-224.

(43) Kol, M. A.; van Laak, A. N.; Rijkers, D. T.; Killian, J. A.; De Kroon, A. I.; De Kruijff, B. Phospholipid flop induced by transmembrane peptides in model membranes is modulated by lipid composition. *Biochemistry* **2003**, *42*, 231-237.

(44) G., R. A.; V., C. V.; C., K. M.; L., W. H.; Dowhan, W.; De Kruijff, B. Regulation of Lipid Polymorphism Is Essential for the Viability of Phosphatidylethanolamine-deficient Escherichia coli Cells. *The Journal of biological chemistry* **1994**, *269*, 28670-28675.

(45) Glukhov, E.; Stark, M.; Burrows, L. L.; Deber, C. M. Basis for selectivity of cationic antimicrobial peptides for bacterial versus mammalian membranes. *The Journal of biological chemistry* **2005**, *280*, 33960-33967.

(46) Marcoux, J.; Wang, S. C.; Politis, A.; Reading, E.; Ma, J.; Biggin, P. C.; Zhou, M.; Tao, H.; Zhang, Q.; Chang, G.; Morgner, N.; Robinson, C. V. Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 9704-9709.

(47) Lee, A. G. How lipids and proteins interact in a membrane: a molecular approach. *Molecular bioSystems* **2005**, *1*, 203-212.

(48) Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochimica et biophysica acta* **2014**, *1838*, 1548-1559.

(49) Hyung, S. J.; Robinson, C. V.; Ruotolo, B. T. Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes. *Chemistry & biology* **2009**, *16*, 382-390.

(50) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K.; Benesch, J. L.; Robinson, C. V. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Analytical chemistry* **2015**, *87*, 4370-4376.

(51) Gong, H.; Li, J.; Xu, A.; Tang, Y.; Ji, W.; Gao, R.; Wang, S.; Yu, L.; Tian, C.; Li, J.; Yen, H. Y.; Man Lam, S.; Shui, G.; Yang, X.; Sun, Y.; Li, X.; Jia, M.; Yang, C.; Jiang, B.; Lou, Z.; Robinson, C. V.; Wong, L. L.; Guddat, L. W.; Sun, F.; Wang, Q.; Rao, Z. An electron transfer path connects subunits of a mycobacterial respiratory supercomplex. *Science* **2018**, *362*.

(52) Hernandez, H.; Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols* **2007**, *2*, 715-726.

(53) 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. *Analytical chemistry* **2002**, *74*, 1402-1407.

(54) Ruotolo, B. T.; Benesch, J. L.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nature protocols* **2008**, *3*, 1139-1152.

(55) Gault, J.; Donlan, J. A.; Liko, I.; Hopper, J. T.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.;

Belov, M.; Damoc, E.; Makarov, A.; Robinson, C. V. High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nature methods* **2016**, *13*, 333-336.

3 Lipodisq: a novel detergent-free method to investigate membrane proteins

3.1 Introduction

As discussed in the previous chapter, it is a lengthy and expensive trial-and-error process to select an appropriate condition to preserve the lipid-protein interaction in detergent micelles. Moreover, research has shown that detergent removes lipids from membrane proteins resulting in loss of its biological activity and stability. As a mimic of a lipid bilayer, detergent micelles cannot provide lateral pressure as lipids in membranes, which may lead to the conformational change of membrane proteins or interfere with the membrane protein activity[1,2].

To avoid the influence of detergent we used lipodisqs to extract membrane proteins directly from the original membrane. The amphipathic SMA co-polymer used for lipodisq formation consists of hydrophobic styrene moieties and hydrophilic maleic acid, hydrolysed from styrene-maleic anhydride (SMAnh). The anhydride units are converted into two carboxlyate groups by reaction with the base (NaOH) (Figure 3-1 A)[3]. Both forms of the co-polymer are widely utilised in the plastics industry. In recent years, SMA co-polymer has been found to interact with phospholipids, forming a discoidal polymer-lipid complex[4]. This structure is stabilised by the intercalation of hydrophobic SMA styrene moieties between the acyl chains of two neighbouring lipids in membranes, while the hydrophilic maleic acid faces the aqueous phase (Figure 3-1 B)[5]. Research showed that this encapsulated lipid particle can preserve the biophysical character of the original membrane[5,6], with little preference in specific lipid species[7,8], enabling purification of membrane proteins with their intact native lipid

environment. Therefore, the lipodisq method is a promising tool to investigate lipid-protein interactions and better understand the structure and function of membrane proteins in their native environment.

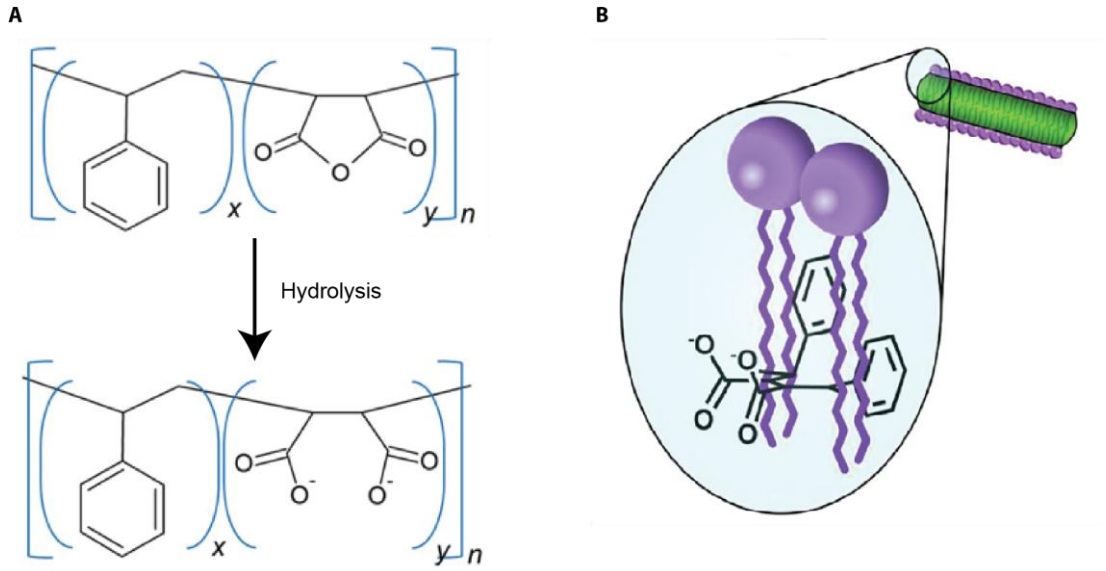


Figure 3-1 SMA co-polymer and interactions with phospholipids in lipodisqs. A, The hydrolysis reaction of SMAnh co-polymer to the SMA co-polymer. Figure is from Rothnie *et al.*[9] **B,** schematic illustration of the phenyl groups of SMA inserted between the acyl chains of two lipids. Figure is from Stroud *et al.*[10].

In the model proposed by Scheidelaar *et al.*, lipodisqs are formed from a self-assembly process with three stages[6]. In the first stage, anionic SMA co-polymers adsorb to the surface of the membrane. Next, SMA co-polymers penetrate into the membrane's hydrophobic core. Finally, the inserted SMA co-polymers lead to the destabilisation of the membrane and formation of lipodisqs surrounded by SMA co-polymer belts (Figure 3-2). This process is driven by the hydrophobic interaction between the lipid acyl chains and the co-polymer styrene moieties[5], and influenced by multiple factors such as SMA concentration, incubation temperature and membrane lipids properties, particularly[6]. Considering the complexity of the lipid environment of

different membrane proteins, the solubilisation conditions of SMA co-polymers need to be investigated.

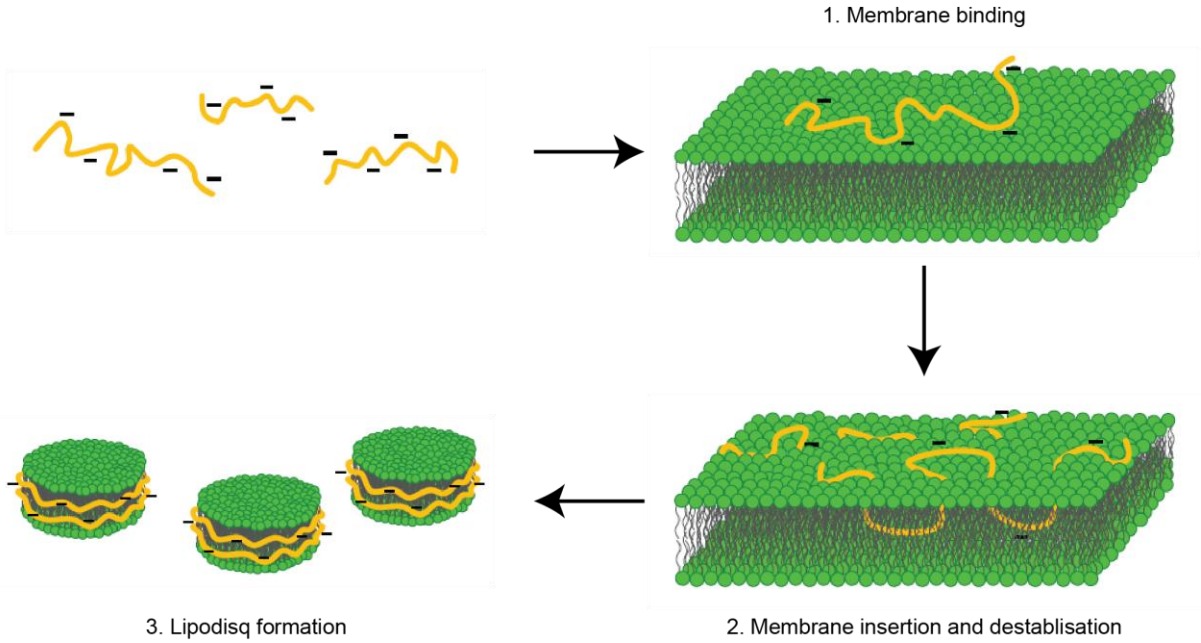


Figure 3-2 The three-stage procedure of membrane solubilisation by SMA co-polymers. Hydrolysed SMA co-polymers bind, insert and destabilise the lipid bilayer. The newly formed lipodisqs are stabilised by the surrounding SMA co-polymer belts. Figure is modified from Dorr *et al.*[3]

The lipodisq method has been successfully applied in a series of biophysical techniques for the investigation of membrane proteins (Section 1.1.3). In this chapter, we attempt to use a native MS approach to investigate the membrane protein lipodisqs prepared in-house, and compare the results with the sample in detergent micelles. With the optimisation of the lipodisq preparation and the adjustment of the MS instrument, we successfully obtained well-resolved spectra of membrane proteins encapsulated in a polymer-lipid disc, and identified their binding lipids, which may be critical for the structure and function of proteins. In our experiment, we selected SMA (2:1) co-polymer to solubilise the parent membrane rather than the other common variant, SMA (3:1) co-polymer. This was because previous research revealed that

SMA (2:1) behaves better in the solubilisation efficiency and in the preservation of membrane physical properties[5,11,12].

3.2 Membrane protein lipodisq formation and characterisation

3.2.1 Lipodisq solubilisation and purification

An initial test was performed to assess the solubilisation capacity of SMA (2:1) co-polymer, and then this was compared with that of detergent, by measuring the fluorescence of the solubilised GFP tagged recombinant MdtK. Aliquots of *E. coli* membrane overexpressing MdtK were incubated with 2.5% SMA (2:1) and 1% detergent DDM individually at room temperature as recommended[9,13]. The mixtures were incubated for 1, 2 and 3 hours respectively and subjected to fluorescence measurement after filtration. The fluorescence intensities of MdtK proteomicelles after incubation for over 2 hours were higher than MdtK lipodisq, but the differences were not obvious (Figure 3-3 A). For MdtK lipodisq, fluorescence intensities were similar over all three time periods. However, an increasing trend of prolonged incubation time was observed in MdtK proteomicelle. This demonstrates that the solubilisation capacity of SMA (2:1) and DDM were similar under test conditions, but that lipodisq formation was much faster than that of proteomicelles.

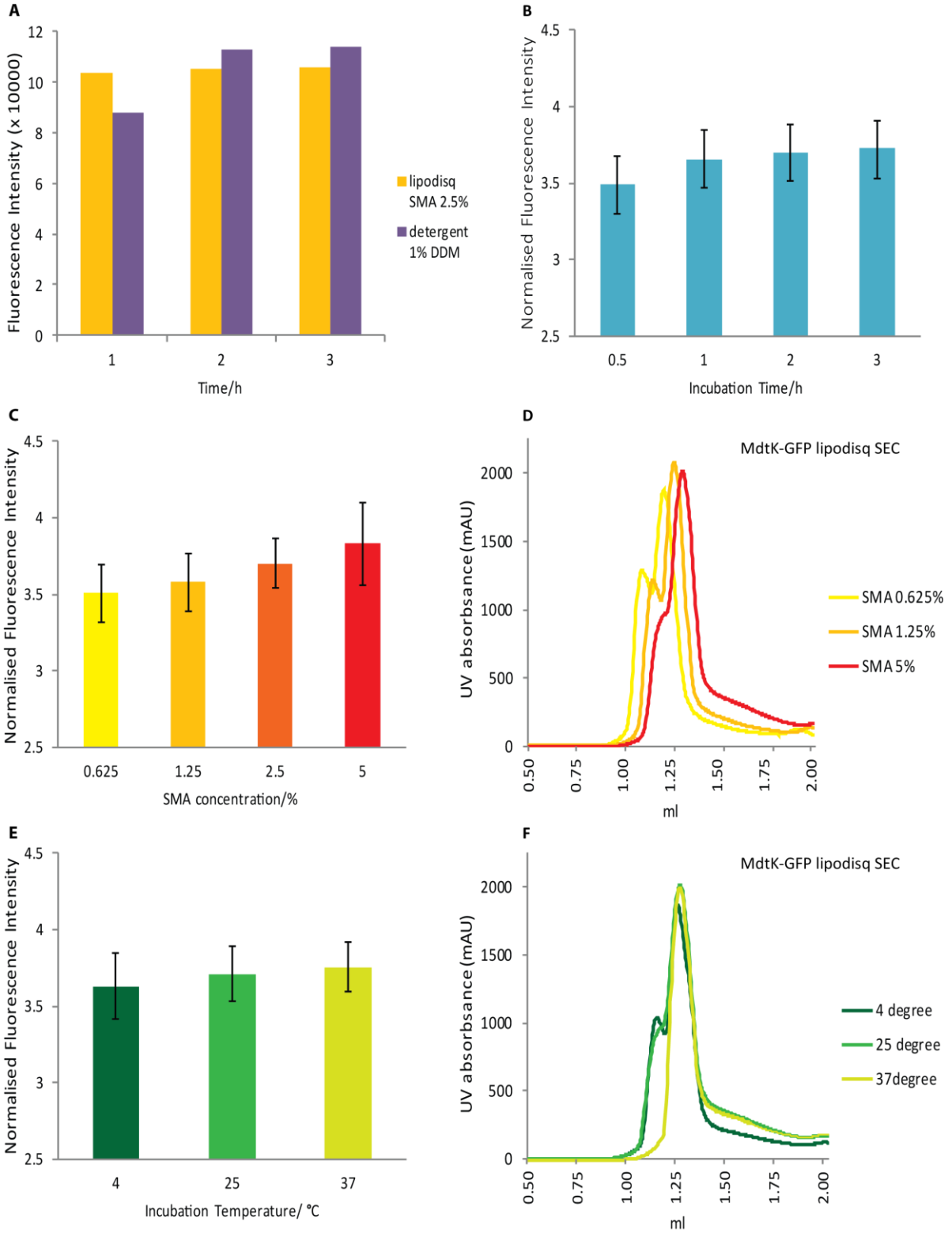

A
Fluorescence Intensity (x 10000)
12
10
8
6
4
2
0
1
2
3
Time/h
lipodisq SMA 2.5%
detergent 1% DDM
B
Normalised Fluorescence Intensity
4.5
4
3.5
3
2.5
0.5
1
2
3
Incubation Time/h
C
Normalised Fluorescence Intensity
4.5
4
3.5
3
2.5
0.625
1.25
2.5
5
SMA concentration/%
D
MdtK-GFP lipodisq SEC
UV absorbsance (mAU)
2000
1500
1000
500
0
0.50
0.75
1.00
1.25
1.50
1.75
2.00
ml
SMA 0.625%
SMA 1.25%
SMA 5%
E
Normalised Fluorescence Intensity
4.5
4
3.5
3
2.5
4
25
37
Incubation Temperature/ ℃
F
MdtK-GFP lipodisq SEC
UV absorbsance (mAU)
2000
1500
1000
500
0
0.50
0.75
1.00
1.25
1.50
1.75
2.00
ml
4 degree
25 degree
37degree

Figure 3-3 MdtK solubilised by SMA (2:1) co-polymer from membrane. A, the fluorescence intensity of crude MdtK incubated with 2.5% SMA (2:1) or 1% DDM at 25°C, after 1, 2 and 3 hours respectively. The solubilised particles were measured after each incubation following filtration through 0.22 µm filter paper after each solubilisation incubation. **B,** the normalised fluorescence intensity of MdtK lipodisq solubilised for 0.5, 1, 2 and 3 hours with 2.5% SMA (2:1) at 25°C. **C,** the normalised fluorescence intensity of MdtK lipodisq solubilised with 0.625%, 1.25%, 2.5% and 5% SMA (2:1) for 2 hours at 25°C. **D,** the combined SEC chromatograms of MdtK lipodidsq from different co-polymer concentrations in **C**. **E,** the normalised fluorescence intensity of MdtK lipodisq solubilised at 4°C, 25°C and 37°C with 2.5% SMA (2:1) for 2 hours. **F,** the combined SEC chromatograms of MdtK lipodidsq from different incubation temperatures in **E**. In **B, C** and **E,** the fluorescence intensities were normalised by negative control, which was the same amount of membrane diluted with IMAC buffer A and treated together with other mixtures for solubilisation. Error bars represent the standard deviation of three repeated experiments.

To optimise solubilisation capacity and efficiency of SMA (2:1) co-polymer on native membrane, we refined the solubilisation condition from three impact factors: incubation time, SMA concentration and incubation temperature. Firstly, we investigated the effect of various incubation times from 0.5 to 3 hours. The increase of fluorescence intensity gradually slowed down with time (Figure 3-3 B). The fluorescence after 1 hour incubation was more intensive than that after 0.5 hour incubation, and was relatively stable for the next 2 hours. Consistently, Figure 3-3 A shows the fluorescence intensities of MdtK lipodisq in 1, 2 and 3 hours were at the similar level. This suggests that SMA (2:1) is able to solubilise MdtK from membrane in a short period of time, and 1 to 2 hours could be enough for the formation of lipodisq. Next we examined different concentration of SMA (2:1), from 0.625% to 5%. No higher concentration was tested due to the limitation of aqueous solubility of SMA. The fluorescence intensity increased with increasing co-polymer's concentration, indicating more MdtK was solubilised (Figure 3-3 C). From the following SEC

chromatography of lipodisq in Figure 3-3 D, increasing SMA concentration led to a slight right shift of lipodisq elution time, which meant a smaller particle formed in SMA with higher concentration. This result was in agreement with the previous research, which reported that the average size of the lipodisq can be adjusted by modifying the ratio of lipids and co-polymers[14]. An unseparated shoulder peak appeared before the main peak of lipodisq, and reduced stepwise with the increase of SMA (Figure 3-3 D). The signal might come from partially solubilised MdtK, which converted into lipodisq particles when more SMA was introduced. Therefore, a high concentration of SMA was favoured for protein solubilisation. Finally different incubation temperatures were introduced. With the rise of temperature, the fluorescence intensity increased slightly (Figure 3-3 E). The partially solubilised signal decreased as the temperature increased (Figure 3-3 F). The shoulder peak disappeared at 37°C, suggesting that MdtK was fully solubilised in a uniform particle size. Moreover, the lipodisq elution time did not change with the temperature, so the temperature had no effect on the lipodisq size. Therefore, a high incubation temperature was preferred to accelerate the rate of solubilisation. This result was supported by the research of Scheidelaar *et al.*, in which the authors proposed that increased SMA concentration and relatively high temperature can promote the speed and yield of vesicle solubilisation[6]. In conclusion, a relatively high SMA concentration (5%) and temperature (37°C) were used for lipodisq preparation in this project.

With optimised solubilisation conditions, several types of membrane protein were successfully extracted from *E. coli* membrane, and the presence of these membrane proteins in lipodisq was confirmed by proteomics of stained gel band (Figure 3-4 A). Below the 47 kDa marker, two bands existed in all membrane protein lipodisq after

IMAC purification; they were identified as CyoA and CyoB by proteomics. Both proteins are subunits of the four-subunit cytochrome *bo*$_3$ ubiquinol oxidase complex, which is a common but vital proton pump in the inner membrane of *E. coli*[15,16]. It is possible for SMA to form CyoA/B lipodisq during solubilisation, and the contaminated lipodisq co-purified together with the target lipodisq because of the non-specific binding of IMAC resin.

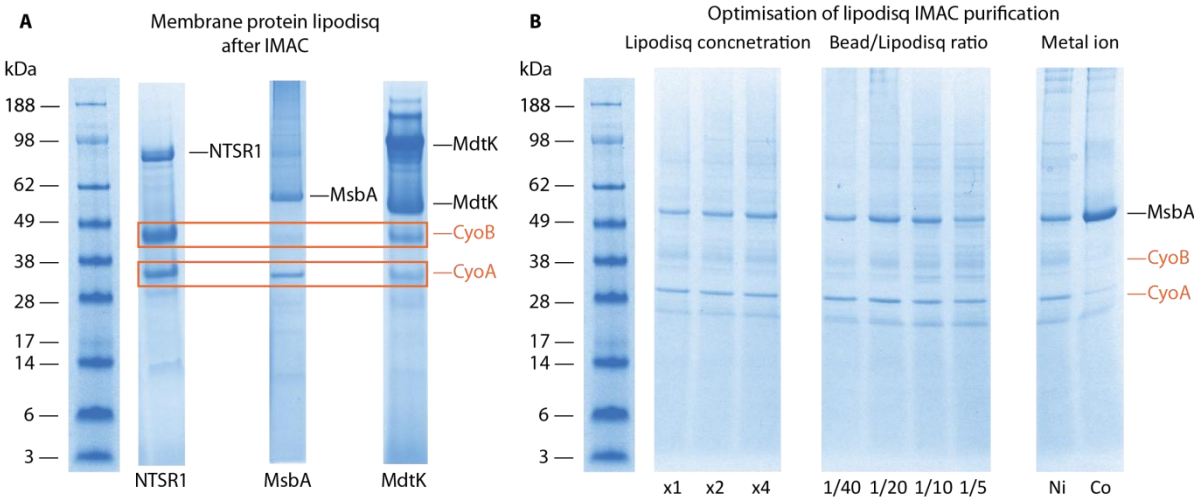


Figure 3-4 The IMAC purification of membrane protein lipodisqs. A, SDS-PAGE analysis of lipodisq of representative self-prepared membrane proteins following the established purification method[9,13]. The indicated bands were subjected to LC-MS/MS based proteomics identification. **B,** the IMAC purification optimisation of MsbA lipodisq. Conditions were adjusted by changing the lidodisq concentration (left panel), the amount of beads (central panel) and the metal ions within beads (right panel).

To reduce the contamination of other protein species, we optimised imidazole concentration in IMAC purification. The concentration test showed that MdtK was being lost from IMAC beads with low imidazole concentration (< 40 mM), while CyoA and CyoB were not fully removed. Similar results were obtained in previous research[17], and also in our preparations of neurotensin receptor 1 (NTSR1) and MsbA. We assumed that the abundant lipid in lipodisq may block the spatial access of

partial protein and interfere with the interactions between His tags and metal ions on beads. Therefore, we adjusted the IMAC incubation conditions from sample concentration, IMAC resin amount and metal species, using MsbA lipodisq as a control experiment. Lipodisq concentration was altered by diluting with IMAC buffer A two and four times before IMAC incubation. The gel pattern of the three samples (Figure 3-4 B, left panel) was identical, revealing that the binding efficiency of IMAC resin was not affected by the protein concentration. Next, MsbA lipodisq mixed with IMAC beads was optimised as a volume ratio from 5:1 to 40:1. The band of CyoA/B faded gradually with reducing of IMAC bead use (Figure 3-4 B, central panel), implying less non-specific bound impurity. Thus the low bead/lipodisq ratio was preferred in impurity removal. We also tested two types of metal ion of IMAC resin, nickel and cobalt. Comparing the purification results of the two metals, MsbA lipodisq captured by cobalt was greater than that captured by nickel and less CyoA/B remained (Figure 3-4 B, right panel). Therefore, cobalt resin was the better option for lipodisq purification.

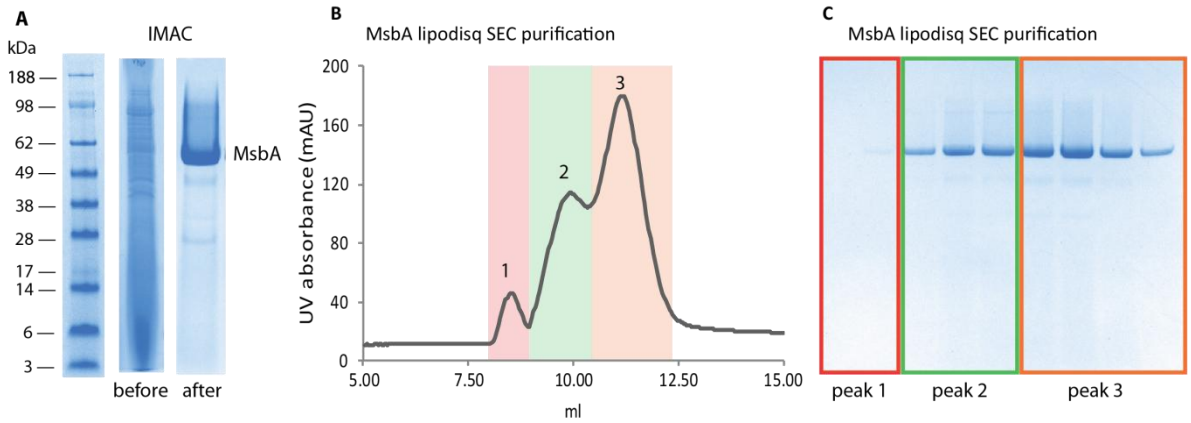


Figure 3-5 The optimised purification of MsbA lipodisq. A, SDS-PAGE analysis of MsbA lipodisq before and After IMAC purification. **B** and **C,** the SEC purification of MsbA lipodisq. The fractions containing different signal peaks in chromatogram (**B**) were highlighted and analysed in SDS-PAGE (**C**), labelled with the same colour as in **B.**

A comparison of the mixture after solubilisation shows that MsbA lipodisq was enriched and isolated from other impurities in optimised IMAC purification (Figure 3-5 A). However, we still got multiple peaks with different intensities in the chromatograph of the following SEC purification, similar to that observed in Figure 3-3 D. For example, 3 different peaks appeared in the SEC chromatograph of MsbA lipodisq in Figure 3-5 B. Fractions from different peaks were collected and subjected to gel analysis (Figure 3-5 C), and no protein was detected in the fractions of peak 1. Since the phenyl moieties of lipodisq scaffold coordinate with the metal ions on the surface of IMAC resin[3,9,13], protein-free lipodisq may co-purify together with His tag-based protein lipodisq through IMAC. These lipodisqs may be detected in SEC afterwards due to the ultraviolet (UV) absorption of the aromatic groups[13,18,19]. The gel of fraction from peak 2 was the same as peak 3: only target protein MsbA was detected (confirmed by proteomics). Thus we assumed that both peaks were signals of MsbA lipodisq, and the different elution time was as a result of the different particle sizes.

We also obtained similar SEC chromatographs of other membrane protein lipodisqs such as MdtK and *P. fluorescens* MATE (PfMATE). To understand the effect of particle size, lipidomics was applied to examine the lipodisqs with different sizes, and the results revealed the identical lipid composition. Additionally, we captured the identical spectra of lipodisqs from different peaks in native MS, indicating that the behaviour of lipodisq protein was not affected by the particle size. Therefore, we suggested that the character of membrane protein in lipodisq is not dependent on disc size.

3.2.2 Evaluation of membrane protein lipodisq

It is known that the quantification of a membrane protein is quite challenging. For the protein in lipodisq, the concentration measurement was affected by the surrounding lipid and SMA co-polymers in some methods[19,20]. To obtain an accurate concentration of lipodisq membrane protein, three different methods to quantify membrane protein were addressed here to measure the concentration of native lipodisq MsbA.

Methods	MsbA-1 (mg/ml)	MsbA-2 (mg/ml)	MsbA-1c (mg/ml)	MsbA-2c (mg/ml)	r^2
UV280 absorbance	0.3	0.45	5.91	9.20	
Pierce™ 660nm assay	0.032	0.066	1.240	2.260	0.999
SDS-PAGE assay	0.108	0.191	1.259	2.388	0.997

Table 3-1 Comparisons of three approaches for the quantification of membrane protein in lipodisqs. MsbA-1 and MsbA-2 were two samples from two batch preparations. MsbA-1c and MsbA -2c were the concentrated samples of MsbA-1 and MsbA-2. The coefficient of deternimation (r^2) was calculated from the standard curve of BSA concentration measurement**.**

UV spectrophotometry absorption at 280 nm is a fast and easy method, and commonly used for protein quantification[12,13,21,22]. However, as mentioned previously, the aromatic group on the lipodisq scaffold has notable absorbance at 280 nm and therefore interferes with UV detection of membrane protein. The MsbA concentrations estimated by UV absorption were always much higher than those using other methods (Table 3-1), which demonstrates that it is inappropriate to quantify SMA solubilised membrane protein using UV absorbance[7,19].

Pierce™ 660nm Protein Assay is a commercially available kit which shows rapid response in the presence of high concentrations of most detergents, reducing agents and other commonly used reagents[23]. This colourimetric method quantifies protein by reference to absorbance of a series of standard protein dilutions assayed alongside the unknown target, which is similar to the third protein quantification method, SDS-PAGE assay. In this method, protein concentration is estimated via densitometric comparison of the protein bands after staining with the concentration gradient of standard protein[7]. Bovine serum albumin (BSA) is used as standard protein to quantify two batches of lipodisq MsbA in the two assays.

We measured the same lipodisq MsbA by 660 nm absorption and SDS-PAGE assays. In the low concentration sample, the results were quite different in the two different methods, but the sample concentrations obtained were similar after 10 times concentration in the original volume. Thus the protein quantification result from the SDS-PAGE assay was more reliable. Since the structure of lipodisq is destroyed in the process of electrophoresis, lipids, SMA molecules and other potential impurities in solution are separated from MsbA in SDS-PAGE gel. Furthermore, protein made contact with staining reagent more sufficiently in the denaturing environment, which supplied a more accurate concentration by densitometric measurement. In summary, SDS-PAGE is a more accurate method of obtaining protein concentration in a complicated system like lipodisq.

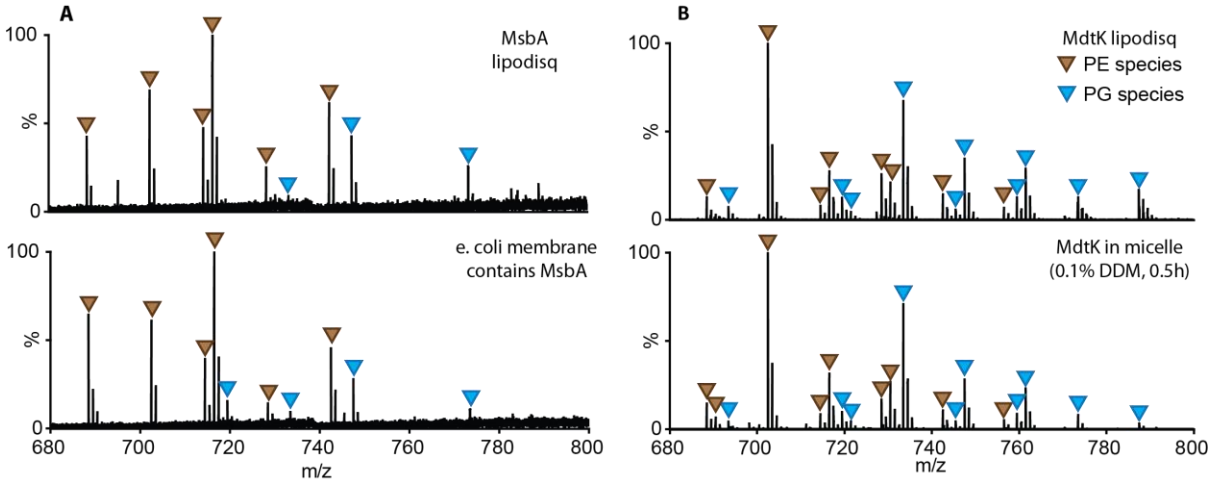


Figure 3-6 The phospholipid species in native lipodisq, membrane and proteolipomicelles. A, the LC-MS spectra of phospholipid species in MsbA native lipodisq and membrane of *E. coli* (C43) cells overexpressing MsbA. **B,** the LC-MS spectra of phospholipid species in MdtK native lipodisq and MdtK in DDM micelle. MdtK in detergent had limited delipidation in a mild conditions (0.1% DDM for 0.5 hours) and was purified following the progressive delipidation protocol in Section 2.7. Lipidomics was performed following the procedure described in Section 2.7.

Lipidomic analysis was performed to examine the local lipid environment of the membrane protein in lipodisq (Figure 3-6). The MsbA lipodisq spectrum was identical with that of the membrane from the expressing cell, indicating that lipid compositions in the two systems were similar (Figure 3-6 A). This is in agreement with previous studies revealing that the SMA co-polymer does not preferentially solubilise specific lipid species[6,7].

We also investigated the composition of lipids from native lipodisq and lipidated membrane protein in detergent. MdtK in micelles was from a mild condition of progressive delipdiation, in which multiple lipids including PGs and PEs were preserved after delipidation and purification (Figure 2-4 C, top). The similar spectra revealed that MdtK native lipodisq and proteolipomicelle possessed the same lipid

composition (Figure 3-6 B), implying that both of the two methods solubilise membrane protein from original membrane directly and maintain a near-physiologic environment for embedded membrane protein, critical for structural and functional characterisation.

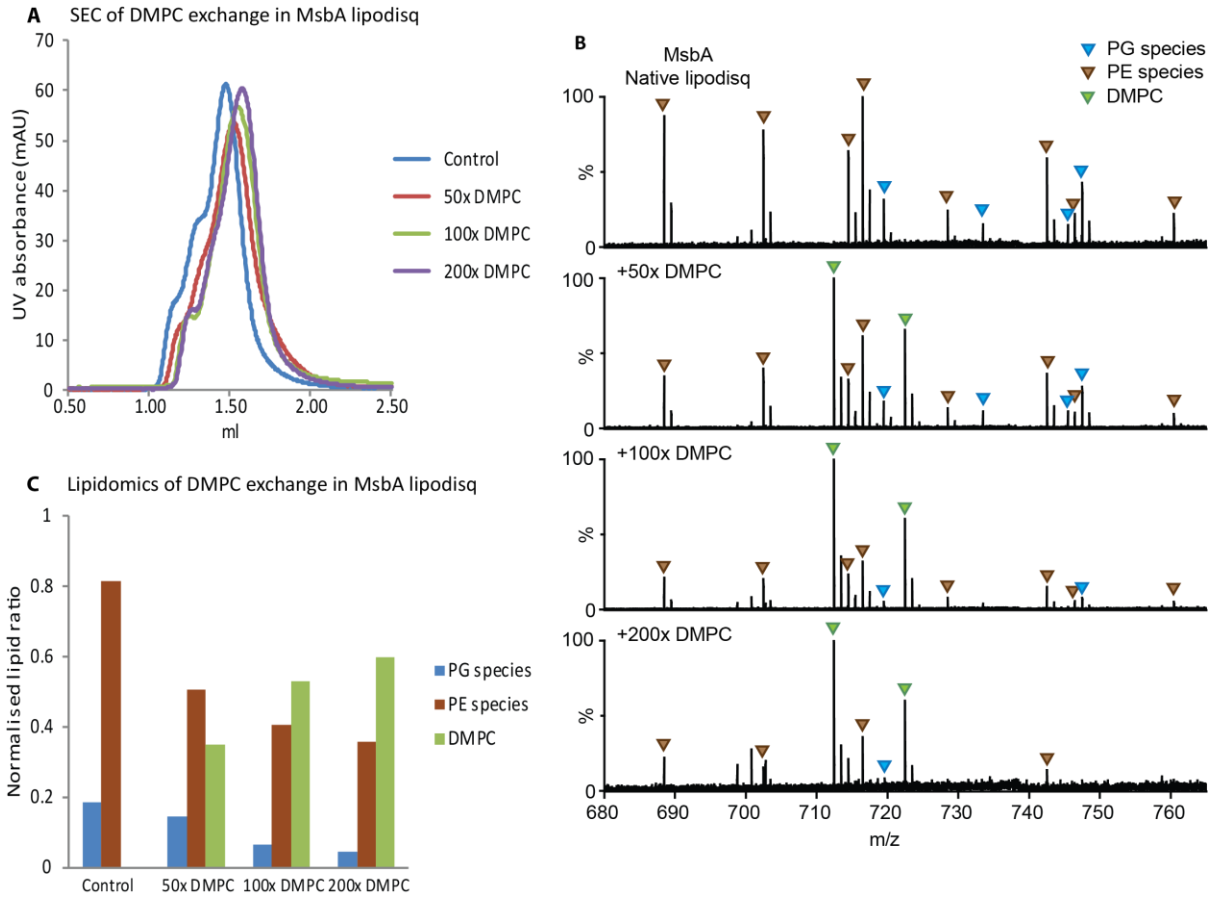


Figure 3-7 Exogenous DMPC exchanged with the phospholipid in the lipodisq of MsbA. A, the combined SEC chromatograms of DMPC exchanging with the lipids in MsbA lipodisq. MsbA lipodisq was added into different concentrations of DMPC (50, 100 and 200 times the concentration of MsbA lipodisq) individually and incubated at RT with 500 RPM for 2 hours. The control was the same batch of MsbA lipodisq diluted with SEC buffer and treated together with the samples of DMPC exchange. After SEC, the lipodisq samples subjected to LC-MS/MS based lipidomic identification and quantification following Section 2.7. **B,** the LC-MS spectra of identified phospholipid species in a series of DMPC exchanged MsbA lipodisq and compared with that in MsbA native lipodisqs. Lipidomics was performed following the procedure described in Section 2.7. **C,** the normalised ratios of PGs, PEs and DMPC in a series of DMPC exchanged MsbA lipodisqs and compared with that in MsbA native lipodisqs.

Lipidomic analysis also revealed the lipid transfer crossing the SMA co-polymer scaffold of lipodisq, reported previously by researchers[12,24-26]. MsbA native lipodisq was mixed with exogenous phospholipid DMPC (PC(14:0/14:0)) in three different ratios (1:50, 1:100 and 1:200) respectively and incubated for a further 2 hours. The signal of DMPC in lipidomic spectra appeared in the MsbA lipodisq after lipid exchange and the intensity of DMPC became dominant stepwise with the increase of DMPC concentration for lipid exchange (Figure 3-7 B). The normalised lipidomic quantification confirmed that the percentage of DMPC in lipodisq increased with the addition of DMPC for exchange, and no preferential replacement of DMPC on specific lipid species was observed (Figure 3-7 C). With more DMPC transferred into lipodisq, the chromatograph of the SEC purification after lipid exchange showed an elution time right shift, implying a smaller size of lipdisq formed in the presence of DMPC. It indicated that the structural characterisation of lipodisq could be affected by the composition of packed lipids.

Furthermore, several researches demonstrate that membrane proteins and SMA co-polymer form lipodisqs that can be used for crystallography and other biophysical methods[27,28]. These methods have shown that lipodisqs are highly dynamic assemblies that exchange the components with other particles even phases, rather than being static complexes.

3.3 Use of native MS to study membrane proteins in lipodisq

3.3.1 Rhodopsin in DMPC lipodisq

To understand the lipid-protein interaction in lipodisqs, we applied native MS to study membrane protein rhodopsin in DMPC lipodisq. The three types of rhodopsin,

bacteriorhodopsin, archaerhodopsin-3 and archaerhodopsin-4, expressed in *Halobacterium salinarum,* were purified using detergent methods and reconstituted in DMPC liposome[29,30]. All three rhodopsins were well-resolved in the native MS spectra, and the *apo* proteins and lipid bound species were clearly assigned (Figure 3-8).

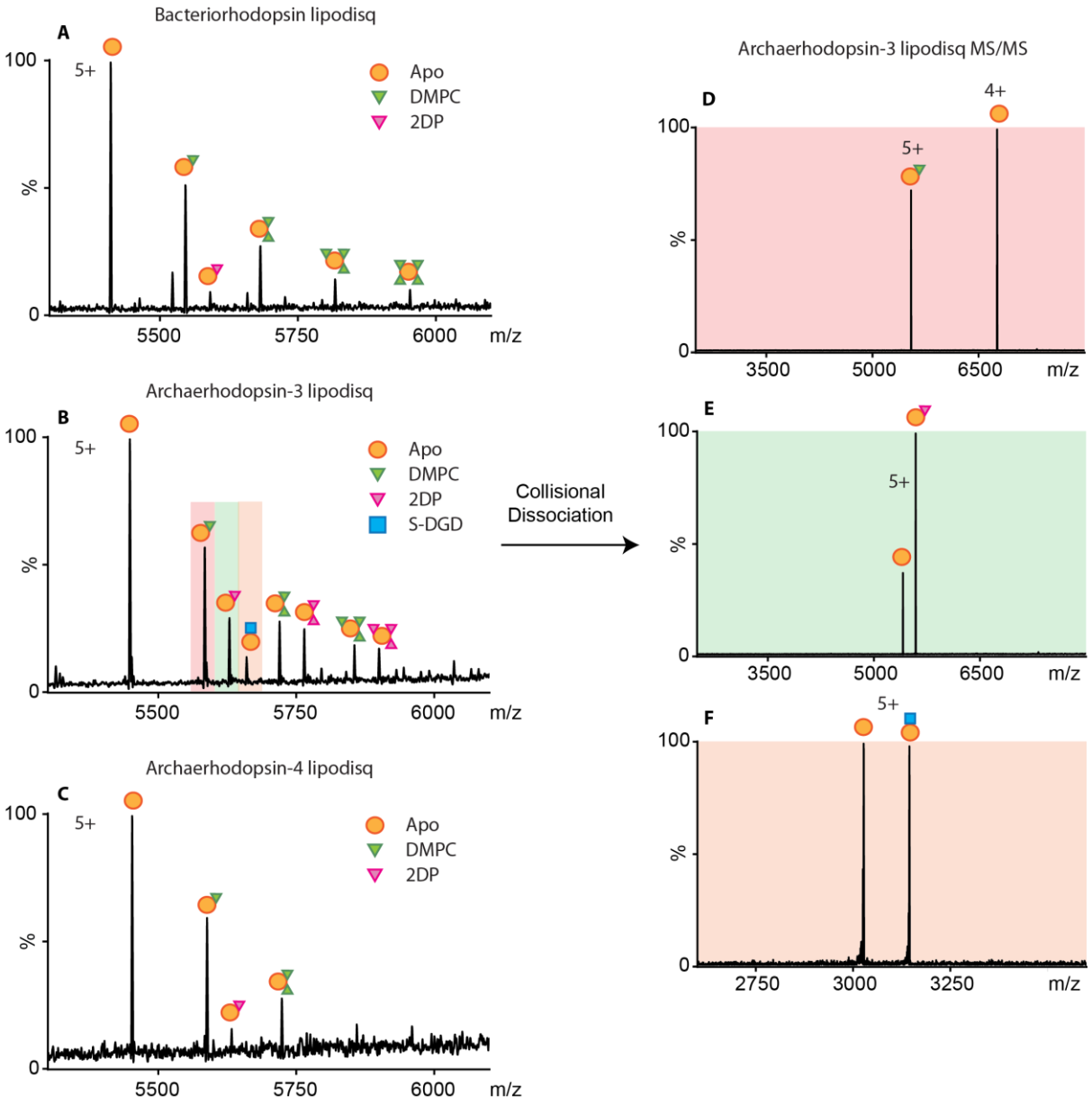


Figure 3-8 The native MS spectra of microbial rhodopsins in lipodisqs. A, B and **C,** the native MS spectra of bacteriorhodopsin (**A**), archaerhodopsin-3 (**B**) and archaerhodopsin-4 (**C**) in DMPC lipodisq. Typical lipid bound species highlighted in **B** were dissociated by extended voltage as described in the Methods, and the MS/MS spectra were recorded in **D, E** and **F,** respectively. The *apo* form of rhodopsin and the binding adducts were assigned and labelled (orange filled circle, rhodopsin; green triangle, DMPC; pink triangle, 2DP; blue square, S-DGD).

As the majority component of lipodisq, more than one DMPC molecules were to bind with rhodopsins (Figure 3-8 A, B and C). Additionally, two other types of lipids, 2DP and S-DGD, were found to interact with rhodopsin, identified by the collisional dissociation in the MS/MS spectra in Figure 3-8 E and F. The reported 2D crystal structure of bacteriorhodopsin confirmed that 2DP located close to its membrane-spanning helices[31], and the glycolipid S-DGD was observed in the lipid annulus of bacteriorhodopsin[32]. However, currently no research has demonstrated the specific interaction between these two lipids and rhodopsin. We assumed that 2DP and S-DGD were non-specific bound lipids remaining from purification.

The native MS spectra of lipodisqs were quite different to those of detergent micelles. Comparing the spectra of bacteriorhodopsin in lipodisqs and proteomicelles, several forms of bacteriorhodopsin were observed in the detergent environment, but not detected in the lipodisq sample (Figure 3-9 A and B). As illustrated in Figure 3-9, most of these forms were immature precursors and may not be present in the membrane. For archaerhodopsin-3, despite exogenous DMPC, both 2DP and S-DGD binding were identified in the spectrum of lipodisq in Figure 3-9 C, but only S-DGD binding was found in rhodopsin in detergent (Figure 3-9 D). Since the lipodisq was from the direct insertion of SMA co-polymer into the membrane, a comparison with the mimic

detergent micelle shows the lipodisq is a more reliable method to reveal the actual state of embedded membrane protein as well as interactions between lipid and protein.

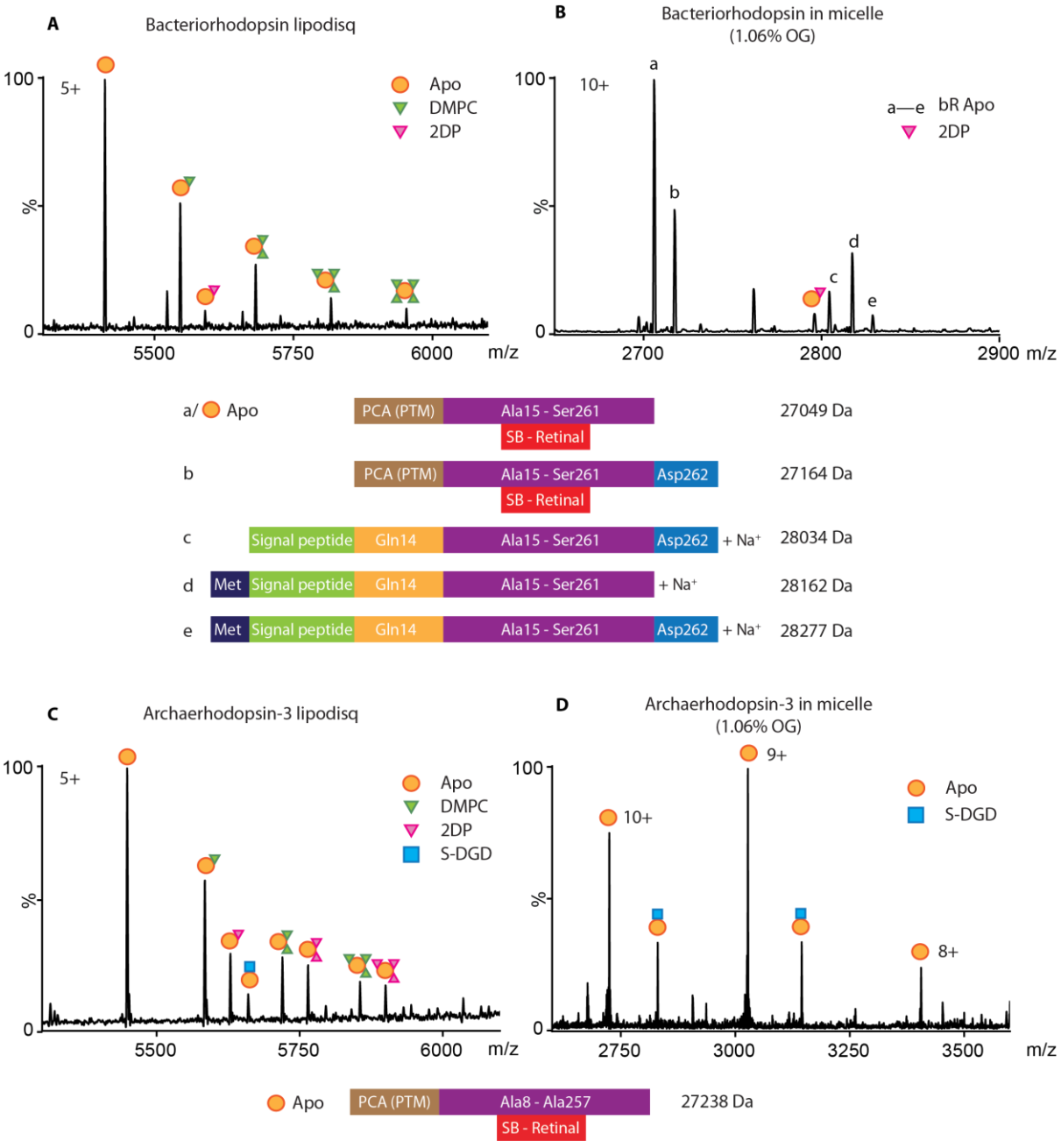


Figure 3-9 The native MS spectra of rhodopsins in DMPC lipodisq and OG micelle. A and **B,** the comparison of native MS spectra of bacteriorhodopsin in lipodisq and micelle. Different forms of *apo* bacteriorhodopsin detected in spectra were identified and shown in the schematic below the spectra. **C** and **D,** the comparison of native MS spectra of archaerhodopsin-3 in lipodisqs and micelles. The *apo* form of rhodopsin and the binding adducts were assigned and labelled (orange filled circle, rhodopsin; green triangle, DMPC; pink triangle, 2DP; blue square, S-DGD).

3.3.2 MdtK in native lipodisq

In contrast to rhodopsin above, native MS was for the in-house native lipodisq of MdtK and a well resolved spectrum was obtained (Figure 3-10 A). The assigned mass of MdtK lipodisq was 102.68 kDa, more than two times greater than that of delipidated MdtK in the DDM micelle (Figure 3-10 B, 50.55 kDa). The most intense charge state was isolated to identify the species observed in the spectrum of lipodisq (Figure 3-10 C) and dissociated by high collisional energy. The species observed in Figure 3-10 A was composed of two monomers with a mass consistent with that of MdtK, and two bound small molecules with a mass around 787 Da (Figure 3-10 D). According to the analysis of MdtK in progressive delipidation in the previous chapter, this bonding molecule could be a PG lipid.

Therefore, similar to the lipid regulating the stoichiometry of LeuT and NhaA[33], we presumed that MdtK exists as a PG mediated dimer in the native environment, which has not yet been reported. In native lipodisq, this character was preserved by the direct solubilisation of SMA co-polymer. In contrast, in the protocol we used for membrane solubilisation in the previous chapter, the introduction of detergent may interrupt the interaction between MdtK and PG, leading to the dissociation of dimeric MdtK. Since no dimer signal was detected from the native MS spectra of exogenous

PG binding with MdtK as shown in Figure 2-11, we proposed that the PG binding with monomeric MdtK cannot induce MdtK dimerisation in a DDM detergent environment.

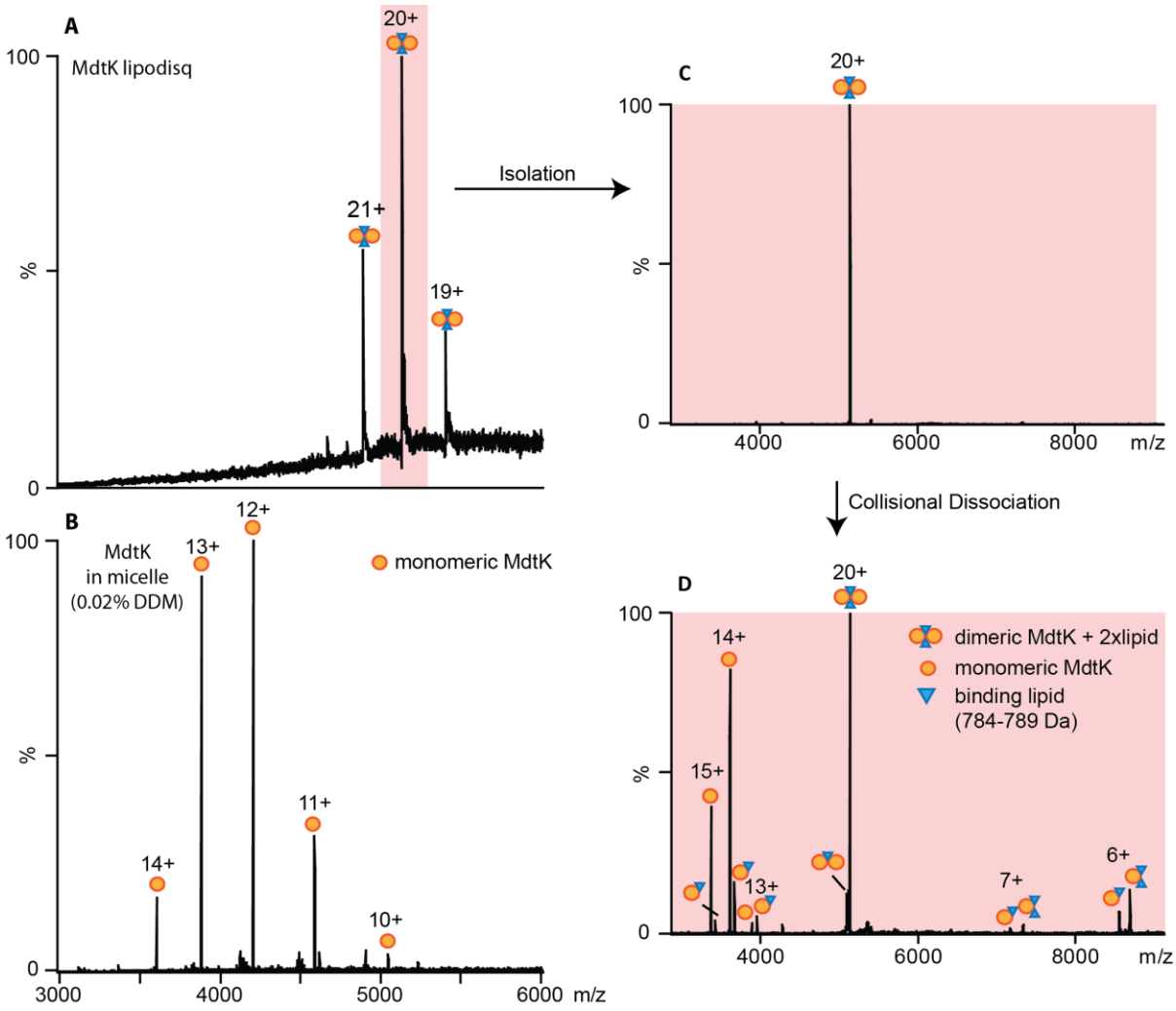


Figure 3-10 The native MS spectra of MdtK in native lipodisq and DDM micelle. A and **B,** comparison of native MS spectra of MdtK in lipodisqs and micelles. The *apo* form of monomeric MdtK (orange filled circle) in 0.02% DDM was assigned in **B**. The signal of lipid bound species highlighted in **A** was isolated and dissociated by extended voltage as described in the Methods, and the spectra were recorded in **C** and **D** respectively**.** The two lipid bound dimeric MdtK was identified by the MS/MS spectra in **D.**

To confirm the role of PG in the oligomerisation of MdtK, we investigated the stoichiometry of MdtK in a PG-deficient environment. We produced a mutant *E. coli* UE54 strain lacking genes for PG synthesis[34], which also affects the downstream

formation of CL[35,36]. As the T7 driven plasmid pET-28a cannot work in the K-12 strain UE54, the protein gene was transferred into a pBAD vector for expression. We successfully harvested MdtK from UE54. The lipidomic analysis confirmed that no PG was present in the cell membrane (Figure 3-11 C, top panel), and the minor component PA enriched to maintain polar localisation of anionic lipids in the cell membrane instead of PG and CL.[37]

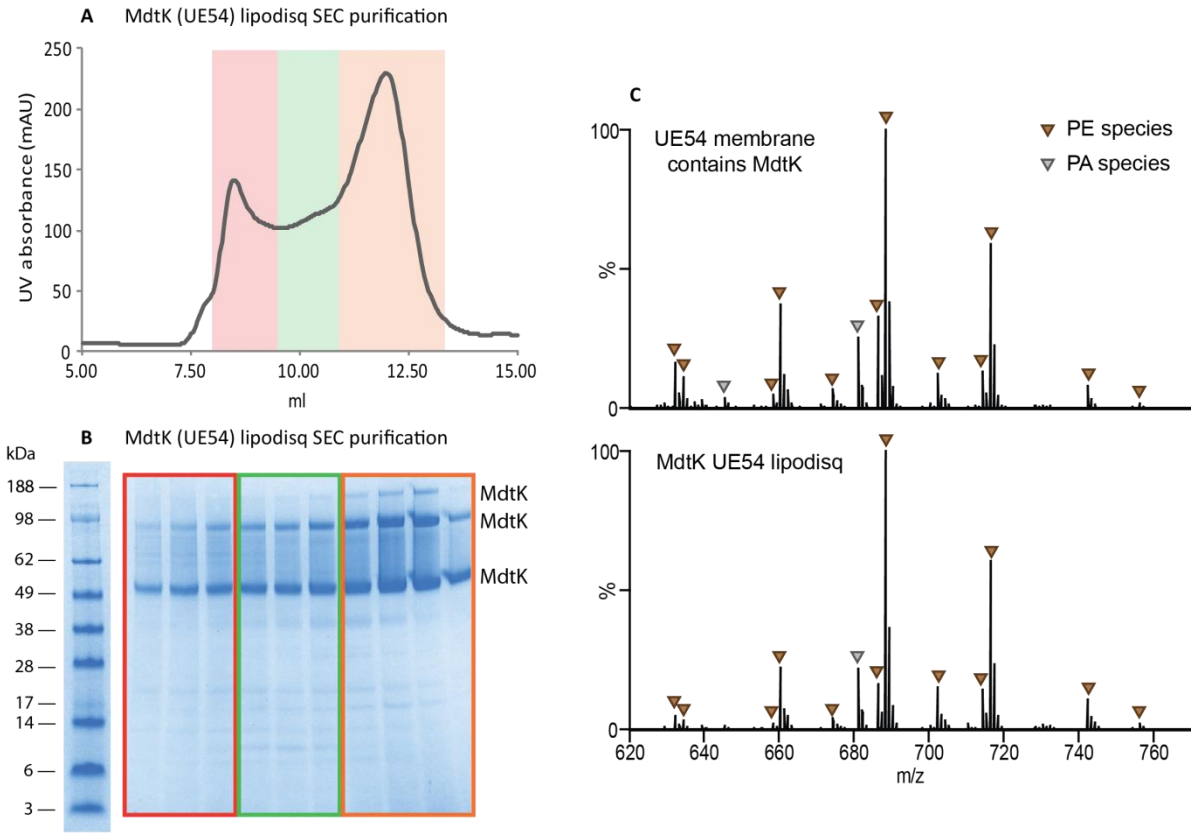


Figure 3-11 The purification of MdtK PG-deficient lipodisq. A and **B,** the SEC purification of MdtK lipodisq prepared from UE54 strain. The fractions containing different signal peaks in chromatogram (**A**) were highlighted and analysed in SDS-PAGE (**B**), labelled with the same colour as in **A.** The indicated bands in **B** were subjected to LC-MS/MS based proteomic identification. **C,** the LC-MS spectra of phospholipid species in MdtK PG-deficient lipodisq and membrane of *E. coli* (UE54) cells overexpressing MdtK. Lipidomics was performed following the procedure described in Section 2.7.

Compared with the native membrane, the PG-deficient membrane was far more difficult to solubilise by the SMA co-polymer. In the C43 strain, the membrane became clear just several minutes after the start of incubation. However, the PG-deficient membrane remained cloudy until the incubation ended, and more insoluble residues remained through the subsequent centrifugal filtration. However, we still obtained lipodisq containing MdtK from the UE54 strain (confirmed by SDS-PAGE in Figure 3-11 C), with divergent particle sizes similar to that in the C43 strain (Figure 3-11 A). The formed PG-deficient lipodisq displayed an identical lipid composition as the original membrane in Figure 3-11 C.

As with the solubilisation process, the disruption and removal of the bulk lipids in lipodisqs using native MS were hard to achieve. Currently, we have not obtained a well-resolved native MS spectrum, and we are therefore adjusting conditions and parameters. When we are able to confirm that MdtK exists in a PG-deficient environment as a monomer using native MS, it will be possible to examine the selectivity and specificity of PG species in mediation of the oligomerisation of MdtK.

3.4 Conclusion

We developed a method of using SMA co-polymer to solubilise membrane proteins directly from the expressing cell membrane in this project, achieving native lipodisq formation and purification on a range of important membrane proteins, including transporters and GPCRs. Importantly, we also proved that this novel detergent-free particle is a potential candidate for native MS study. As displayed in the native MS spectra above, in some cases, rhodopsin and MdtK for instance, membrane proteins behave differently in lipodisqs and micelles. Compared with a mimic environment in

detergent micelles, native lipodisqs preserve the original state of the membrane protein and recent lipid cohort[6,7,18]. By using native MS and relevant biophysical techniques, we are able to gain a better understanding of the lipid effect on the conformation and function of the membrane protein *in vivo* from inspecting native lipodisqs.

However, the properties of SMA co-polymer limit the application of lipodisq for membrane protein study[13,18]. Currently, a series of SMA derivatives have been developed and subjected to protein solubilisation[10]. In contrast to conventional SMA, these functional group modifications provide a wider pH range and polyvalent cation tolerance[18,19,38], which expands the use of the lipodisq method in membrane protein. The possibility of studying these new lipodisqs via native MS is also being assessed.

3.5 Methods

Native lipodisq of membrane protein preparation and purification. SMA (2:1) co-polymer was hydrolysed from styrene/maleic anhydride polymer (67:33, SMA2000, Polysciences) and stored according to published protocols[9,13]. For membrane solubilisation, SMA (2:1) co-polymer was dissolved in SMA buffer (50 mM Tris, pH8.0) to make 10% solution. The membrane fraction was diluted to roughly 8.5 mg/µl with solubilisation buffer (50 mM Tris, 500 mM NaCl and 10% glycerol, pH 8.0). The two solutions were mixed in a volume ratio of 1:1 and incubated at 37 °C with rotation. The mixture was filtered through 0.22 µm filter paper to remove unsolubilised residues. Lipodisqs of membrane protein were purified by IMAC and SEC chromatography in FPLC at 4 °C. 0.05 times of total volume (V) cobalt agarose beads were prewashed with IMAC buffer A (50 mM Tris, 250 mM NaCl and 5% glycerol, pH

8.0) and added into the mixed solution, incubating overnight with rotation. Then the mixture was loaded into an empty gravity flow column, washing with 3V of IMAC buffer A, and eluted with 0.2V IMAC buffer B (50 mM Tris, 500 mM NaCl, 500 mM imidazole and 10% glycerol, pH 8.0). The collected sample was concentrated at 5000 g in a concentrator (100 kDa) and filtered for purification in an SEC column (Superdex 200 increase gel-filtration column) with SEC buffer (50 mM Tris and 150 mM NaCl, pH 8.0). Purified lipodisqs of membrane protein were stored at 4 °C.

MdtK-GFP expression in PG-deficient *E. coli* strain UE54. The cells of UE54 strain (Host: *E. coli* K-12 MG 1655, Genotype: *lpo-5508 Δara-714 rcsF:: miniTn10cat ΔpgsA:: kan*, supplied by Matsuoka lab, Saitama University) was made into calcium competent cells according to the established manual from the Krantz lab[39]. MdtK-GFP gene was inserted into a pBAD22 vector from a pET-28a vector following the protocol of the commercial polymerase chain reaction (PCR) kit (Phusion® Hot Start Flex DNA Polymerase)[40], and transferred in UE54 competent cells for expression. A tip of UE54 cell colonies was mini-cultured in 5 mL of (LB) supplemented with 0.05 mg/ml kanamycin sulphate at 37 °C with 180 RPM for 12 h, and 0.5 ml of cells was added into 100 ml LB supplemented with 0.05 mg/ml kanamycin sulfate to preculture overnight at 37°C with 180 RPM. Then 5 ml of precultured cells grew in 1 L LB supplemented with 0.05 mg/ml kanamycin sulfate at 37 °C with 180 RPM for 4 h (OD (D600) around 0.3 finally). Protein expression was induced by addition of 0.2% L-arabinose, and incubation continued for another 3 h at 37 °C. The cell membrane was harvested and stored following methods in Section 2.7.

Native mass spectrometry. For MS analysis, lipodisqs of membrane proteins after SEC purification were buffer exchanged into MS spray buffer using the biospin column (Zeba^{TM} Spin Desalting Columns, molecular weight exclusion limit 40 kDa). Spray buffer was 200 mM ammonium acetate (pH 8.0). The sample was introduced into a modified Q-Exactive mass spectrometer (Thermo) and ions were transferred into the HCD cell following a gentle voltage gradient as detailed in the description Section 2.7. The spray voltage was maintained at 1.4 kV at a temperature of 200 °C, and the optimised acceleration voltage was 100 V for HCD cell[41]. The collisional dissociation was performed with the voltage ramp from 0-150 V in HCD cells with 100 V for in-source trapping and 50 V for source fragmentation[42]. Backing pressure was maintained around 1.00×10^{-9} mbar. Data were analysed using Xcalibur 2.2 SP1.48.

References

(1) Quick, M.; Shi, L.; Zehnpfennig, B.; Weinstein, H.; Javitch, J. A. Experimental conditions can obscure the second high-affinity site in LeuT. *Nat Struct Mol Biol* **2012**, *19*, 207-211.

(2) Zhou, H. X.; Cross, T. A. Modeling the membrane environment has implications for membrane protein structure and function: Influenza A M2 protein. *Protein Science* **2013**, *22*, 381-394.

(3) Dorr, J. M.; Scheidelaar, S.; Koorengevel, M. C.; Dominguez, J. J.; Schafer, M.; van Walree, C. A.; Killian, J. A. The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J* **2016**, *45*, 3-21.

(4) Tonge, S. R.; Tighe, B. J. Responsive hydrophobically associating polymers: a review of structure and properties. *Adv Drug Deliv Rev* **2001**, *53*, 109-122.

(5) Jamshad, M.; Grimard, V.; Idini, I.; Knowles, T. J.; Dowle, M. R.; Schofield, N.; Sridhar, P.; Lin, Y.; Finka, R.; Wheatley, M.; Thomas, O. R. T.; Palmer, R. E.; Overduin, M.; Govaerts, C.; Ruysschaert, J.-M.; Edler, K. J.; Dafforn, T. R. Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins. *Nano Research* **2014**, *8*, 774-789.

(6) Scheidelaar, S.; Koorengevel, M. C.; Pardo, J. D.; Meeldijk, J. D.; Breukink, E.; Killian, J. A. Molecular model for the solubilization of membranes into nanodisks by styrene maleic Acid copolymers. *Biophys J* **2015**, *108*, 279-290.

(7) Dorr, J. M.; Koorengevel, M. C.; Schafer, M.; Prokofyev, A. V.; Scheidelaar, S.; van der Cruijsen, E. A.; Dafforn, T. R.; Baldus, M.; Killian, J. A. Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K+ channel: the power of native nanodiscs. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 18607-18612.

(8) Jamshad, M.; Lin, Y. P.; Knowles, T. J.; Parslow, R. A.; Harris, C.; Wheatley, M.; Poyner, D. R.; Bill, R. M.; Thomas, O. R.; Overduin, M.; Dafforn, T. R. Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochemical Society transactions* **2011**, *39*, 813-818.

(9) Rothnie, A. J. Detergent-Free Membrane Protein Purification. *Methods Mol Biol* **2016**, *1432*, 261-267.

(10) Stroud, Z.; Hall, S. C. L.; Dafforn, T. R. Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. *Methods* **2018**, *147*, 106-117.

(11) Scheidelaar, S.; Koorengevel, M. C.; van Walree, C. A.; Dominguez, J. J.; Dorr, J. M.; Killian, J. A. Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers. *Biophys J* **2016**, *111*, 1974-1986.

(12) Grethen, A.; Glueck, D.; Keller, S. Role of Coulombic Repulsion in Collisional Lipid Transfer Among SMA(2:1)-Bounded Nanodiscs. *J Membr Biol* **2018**, *251*, 443-451.

(13) Lee, S. C.; Knowles, T. J.; Postis, V. L.; Jamshad, M.; Parslow, R. A.; Lin, Y. P.; Goldman, A.; Sridhar, P.; Overduin, M.; Muench, S. P.; Dafforn, T. R. A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature protocols* **2016**, *11*, 1149-1162.

(14) Zhang, R.; Sahu, I. D.; Liu, L.; Osatuke, A.; Comer, R. G.; Dabney-Smith, C.; Lorigan, G. A. Characterizing the structure of lipodisq nanoparticles for membrane protein spectroscopic studies. *Biochimica et biophysica acta* **2015**, *1848*, 329-333.

(15) Abramson, J.; Riistama, S.; Larsson, G.; Jasaitis, A.; Svensson-Ek, M.; Laakkonen, L.; Puustinen, A.; Iwata, S.; Wikstrom, M. The structure of the ubiquinol

oxidase from Escherichia coli and its ubiquinone binding site. *Nat Struct Biol* **2000**, *7*, 910-917.

(16) Chepuri, V.; Gennis, R. B. The Use of Gene Fusions to Determine the Topology of All of the Subunits of the Cytochrome-O Terminal Oxidase Complex of Escherichia-Coli. *Journal of Biological Chemistry* **1990**, *265*, 12978-12986.

(17) Pollock, N. L.; Lee, S. C.; Patel, J. H.; Gulamhussein, A. A.; Rothnie, A. J. Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer. *Biochim Biophys Acta Biomembr* **2018**, *1860*, 809-817.

(18) Fiori, M. C.; Jiang, Y.; Altenberg, G. A.; Liang, H. Polymer-encased nanodiscs with improved buffer compatibility. *Sci Rep* **2017**, *7*, 7432.

(19) Oluwole, A. O.; Danielczak, B.; Meister, A.; Babalola, J. O.; Vargas, C.; Keller, S. Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer. *Angewandte Chemie* **2017**, *56*, 1919-1924.

(20) Lin, Y.-p. Over-expression and Biophysical Characterisation of Membrane Proteins Solubilised in a Styrene Maleic Acid Polymer. University of Birmingham, 2011.

(21) Stoscheck, C. M. quantitation of protein. *Methods Enzymol* **1990**, *182*, 50-69.

(22) Layne, E. Spectrophotometric and Turbidimetric Methods for Measuring Proteins. *Method Enzymol* **1957**, *3*, 447-454.

(23) Instructions: Pierce 660nm Protein Assay. https://www.thermofisher.com/order/catalog/product/22660.

(24) Cuevas Arenas, R.; Danielczak, B.; Martel, A.; Porcar, L.; Breyton, C.; Ebel, C.; Keller, S. Fast Collisional Lipid Transfer Among Polymer-Bounded Nanodiscs. *Sci Rep* **2017**, *7*, 45875.

(25) Cuevas Arenas, R.; Klingler, J.; Vargas, C.; Keller, S. Influence of lipid bilayer properties on nanodisc formation mediated by styrene/maleic acid copolymers. *Nanoscale* **2016**, *8*, 15016-15026.

(26) Hazell, G.; Arnold, T.; Barker, R. D.; Clifton, L. A.; Steinke, N. J.; Tognoloni, C.; Edler, K. J. Evidence of Lipid Exchange in Styrene Maleic Acid Lipid Particle (SMALP) Nanodisc Systems. *Langmuir : the ACS journal of surfaces and colloids* **2016**, *32*, 11845-11853.

(27) Broecker, J.; Eger, B. T.; Ernst, O. P. Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs. *Structure* **2017**, *25*, 384-392.

(28) Schmidt, V.; Sturgis, J. N. Modifying styrene-maleic acid co-polymer for studying lipid nanodiscs. *Biochim Biophys Acta Biomembr* **2018**, *1860*, 777-783.

(29) Cherry, R. J.; Muller, U. Temperature-dependent aggregation of bacteriorhodopsin in dipalmitoyl- and dimyristoylphosphatidylcholine vesicles. *J Mol Biol* **1978**, *121*, 283-298.

(30) Ding, X.; Sun, C.; Cui, H.; Chen, S.; Gao, Y.; Yang, Y.; Wang, J.; He, X.; Iuga, D.; Tian, F.; Watts, A.; Zhao, X. Functional roles of tyrosine 185 during the bacteriorhodopsin photocycle as revealed by in situ spectroscopic studies. *Biochim Biophys Acta Bioenerg* **2018**.

(31) Mitsuoka, K.; Hirai, T.; Murata, K.; Miyazawa, A.; Kidera, A.; Kimura, Y.; Fujiyoshi, Y. The structure of bacteriorhodopsin at 3.0 A resolution based on electron crystallography: implication of the charge distribution. *J Mol Biol* **1999**, *286*, 861-882.

(32) Lobasso, S.; Lopalco, P.; Angelini, R.; Pollice, A.; Laera, G.; Milano, F.; Agostiano, A.; Corcelli, A. Isolation of Squarebop I bacteriorhodopsin from biomass of coastal salterns. *Protein Expr Purif* **2012**, *84*, 73-79.

(33) Gupta, K.; Donlan, J. A.; Hopper, J. T.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **2017**, *541*, 421-424.

(34) Shiba, Y.; Yokoyama, Y.; Aono, Y.; Kiuchi, T.; Kusaka, J.; Matsumoto, K.; Hara, H. Activation of the Rcs signal transduction system is responsible for the thermosensitive growth defect of an Escherichia coli mutant lacking phosphatidylglycerol and cardiolipin. *J Bacteriol* **2004**, *186*, 6526-6535.

(35) Tan, B. K.; Bogdanov, M.; Zhao, J.; Dowhan, W.; Raetz, C. R.; Guan, Z. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 16504-16509.

(36) Oliver, P. M.; Crooks, J. A.; Leidl, M.; Yoon, E. J.; Saghatelian, A.; Weibel, D. B. Localization of anionic phospholipids in Escherichia coli cells. *J Bacteriol* **2014**, *196*, 3386-3398.

(37) Mileykovskaya, E.; Ryan, A. C.; Mo, X.; Lin, C. C.; Khalaf, K. I.; Dowhan, W.; Garrett, T. A. Phosphatidic acid and N-acylphosphatidylethanolamine form membrane domains in Escherichia coli mutant lacking cardiolipin and phosphatidylglycerol. *The Journal of biological chemistry* **2009**, *284*, 2990-3000.

(38) Ravula, T.; Hardin, N. Z.; Ramadugu, S. K.; Ramamoorthy, A. pH Tunable and Divalent Metal Ion Tolerant Polymer Lipid Nanodiscs. *Langmuir : the ACS journal of surfaces and colloids* **2017**, *33*, 10655-10662.

(39) Making calcium competent cells. http://mcb.berkeley.edu/labs/krantz/protocols/calcium_comp_cells.pdf.

(40) PCR protocol for Phusion® Hot Start Flex DNA Polymerase (M0535). https://international.neb.com/protocols/2012/09/06/protocol-for-phusion-hot-start-flex-dna-polymerase-m0535.

(41) Yen, H. Y.; Hoi, K. K.; Liko, I.; Hedger, G.; Horrell, M. R.; Song, W.; Wu, D.; Heine, P.; Warne, T.; Lee, Y.; Carpenter, B.; Pluckthun, A.; Tate, C. G.; Sansom, M. S. P.; Robinson, C. V. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **2018**, *559*, 423-427.

(42) Gong, H.; Li, J.; Xu, A.; Tang, Y.; Ji, W.; Gao, R.; Wang, S.; Yu, L.; Tian, C.; Li, J.; Yen, H. Y.; Man Lam, S.; Shui, G.; Yang, X.; Sun, Y.; Li, X.; Jia, M.; Yang, C.; Jiang, B.; Lou, Z.; Robinson, C. V.; Wong, L. L.; Guddat, L. W.; Sun, F.; Wang, Q.; Rao, Z. An electron transfer path connects subunits of a mycobacterial respiratory supercomplex. *Science* **2018**, *362*.

4 Using mass spectrometry study a mycobacterial respiratory supercomplex

4.1 Introduction

Respiration is one of the most vital biological processes in almost all life forms. During respiration, electrons are transferred from electron donors to acceptors via an electron transport chain. ATP is synthesised by the proton gradient and provides energy for most cellular reactions[1]. In the respiratory chain, multiple membrane oxidoreductases are assembled into supercomplexes to regulate the process.

The target prokaryotic respiratory supercomplex from *mycobacterium smegmatis (M. smegmatis)* comprises two oxidoreductases: *bcc*-type complex III (cytochrome *c* domain fused bc_3-type ubiquinol: cytochrome *c* oxidoreductase) and aa_3-type complex IV (cytochrome *c* oxidase). The Cryo-EM study reported in H. Gong *et al*. revealed that the supercomplex CIII-CIV is in a linear dimeric arrangement of CIV_1-$CIII_2$-CIV_1, where the central CIII homodimer is flanked by a CIV monomer at each side (Figure 4-1 A)[2]. As shown in Figure 4-1 B, CIII complex is composed of QcrA, QcrB and QcrC three subunits with 3, 8 and 1 transmembrane helices, respectively. CIV complex consists of six subunits, CtaC, CtaD, CtaE, CtaF, CtaI and CtaJ, and the last two are new components identified in *M. smegmatis*. There are 2 transmembrane helices from CtaC, 12 from CtaD, 5 from CtaE, 4 from CtaF and 1 from CtaJ. Besides, two novel subunits LpqE and PRSAF1 are assigned within the interface between CIII and CIV (Figure 4-1 A), together with numerous phospholipids to mediate their interactions (Figure 4-1 B). Additionally, a superoxide dismutase subunit SodC is observed at the top of the CIII complex dimer, at the periplasmic face

(Figure 4-1 A), protecting the respiratory system against reactive oxygen species formed by CIII complex (Figure 4-2)[2].

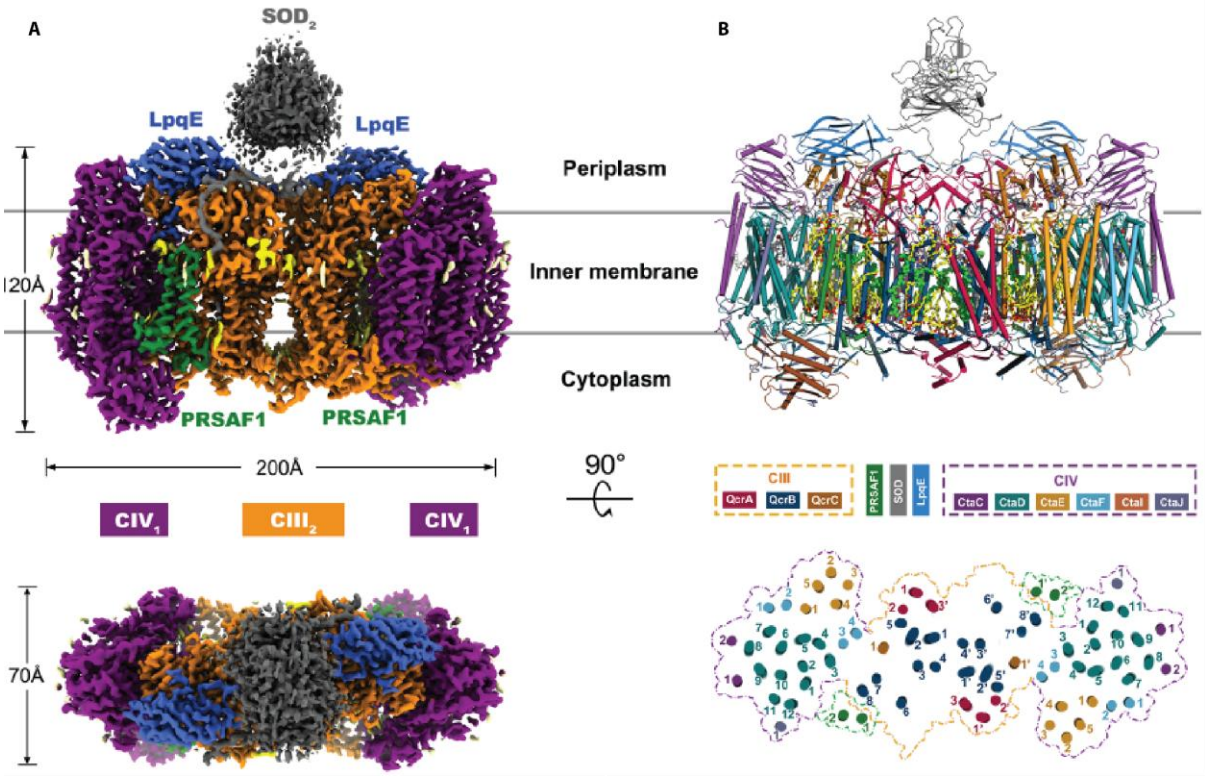


Figure 4-1 The overall structure of the *bcc-aa$_3$* type respiratory CIII-CIV supercomplex from *Mycobacterium smegmatis* (PDB ID: 6ADQ). A, the cryo-EM map of the supercomplex shows a linear twofold dimerised form of CIV_1–$CIII_2$–CIV_1 from the side view (top) and the top view (bottom).The complex CIII, CIV and protein LpqE, PRSAF1, SOD are coloured respectively (orange, CIII; purple, CIV; blue, LpqE; green, PRSAF1; gray, SOD). **B**, Representation of the supercomplex from the side view (top) and the cross-section view (bottom). The subunits are labelled with the colour displayed (middle). The binding MKs and phospholipids in the side view are presented as bright green coloured solid spheres and yellow sticks. The boundary of the supercomplex is depicted with dashed lines in the corresponding colour used in **A.** Figures are from H. Gong *et al.* [2].

The cryo-EM study also demonstrated the full electron transfer pathway within the respiratory supercomplex (Figure 4-2). Starting from the electron donor menaquinol (MKH_2) binding in CIII, one electron is transferred via the [2Fe-2S] cluster, heme c_{D2}, heme c_{D1}, Cu_A, heme *a* in turn and arrives at the terminal heme a_3:Cu_B reaction

centre in CIV for oxygen reduction into water. The other electron from the same donor is passed to heme b_L then heme b_H and finally to the menaquinone (MK) for hydrogenation to complete the Q-cycle in the same subunit. The whole process of electron translocation between prosthetic groups is buried inside the integral complex without cytochrome *c* shuttle and conformational transition[2].

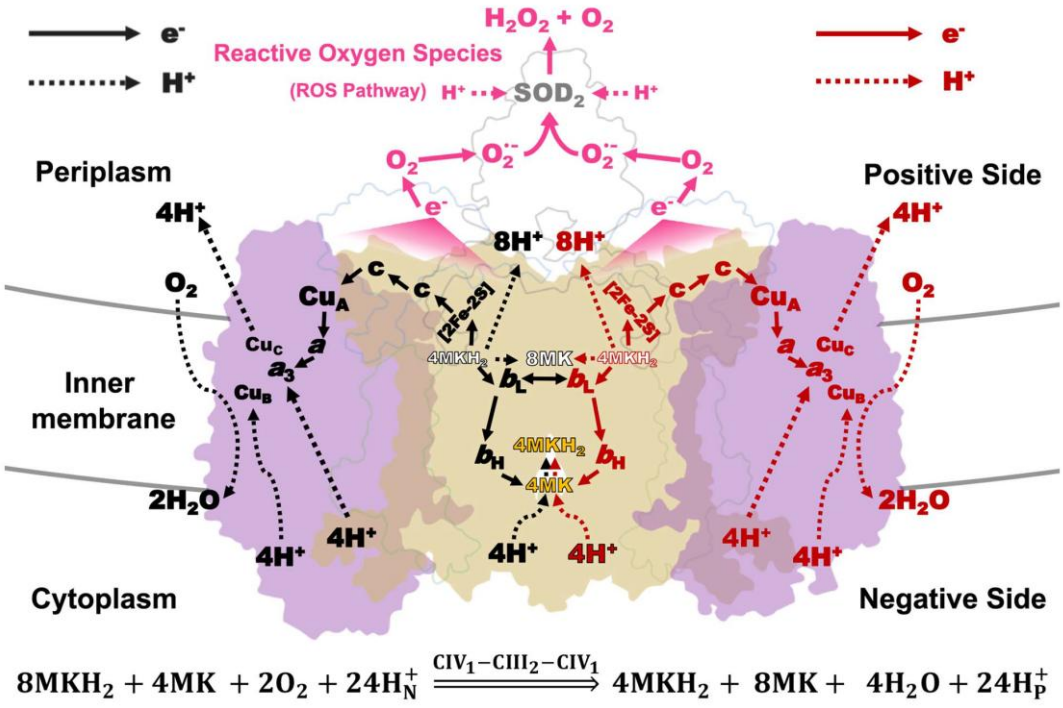


Figure 4-2 Schematic diagram of the electron transfer pathway in the *bcc-aa₃* type respiratory CIII/CIV supercomplex from *Mycobacterium smegmatis*. The electrons donated from MKH_2 are transferred from CIII to CIV for oxygen reduction and MK via heme *b* groups to accomplish Q cycle in CIII complex. The entire respiratory reaction is attached. The diagram also displays the scavenging mechanism of superoxide dismutase complex on reactive oxygen species. Figure is from H. Gong *et al.*[2].

However, some details of the supercomplex components cannot be observed directly using the cryo-EM method. For instance, the types of bioactive lipids inside the supercomplex are deduced from the density map of Cryo-EM, since their length and the degree of unsaturation are still unknown. For a better understanding of the

electron translocation in the *bcc-aa*$_3$ type respiratory supercomplex of mycobacteria, this chapter looks at the application of MS in the investigation of the supercomplex. To verify the identity of each component, we firstly use MS to determine the exact molecular weight of the protein subunits and the bound lipid molecules in the supercomplex. Secondly, native MS is also used to inspect the protein-protein and protein-lipid interactions, and these are then compared with those observed within cryo-EM. The analysis of component identity and complex structure using MS is expected to assist in the further study of the respiratory supercomplex of mycobacteria.

4.2 The identification of supercomplex CIII-CIV subunits

By optimising parameters in the native mass spectrometer, we successfully obtained the fully-resolved spectrum (Figure 4-3 A) of the complex with a determined mass of 873.44 kDa, which is much larger than the value of structural model (*c.a.* 720 kDa)[2]. According to the SDS-PAGE hydrolysis and in-gel proteomics in Figure 4-3 B, the identification of the known subunits confirmed the high purity of the supercomplex in the sample for native MS, proving that the assigned species with the very large mass is the intact supercomplex CIII-CIV. It is worth noting that there are additional protein bands in gel absent in the current model. Therefore we proposed that besides the contributions of the co-purified lipid species and detergents surrounding the supercomplex, it is possible that there are unidentified subunits presenting in the supercomplex system.

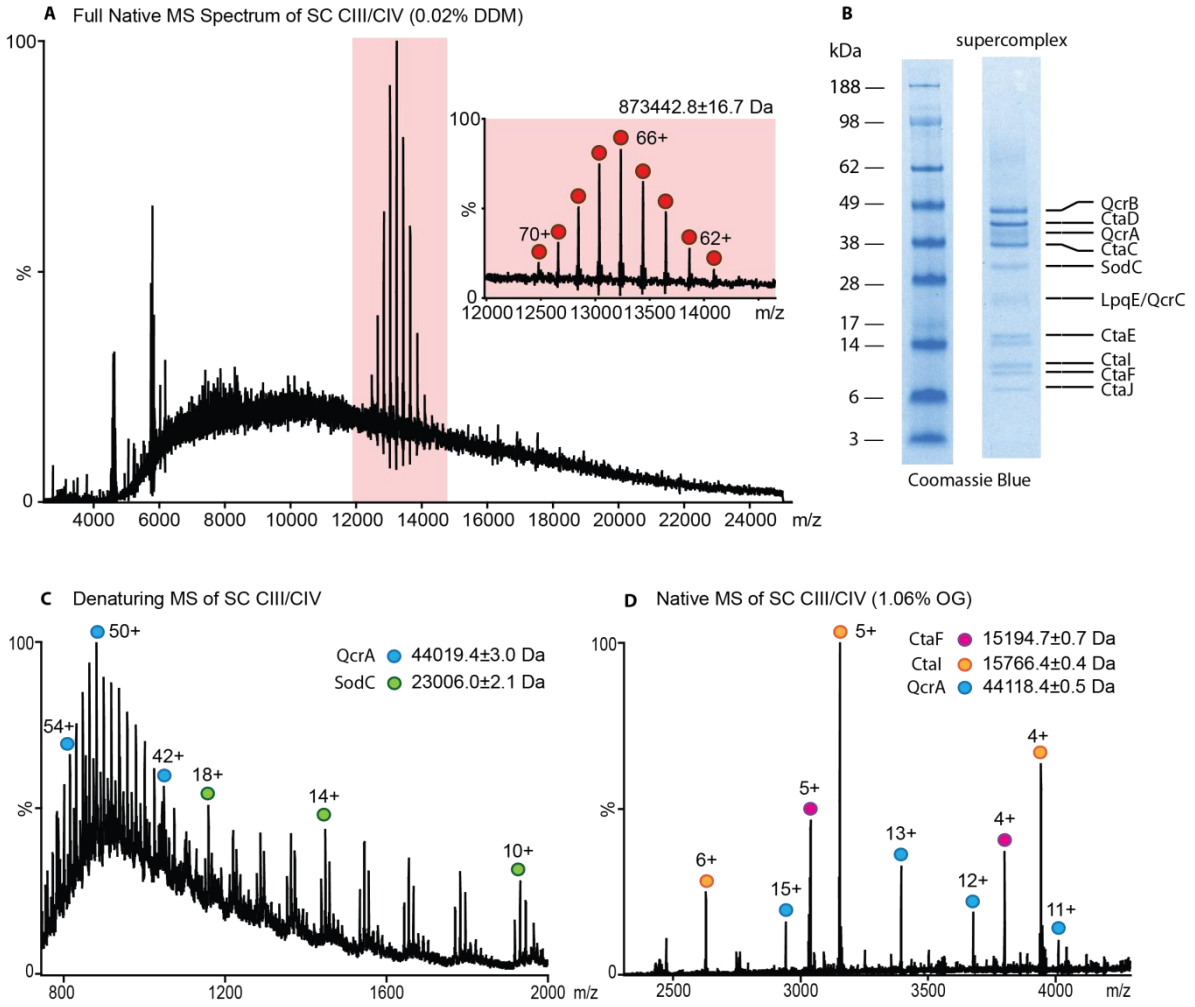


Figure 4-3 The identification of the supercomplex CIII/CIV by mass spectrometry. A, the full native MS spectrum of the intact $CIII_2CIV_2SOD_2$ supercomplex in MS buffer containing 0.02% DDM. The intact supercomplex was assigned in the amplified inset spectra. **B,** SDS-PAGE analysis of purified supercomplex from SEC. The bands of labelled subunits were subjected for LC-MS/MS based proteomics identification. **C** and **D,** the representative spectra of the subunits identification via two different approaches: denaturing LC-MS (**C**) and native MS (**D**). Signals of identified subunits were labelled and assigned in the spectra.

To interrogate the supercomplex components, their stoichiometry and interactions, we attempted to isolate and dissociate the charge state of the intact supercomplex. However, the dissociation was so difficult that we did not record the species with smaller mass in the MS/MS spectrum even with the application of the highest possible voltage of the HCD cell. Therefore, we introduced a novel denaturing LC-MS

method for the determination of subunit molecular weight. Without preinjection acidification or resuspension, the native-state supercomplex was subjected to the LC-MS, and the integration was denatured into smaller subunits and peptide fragments. Different components were separated using a chromatographic column and analysed by the subsequent MS, similar to the reported rapid LC-MS method for membrane protein accurate molecular weight determination[3]. As shown in the representative spectrum of Figure 4-3 C, only QcrA and SodC were well-resolved in this spectrum due to the different retention time of subunits in the column. In the region of 800—1000 m/z, QcrA was observed to possess adductive peaks of additional 510 Da, which was confirmed as DDM detergent. In the region of over 1200 m/z, we detected the presence of SodC with multiple levels of glycosylation, which will be discussed in the next section. Using the denaturing LC-MS method, we successfully identified ten of the twelve subunits of the supercomplex with an error margin of less than 5 Da (Table 4-1).

In the meantime, the native MS condition was adjusted for the acquisition of the signal from subunits. A relatively low sample concentration was used as the assembled subunits tended to disperse in low concentration. For complex dissociation, higher acceleration voltages were applied in the instrument. Like the procedure in progressive delipidation, a detergent screening by buffer exchange was performed to select an appropriate detergent. We obtained native MS spectra from several detergents, and identified eight subunits of the supercomplex in spectra from five different detergents (Table 4-1). For instance, in the spectrum recorded from the supercomplex in 2×CMC OG (Figure 4-3 D), CtaF, CtaI and QcrA were clearly resolved and assigned, and CtaI was also observed in the spectra of C_8E_4 and LDAO.

	Molecular weight (kDa)			
Protein Name	**Theoretical Value**	**Built in Model**	**Denaturing LC-MS**	**Native MS**
QcrA	46.33	42.16	44.01	44.11
QcrB	60.21	59.05		59.25
QcrC	30.82	23.16	24.59	24.72
PRSAF1	11.32	10.45	11.04	
CtaC	38.04	35.17	36.24	
CtaD	64.11	61.69		
CtaE	22.18	22.18	22.09	22.09
CtaF	15.17	15.17	15.19	15.20
CtaI	15.90	14.80	15.77	15.78
CtaJ	8.35	7.18	8.38	
SodC	23.21	21.20	23.00	23.00
LpqE	19.11	15.25	16.60	16.61
Unknown Subunit			9.39	
				39.96
				24.17

Table 4-1 Identification of the supercomplex CIII/CIV subunits by mass spectrometry. The theoretical and modelling molecular weights of subunits were calculated based on the published sequence and structure of the supercomplex (PDB ID: 6ADQ)[2]. The masses of subunits were acquired in denaturing LC-MS and native MS following the protocols described in Section 4.6. The mass of SodC determined by MS was the lowest mass of glycosylated SodC assigned in the denaturing and native MS spectra respectively.

The molecular weights of the supercomplex subunits determined by denaturing LC-MS and native MS are listed in Table 4-1. QcrB and CtaD were not observed in

the LC-MS spectra. These two subunits have the largest masses and number of transmembrane helices (12 for QcrB, 8 for CtaD) among the twelve identified subunits. Therefore, we supposed that the current protocol was not ideal for the elution of these proteins with high hydrophobicity. Some chromatography conditions could be optimised for hydrophobic proteins, such as extending the gradient elution time, switching to the column with higher resolution, and increasing the eluent strength. Relatively high column temperature and low buffer pH could be examined for the improvement of complex dissociation. It should be noted that these condition adjustments may affect the separation and identification of other more hydrophilic proteins. Additionally, unlike the signal of plastic contamination, the subunits of the supercomplex eluted with limited efficiency. Obvious carryover was observed in the subsequent runs after protein injection, which was reported previously[4,5]. Compared to denaturing LC-MS, more subunits were not detected in native MS spectra (PRSAF1, CtaC, CtaD and CtaJ). With the exception of QcrB and LpqE, each of the remaining six subunits was confirmed in the spectra from at least two types of detergent. Meanwhile, ligand bound species, subunit associations and stoichiometries were also recorded through native MS, which will be discussed in the later sections.

As shown in Table 4-1, there were seven subunits (QcrA, QcrC, CtaE, CtaF, CtaI, SodC and LpqE) identified by both denaturing and native MS, and the mass differences between the two methods were quite small. Considering the presence of non-covalent adducts in the sample, in QcrA and QcrC, the mass obtained in native MS was marginally greater than that from denaturing LC-MS. In addition to the known subunits of the supercomplex, we also identified several species with unmatched

mass by the two methods. We cannot confirm that these species were subunits of the supercomplex as this is not supported by other methods or structural information.

Next, the molecular weights of the supercomplex subunits determined by MS were compared with the corresponding values in the theoretical and structural model. In the case of CtaE, CtaF and SodC, the masses from the sequence, modelling and the two MS methods were almost identical, demonstrating the feasibility and accuracy of molecular weight identification by denaturing LC-MS and native MS. For CtaI, CtaJ and PRSAF1, the result using MS was similar to the theoretical value, indicating that the flexible residues were overlapped by the moieties of vicinal subunits in modelling. Conversely, the experimental mass of QcrB is closer to the modelling value, showing that the invisible sequences in the crystal structure might be truncated in PTM. However, in QcrA, QcrC, CtaC and LpqE, their experimental molecular weights were between the theoretical values and those of models. Our calculation revealed that the experimental masses of these subunits were consistent with the theoretical molecular weight of unmodelled, truncated residues at the sequence terminus of each subunits: QcrA, 20 residues off at N-terminal; QcrC, 57 residues off at N-terminal; CtaC, 14 residues off at N-terminal and 3 residues off at C-terminal; LpqE, 23 residues off at N-terminal and 2 residues off at C-terminal.

In summary, despite the effect from the unidentified subunit CtaD, the actual molecular weight of the supercomplex was still much smaller than the detected mass of the whole complex using native MS. Further optimisation of denaturing LC-MS and

native MS could be used, potentially together with other biochemical assays to investigate the potential additional subunits.

4.3 Lipidomic analysis of supercomplex CIII-CIV

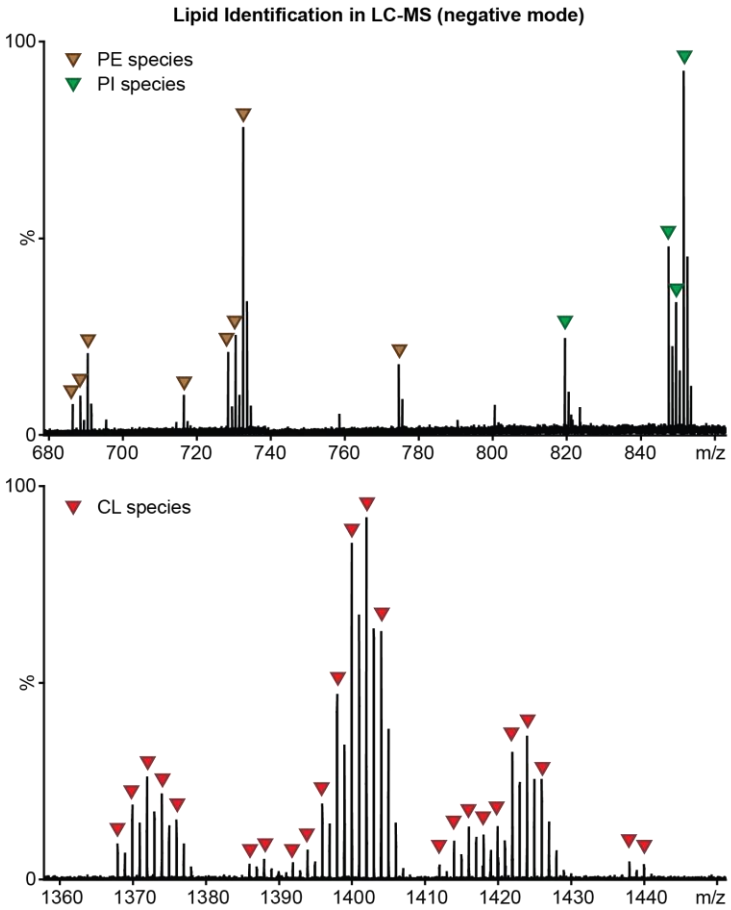
4.3.1 The lipid composition of the supercomplex

Lipidomic analysis was performed to examine the co-purified lipids of the supercomplex. 8 PE species, 4 phosphatidylinositol (PI) species and 26 CL species were extracted from the respiratory supercomplex CIII-CIV, with no other class of phospholipids detected by LC-MS (Figure 4-4). Independent lipidomic analysis confirmed that the majority of lipids in the purified supercomplex were PE, PI and CL[2]; the dominant species in each type of lipid corresponds to the intensive peaks in the spectra we obtained in Figure 4-4.

As mentioned above, phospholipids, especially CL are known to play an important role in the assembling and stabilisation of supercomplexes[6]. According to the structure built in H. Gong *et al.*, 4 PE, 8 PI and 18 CL are identified to attach the transmembrane space of the dimeric supercomplex[2]. Among these lipids, two PE molecules are located in the cavity formed by two monomeric CIII complexes, and four CL molecules are observed to fill the large space between CIII and CIV, which may contribute to the stability and assembly of the supercomplex (Figure 4-1 B).

Additionally, some PE species were detected in the denaturing LC-MS spectra of the supercomplex with a low resolution and intensity, which was also reported in the rapid LC-MS method[3]. As the denaturing method was developed for molecular weight

determination of the membrane protein, the lipid information from the denaturing method was not utilised for the lipidomic analysis of the supercomplex.



PE(38:0)	774.60
PE(35:0)	732.55
PE(35:1)	730.54
PE(35:2)	728.52
PE(34:1)	716.52
PE(32:0)	690.51
PE(32:1)	688.49
PE(32:2)	686.47

PI(35:0)	851.56
PI(35:1)	849.55
PI(35:2)	847.53
PI(33:2)	819.50

CL(76:14)	1491.94
CL(76:15)	1489.93
CL(72:12)	1439.99
CL(72:13)	1437.97
CL(70:5)	1425.97
CL(70:6)	1423.95
CL(70:7)	1421.94
CL(70:8)	1419.93
CL(70:9)	1417.91
CL(70:10)	1415.99
CL(70:11)	1413.97
CL(70:12)	1411.96
CL(68:2)	1403.99
CL(68:3)	1401.97
CL(68:4)	1399.96
CL(68:5)	1397.94
CL(68:6)	1395.93
CL(68:7)	1393.91
CL(68:8)	1391.89
CL(68:10)	1387.96
CL(68:11)	1385.94
CL(66:2)	1375.96
CL(66:3)	1373.94
CL(66:4)	1371.93
CL(66:5)	1369.91
CL(66:6)	1367.90

Figure 4-4 The lipidomic analysis of the supercomplex CIII/CIV. The LC-MS spectra of phospholipid and CL species extracted from the supercomplex. PE, PI and CL species are labelled with brown, green and red triangles respectively. The lipid species extracted from the supercomplex present in the spectra are listed on the right hand side. Lipidomics was performed following the procedure described in Section 2.7.

4.3.2 Menaquinone identification

As the key electron carrier in respiratory electron transport systems, menaquinone (MK) and its derivatives from the respiratory supercomplex were identified in LC-MS in positive ion mode. Five variants of MK were detected in the MS spectra, and four of them were partially saturated species (Figure 4-5 A, C and D). MK-9(H_2) was the

predominant molecule, and only traces of the unhydrogenated species MK-9 was observed in the supercomplex (aligned with pure MK-9 in Figure 4-5 B). This result is supported by previous research, demonstrating that MK is the only electron carrier in the electron transport system of mycobacteria, and the majority species in mycobacteria is MK-9(II-H_2)[7,8]. The MK-9(II-H_2) molecule possesses a side chain of nine isoprene units with one double bond hydrogenated at the β-position unit[9].

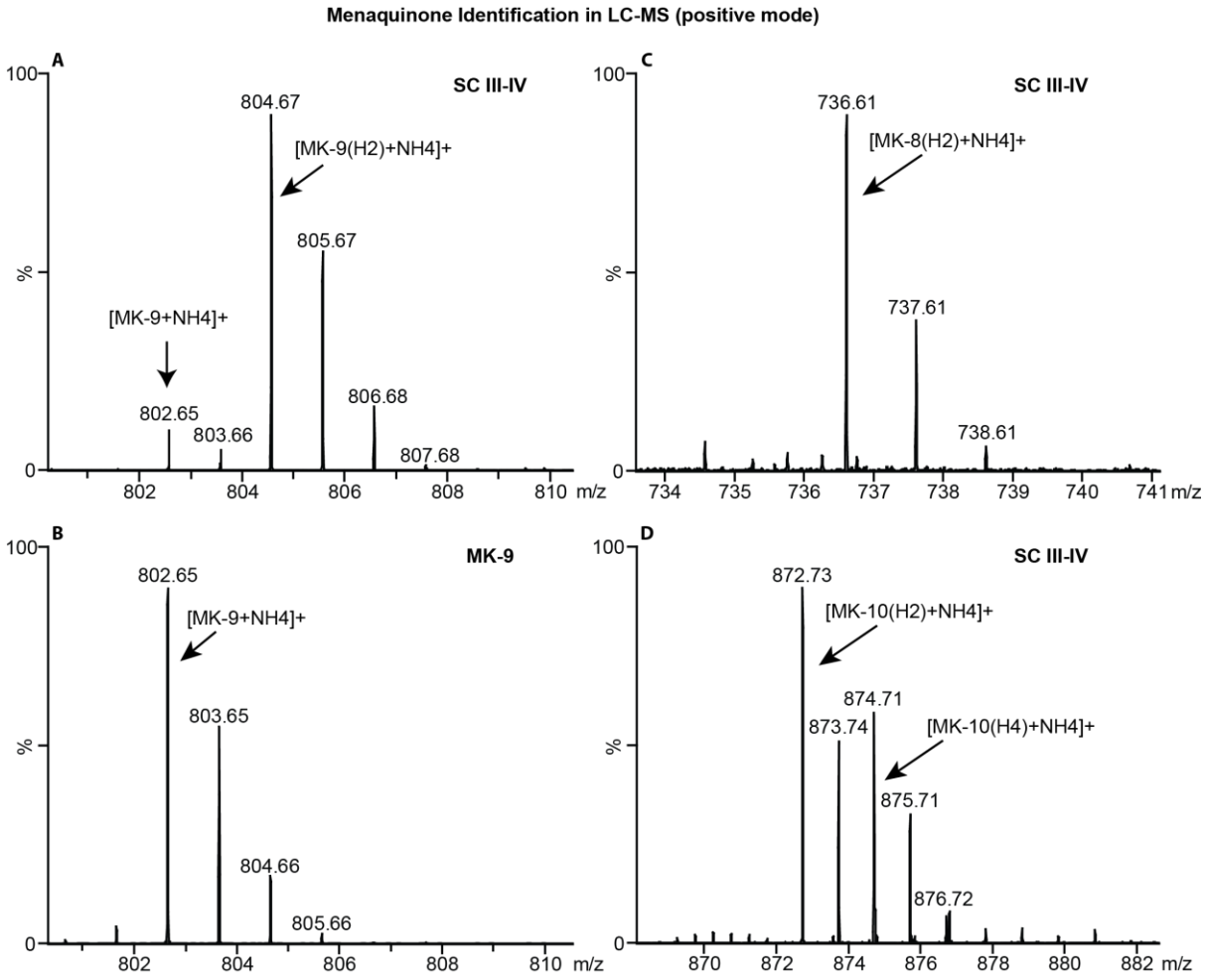


Figure 4-5 The identification of variants of MKs in supercomplex CIII/CIV. A, C and **D,** the LC-MS spectra of various MK species extracted from the supercomplex. **B,** the LC-MS spectrum of commercial pure MK-9 supplied by collaborators. The MK's identification was performed in positive ion mode following the lipidomic analysis protocol in Section 2.7.

The discussion in the previous section revealed that the binding MK molecules are involved in the electron translocation in the respiratory supercomplex. In the structure of dimeric CIII complex displayed in Figure 4-8 D, the reductive menaquinol (MKH_2) bound site on QcrB, supplies electrons for oxygen reduction in CIV and MK reduction at Q_N in CIII[2]. Based on the research from our collaborator, the edge-to-edge distances from MKH_2 to downstream heme b_L and [2Fe-2S] cluster, and MK to upstream heme b_L are 16, 13 and 5 Å respectively, which are short enough for the electron passing.

It was worth noting that the amount of partially saturated MK variants was much more than that of unsaturated ones in the supercomplex. Previous research has reported that the efficiency of electron transport in mycobacteria is increased by the reduction reaction in the biosynthesis process via reductase MenJ, converting MK to MKH_2, which provides the endogenous electron donor for the respiratory supercomplex[10]. In summary, this mechanism regulates the composition of the MK pool in mycobacteria, maintains the respiratory state, and additionally, affects the bacterial survival[10,11].

4.4 Use of native MS to investigate the components of supercomplex CIII-CIV

4.4.1 The periplasmic superoxide dismutase SodC in the supercomplex

Despite the molecular weight of subunits from the supercomplex, native MS was able to provide information on the subunits and their interactions in the system. In Figure 4-6 A, native MS identified a glycosylated dimeric protein component SodC, which was also observed on the periplasmic side of the CIII complex dimer in the EM map (Figure 4-1). Collisional dissociation was performed on the most intense peak in

Figure 4-6 A for the stoichiometry of SodC, and the dissociated signal in the MS/MS spectrum under high energy confirmed the presence of dimeric SodC in the supercomplex (Figure 4-6 B, bottom). Moreover, the native MS spectrum of dimeric SodC clearly displayed the extensive glycosylation and the copper-binding of SodC in the supercomplex. In the inset spectrum in Figure 4-6 A, the neighbouring major peaks of charge state 13+ differed by 162 Da, which corresponded to a hexose unit. Variable levels of glycosylation were observed, with the protein cofactor copper ion attached to SodC in each glycosylated state. The pattern of monomeric SodC (5+ charge state) revealed three major and two minor SodC glycoforms existed in the supercomplex, which was also identified in the SodC spectrum from the denaturing LC-MS (Figure 4-3 C). According to the previous research on SodC using *Mycobacteria tuberculsis (M. tuberculsis),* SodC was *O*-glycosylated with mannose on Thr and Ser residues in the region near the N-terminal[12]. Therefore, considering the sequence similarity of SodC in the two mycobacteria in Figure 4-6 C, it is proposed that the SodC of the supercomplex from *M. smegmatis* is modified with mannose at the N-terminal portion, and the four potential specific glycosylation sites were highlighted in the 3D density map[2].

The Cu/Zn containing superoxide dismutase is a key enzyme for the survival of bacteria, providing local protection against oxy-radical damage by endogenous and/or exogenous reactive oxygen species[12-14]. As an integral part of this respiratory supercomplex, the assembled SodC was reported to preserve superoxide dismutase activity[2], similar to the other mycobacterial superoxide dismutases associated with the membrane and respiratory supercomplex[13,15]. According to the previous study on *M. tuberculsis* SodC, its enzymatic activity is not affected by the N-terminal

O-glycosylation[12]. Therefore, we assumed that the catalytic activity of SodC may not be directly influenced by the state of glycosylation.

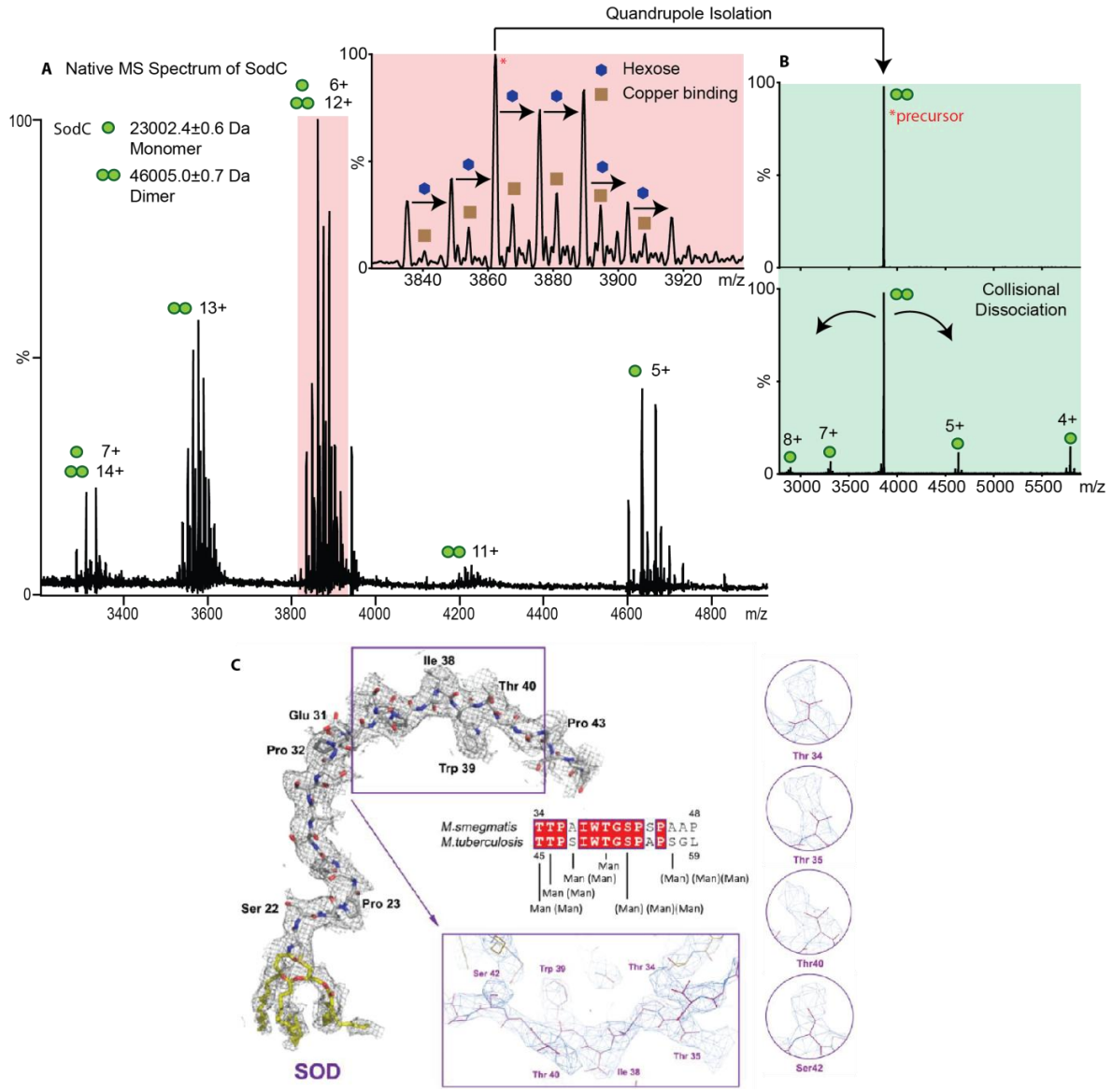


Figure 4-6 The glycosylation and dimerisation of superoxide dismutase SodC in the supercomplex. A, the native MS spectrum of SodC in MS buffer containing 0.02% DDM. The glycosylation and copper-binding were identified in the amplified insect spectrum. The signal of dimeric SodC labelled in the insect spectrum in **A** was isolated and dissociated by extended voltage as described in the methods, and the spectra were recorded in **B. C,** the cryo-EM density of monomeric SodC from published research[2]. A sequence alignment of the N-Terminal clustering of the O-glycosylation sites in the SodC from *M. smegmatis* and *M. tuberculsis.* Based on the previous research on SodC[12], the specific glycosylation sites were identified and amplified in violet circles.

4.4.2 The supercomplex subunits association

As discussed in the previous section, using the native MS, we observed that the functional subunit SodC was present in the supercomplex as an assembled dimer (Figure 4-7 A). Meanwhile, other associated subunits were also identified in the native MS spectra. In Figure 4-7 B and C, we obtained the spectra of the intact CIV complex and dimeric CIII complex. The mass of CIII dimer identified by native MS matches the sum of the molecular weight of QcrA, QcrB and QcrC subunits. For CIV, although the molecular weight of CtaD subunit was not confirmed by MS, we estimated the mass of the CIV complex with the theoretical value of CtaD and obtained the molecular weight of other subunits, which is consistent with the mass captured as integral in native MS. In contrast to SodC, the signal of the two assembled complexes was relatively weak, and no dissociated subunit was recorded in the spectrum.

With reference to the revealed structure of the supercomplex (Figure 4-7 D), the CIV complexes were located on each side of the central CIII dimer in a symmetrised linear arrangement. In the supercomplex, the assembly and stability of the associated architecture were contributed from phospholipids located around the TMD of CIII and CIV and the interface between CIII and CIV[6]. As revealed in the cryo-EM study, few lipids were found to fill the space between the intrinsic subunits of CIII and CIV, indicating a compact integration inside the two complexes[1]. In contrast, SodC linked to the TMD of CIII complex via its lipid modified N-terminal fragment, allowing a flexible orientation for the subunit in the periplasm (Figure 4-7 D). Therefore, compared with the associated CIII and CIV complexes, the soluble SodC assembly is much easier to be identified and dissociated with native MS.

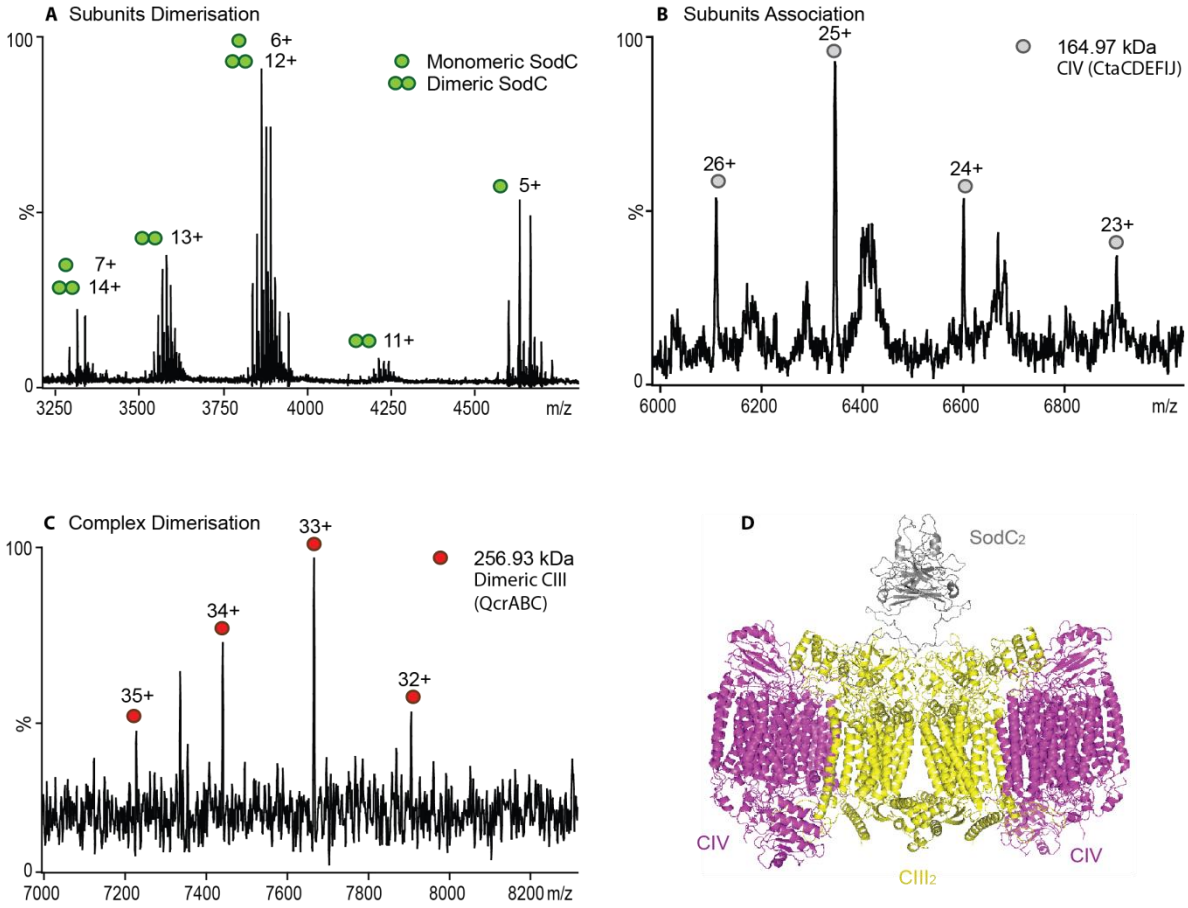


Figure 4-7 The association of supercomplex subunits. A, the native MS spectrum of monomeric and dimeric SodC in MS buffer containing 0.02% DDM. **B,** the native MS spectrum of CIV complex in MS buffer containing 0.007% DMNG. **C,** the native MS spectrum of dimeric CIII complex in MS buffer containing 0.02% DDM. **D,** the arrangement of CIII, CIV complex and associated SodC in supercomplex. Protein LpqE and PRSAF1 were not shown.

4.4.3 The prosthetic groups in supercomplex

Native MS also provided information of the bound prosthetic groups. In Figure 4-8 A, CIV complex was identified with an adduct with a mass over 1.65 kDa. Considering the mass of the bound prosthetic groups and phospholipids of CIV (Figure 4-8 C), the adduct binding species should be a di-heme *a* CIV complex. Similarly, CIII subunit QcrB bound with two heme *b* was also identified in the native MS spectrum in Figure 4-8 B.

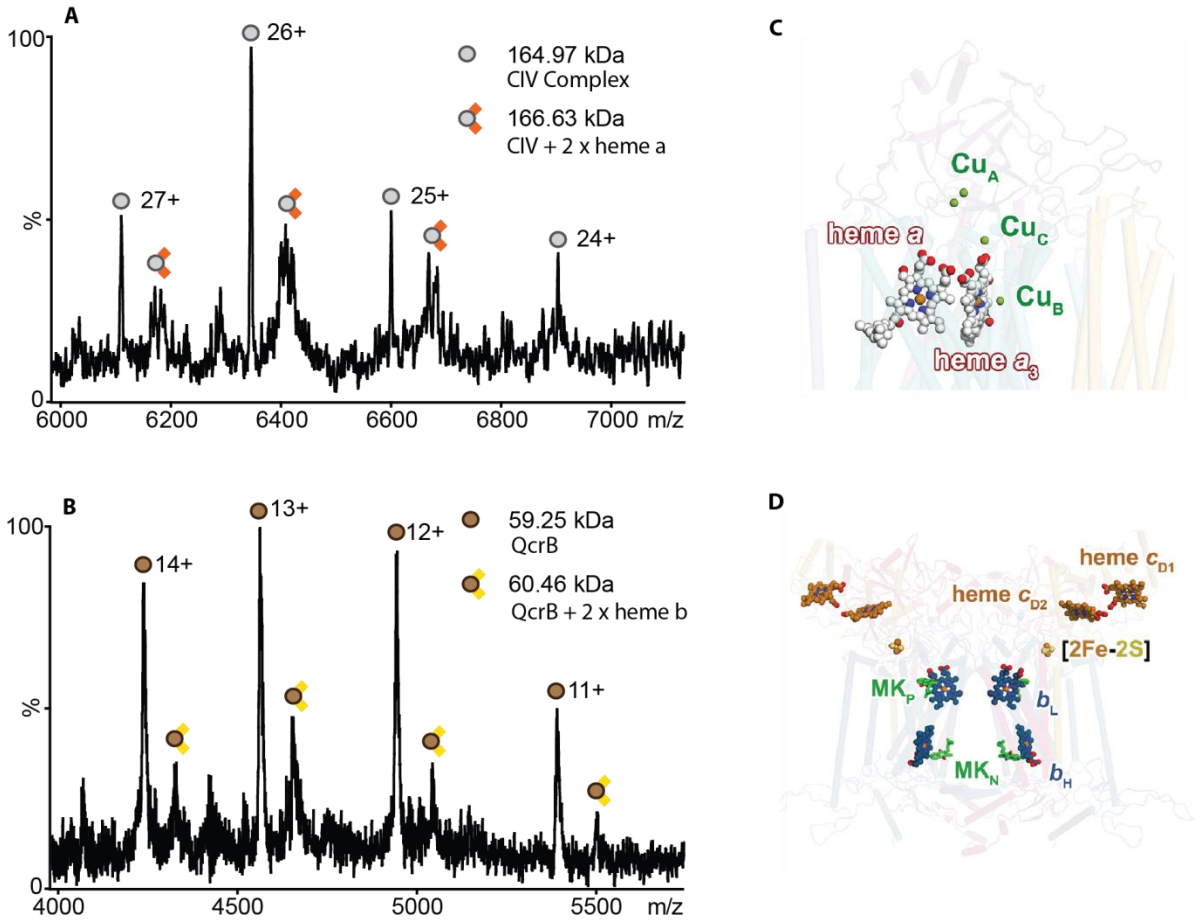


Figure 4-8 The prosthetic groups binding in the supercomplex CII/CIV. A, the native MS spectrum of CIV complex binding with two heme a in MS buffer containing 0.007% DMNG. **B,** the native MS spectrum of subunit QcrB binding with two heme b in MS buffer containing 0.02% DDM. **C** and **D,** the spatial location of relevant prosthetic groups in CIV and CIII complex respectively. Figure is from H. Gong *et al.*[2].

Our results were supported by observations in cryo-EM. As shown in Figure 4-8 C and D, the prosthetic groups heme a/a_3 and heme b_H/b_L were identified in CIV and CIII complex respectively, playing important roles in the electron transfer in the supercomplex. The complete electron transfer pathway in this respiratory CIII/CIV supercomplex was illustrated in the published paper of H. Gong *et al.*[2]. In the mechanism they proposed (Figure 4-2), heme b_L receives one electron from donor MK_P and delivers it to receiver MK_N via heme b_H. The edge-to-edge distance between

the two heme *b* groups is 12 Å, which ensures the electron translocation in QcrB subunit. Within CIV, the electron from Cu_A in QcrC is transferred through heme *a* to the coordinated group heme a_3-Cu_B, the oxygen reduction centre. The short distance between the two heme *a* groups (13 Å) allows the rapid inter-heme electron transfer.

4.5 Conclusion

The accurate masses of supercomplex components were obtained by MS approaches. We confirmed the molecular weights of protein subunits in both native and denaturing methods, and the results are identical. However, some values are quite distinct from the structural model based on cryo-EM density map, indicating that residues that superimpose in complex structures interfere with the cryo-EM observation, which may affect the identification of critical structures or adducts. Lipidomics identified the lipid composition of the transmembrane complex. Combining this information with their location illustrated in cryo-EM, we are able to figure out the function of binding lipids in the mechanism of the respiratory supercomplex.

As a powerful technique to investigate intact membrane protein assemblies, native MS revealed the interactions of the components inside the respiratory supercomplex, including interactions between protein and protein, and protein and small molecule. In contrast to the spatial location directly observed by cryo-EM, the interactions maintained in native MS are important in the structure and function of the supercomplex.

In summary, the study of the mycobacterial respiratory supercomplex showed that the comprehensive understanding of protein assemblies can be achieved by the application of cryo-EM coupled with MS methods. Cryo-EM directly profiles the macromolecule and MS provides the intracomplex information of individual components and their interactions. These findings provide a foundation for drug design for disease caused by analogous bacteria.

4.6 Methods

Denaturing LC-MS analysis. The supercomplex fractioned from SEC purification was loaded into a 5-µl sample loop using an autosampler by a full-loop method with an overfill factor of 1.4, and then transferred to a 1-µl injection loop using a loading pump at a flow rate of 5 µl/min with 99% LC buffer A (0.1% (v/v) formic acid) and 1% LC buffer B (IPA: ACN: H_2O as ratio of 10: 9: 1, 0.1% (v/v) formic acid). The sample in the injection loop was injected onto a monolithic column (ProSwift^{TM} RP-4H, Analytical, 100 µm × 25 cm, Thermo Fisher Scientific) using a nanopump at a flow rate of 300 nl/min. The supercomplex was denatured and dissociated through the monolithic column at 40 °C by a gradient starting from 99% LC buffer A and 1% LC buffer B. After 3 min, LC buffer B was ramped to 99% over 32 min, and held for 10 min, then reduced to 1% again for 2 min. The column eluent was delivered via a dynamic nanospray source to a hybrid LTQ Orbitrap mass spectrometer (Thermo Scientific). The conditions used for acquisition were: spray voltage 1.8 kV and capillary temperature 175 °C. The LTQ-Orbitrap XL was operated in positive ion mode. Survey full-scan MS spectra were acquired in the orbitrap (m/z 350 − 2,000) with a resolution of 60,000.

Native mass spectrometry. The supercomplex fractioned from SEC purification was buffer exchanged into MS spray buffer using biospin column (Zeba^{TM} Spin Desalting Columns, molecular weight exclusion limit 40 kDa). Spray buffer was 200 mM ammonium acetate (pH 8.0) containing 2×CMC of relevant detergents. The sample was introduced into a modified Q-Exactive mass spectrometer (Thermo) and ions were transferred into the HCD cell following a gentle voltage gradient as the description in Section 2.7. The capillary voltage was maintained at 1.2 to 1.4 eV at temperature of 200 °C. The optimised acceleration voltage for intact supercomplex is 150 V and 100-120 V in the source and HCD cell, respectively. The collisional dissociation of dimeric SodC was performed with the voltage ramp from 0-150 V in HCD cells. Backing pressure was maintained around 1.20×10^{-9} mbar. Data was analysed using Xcalibur 2.2 SP1.48[2].

References

(1) Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **1961**, *191*, 144-148.

(2) Gong, H.; Li, J.; Xu, A.; Tang, Y.; Ji, W.; Gao, R.; Wang, S.; Yu, L.; Tian, C.; Li, J.; Yen, H.-Y.; Man Lam, S.; Shui, G.; Yang, X.; Sun, Y.; Li, X.; Jia, M.; Yang, C.; Jiang, B.; Lou, Z.; Robinson, C. V.; Wong, L.-L.; Guddat, L. W.; Sun, F.; Wang, Q.; Rao, Z. An electron transfer path connects subunits of a mycobacterial respiratory supercomplex. *Science* **2018**.

(3) Lippens, J. L.; Egea, P. F.; Spahr, C.; Vaish, A.; Keener, J. E.; Marty, M. T.; Loo, J. A.; Campuzano, I. D. G. Rapid LC-MS Method for Accurate Molecular Weight Determination of Membrane and Hydrophobic Proteins. *Analytical chemistry* **2018**, *90*, 13616-13623.

(4) Whitelegge, J.; Halgand, F.; Souda, P.; Zabrouskov, V. Top-down mass spectrometry of integral membrane proteins. *Expert Review of Proteomics* **2006**, *3*, 585-596.

(5) Souda, P.; Ryan, C. M.; Cramer, W. A.; Whitelegge, J. Profiling of integral membrane proteins and their post translational modifications using high-resolution mass spectrometry. *Methods* **2011**, *55*, 330-336.

(6) Magalon, A.; Arias-Cartin, R.; Walburger, A. Supramolecular organization in prokaryotic respiratory systems. *Adv Microb Physiol* **2012**, *61*, 217-266.

(7) Dunphy, P. J.; Gutnick, D. L.; Phillips, P. G.; Brodie, A. F. A New Natural Naphthoquinone in Mycobacterium Phlei . Cis-Dihydromenaquinone-9 Structure and Function. *Journal of Biological Chemistry* **1968**, *243*, 398-&.

(8) Azerad, R.; Cyrot-Pelletier, M. O. Structure and configuration of the polyprenoid side chain of dihydromenaquinones from Myco- and Corynebacteria. *Biochimie* **1973**, *55*, 591-603.

(9) Azerad, R.; Cyrot, M. O.; Lederer, E. Structure of the dihydromenaquinone-9 of Mycobacterium phlei. *Biochem Biophys Res Commun* **1967**, *27*, 249-252.

(10) Upadhyay, A.; Fontes, F. L.; Gonzalez-Juarrero, M.; McNeil, M. R.; Crans, D. C.; Jackson, M.; Crick, D. C. Partial Saturation of Menaquinone in Mycobacterium tuberculosis: Function and Essentiality of a Novel Reductase, MenJ. *ACS Cent Sci* **2015**, *1*, 292-302.

(11) Honaker, R. W.; Dhiman, R. K.; Narayanasamy, P.; Crick, D. C.; Voskuil, M. I. DosS responds to a reduced electron transport system to induce the Mycobacterium tuberculosis DosR regulon. *J Bacteriol* **2010**, *192*, 6447-6455.

(12) Sartain, M. J.; Belisle, J. T. N-Terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC. *Glycobiology* **2009**, *19*, 38-51.

(13) D'Orazio, M.; Folcarelli, S.; Mariani, F.; Colizzi, V.; Rotilio, G.; Battistoni, A. Lipid modification of the Cu,Zn superoxide dismutase from Mycobacterium tuberculosis. **2001**, *359*, 17-22.

(14) Keith, K. E.; Valvano, M. A. Characterization of SodC, a periplasmic superoxide dismutase from Burkholderia cenocepacia. *Infect Immun* **2007**, *75*, 2451-2460.

(15) Suthammarak, W.; Somerlot, B. H.; Opheim, E.; Sedensky, M.; Morgan, P. G. Novel interactions between mitochondrial superoxide dismutases and the electron transport chain. *Aging Cell* **2013**, *12*, 1132-1140.

5 Conclusion

Because of their dynamic and transient interactions, it is a challenging task to describe a comprehensive profile of membrane proteins and their lipid environments. However, recent developments in native MS have enabled the preservation of the non-covalent interactions in integral membrane complexes, including protein-protein and protein-lipid interactions, in the gas phase. This successful technique provides high resolution and sensitivity in the mass analysis of macromolecules, allowing interrogation of structure and function of membrane proteins.

As the most compatible membrane mimetics for native MS, detergents are utilised to replace the lipids surrounding membrane proteins. Particles with fewer lipids are usually preferred for MS, since they enable a better recognition of the target protein and interrelated ligands. However, this method may lead to the stripping of critical lipids from the protein and, as a result, alteration of its stability and activity. To sustain and detect these non-specific protein-lipid interactions and understand membrane proteins correctly, a progressive lipid removal process was proposed to control the lipidation level of membrane proteins. The lipid shells of membrane proteins are peeled off layer by layer by changing detergent conditions or prolonging the incubation time. The stepwise stages of lipidation were captured by native MS. In contrast to other reconstituted systems such as liposomes, bicelles and nanodiscs, the lipoproteins obtained from progressive delipidation preserve their original neighbouring lipid inside the detergent micelles, creating a more native-like mimetic system. A full understanding of lipid identification, quantification and distribution can

be obtained via LC-MS based quantitative lipidomics on the lipoproteins from different stages of the delipidation. Moreover, with the introduction of other biophysical assays, as delipidation proceeds, it is possible to observe the change in protein properties and thus determine the impact of individual lipid components on the structure, function and oligomerisation of the membrane protein.

Adjacent membrane lipids can also be stabilised using lipodisqs, an alternative mimetic system. From an empty PC lipid disc to a native membrane protein lipodisq, this recently developed method has proven promising in the interrogation of protein-lipid interactions. The attractive advantages of novel lipodisqs method are as follows: firstly, in contrast to other disc-based mimetics, lipodisqs are a detergent-free system avoiding the detergent effect during extraction and purification. Furthermore, although optimal lipodisq preparation varies slightly with different target membrane proteins, compared to the time and cost involved using detergents, lipodisqs offer a relatively standard protocol for membrane protein isolation, which simplifies sample preparation. However, despite successful application of lipodisqs in many biophysical techniques, it is still a relatively new technique in native MS which may require instrument modification for particle ionisation and liberation in the gas phase. Additionally, the heterogeneous composition of the lipodisq leads to a more complex spectra outcome, and therefore, better resolution and deconvolution are required for spectra interpretation.

MS can, however, be combined with other techniques and methodologies. Such a mass-based method can provide additional information where visual resolution is

limited, for example in techniques like cryo-EM and X-ray crystallisation. In the elucidation of the respiratory supercomplex, for example, MS can provide confirmation of protein subunits and endogenous lipids as well as protein-protein and protein-ligand interactions in the supercomplex. This kind of combination of multi-technology gives a full-scale insight of complex biophysical systems.

In summary, MS is a powerful technique in the investigation of membrane protein-lipid interactions. In our study, native MS coupled with LC-MS based quantitative lipidomics is used to explore new analytical methods, examine novel protein-lipid particles and profile a complex membrane protein system. Although there is currently no one MS technology that allows investigation of all interactions between membrane proteins and lipids, with the further development of instrument and methodology, more and more applications of MS will expand into a wider research field, informing drug discovery and human health.

Abbreviation

2DP	3-[(3-methylphosphono-glycerolyl)phosphonyl]-[1,2-di(2,6,10,14-tetramethyl-hexadecan-16-yl]glycerol
ABC	ATP-binding cassette
ACN	Acetonitrile
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
C_8E_4	Tetraethylene glycol monooctyl ether
CCS	Collisional cross section
CID	Collision-induced dissociation
CL	Cardiolipin
CMC	Critical micelle concentration
cryo-EM	Cryo-electron microscopy
CYMAL-6	6-cyclohexyl-1-hexyl-β-D-maltoside
DDM	n-dodecyl-β-D-maltopyranoside
DM	n-decyl-α-D-maltopyranoside
DMNG	Decyl maltose neopentyl glycol
E. coli	*Escherichia coli*
ESI	Electrospray ionisation
Fos-Choline-10	n-Decylphosphocholine
FPLC	Fast protein liquid chromatography
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
HCD	Higher-energy collisional dissociation
IM	Ion mobility
IMAC	Immobilised metal affinity chromatography
IPA	Isopropanol

IPTG	Isopropyl β-d-1-thiogalacto-pyranoside
LB	Luria Broth
LC	Liquid chromatography
LDAO	Lauryldimethylamine-N-oxide
LMNG	Lauryl maltose neopentyl glycol
LPS	Lipopolysaccharide
M. smegmatis	*Mycobacterium smegmatis*
M. tuberculsis	*Mycobacterium tuberculsis*
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionisation
MATE	Multidrug and toxic compound extrusion
MK	Menaquinone
MKH_2	Menaquinol
MPLA	Monophosphoryl Lipid A (Synthetic) (PHAD®)
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSP	Membrane scaffold protein
nESI	Nano-electrospray ionisation
NG	n-nonyl-β-D-glucopyranoside
NMR	Nuclear magnetic resonance
NTSR	Neurotensin receptor
OD	Optical density
OG	n-octyl-β-D-glucoside
OGNG	Octyl glucose neopentyl glycol
P. fluorescens	*Pseudomonas fluorescens*
PC	Phosphatidylcholine

PDB	Protein Data Bank
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PTM	Post-translational modification
Q-ToF	Quadrupole Time-of-Flight
RF	Radio frequency
RP	Reversed-phase
RPM	Rotations per minute
RT	Room temperature
Salipro	Saposin-lipoprotein particles
SapA	Saposin A
S-DGD	Sulfated diglycosyl diether
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SMA	Styrene-maleic acid
SMALP	Styrene-maleic acid lipid particle
SMAnh	Styrene-maleic anhydride
TB	Terrific Broth
TMD	Transmembrane domain
ToF	Time-of-flight
Triton X-100	α-[4-(1,1,3,3-Tetramethylbutyl)phenyl]-ω-hydroxy-poly(oxy-1,2-ethanediyl)
UV	Ultraviolet

Acknowledgements

First and foremost, I very much appreciate my supervisor, Prof. Carol Robinson, for providing me such an invaluable opportunity to know the extraordinary world of native mass spectrometry. It has been a real honour to study under her support and encouragement in my DPhil years.

I acknowledge with thanks to the mentorship from Dr. Cherine Bechara and Dr. Hsin-Yung Yen in my research. Their patience and advice enlightened and encouraged me to explore the unknown.

My thanks also go to all members from Robinson, Benesch and Baldwin groups. Particularly, I am grateful to Dr. Tim Allison, Dr. Idlir Liko, Dr. Di Wu and Dr. Shahid Mehmood for their guidance in experiments; Dr. Helena Hernandez, Dr. Todd Mize, Ms. Lauren Crawley, Dr. Victoria Cullen and Dr. Andrew Dolan for the lab management and maintenance; Dr. Victoria Beilstein-Edmands, Dr. Heidi Gastall, Dr. Iva Pritisanac and Dr. Kin Kuan Hoi for their suggestions in DPhil life and study. I also want to thank Joe, Mike, Jani, Mark, Miranda, Anna, Josh, Xingyu, Zihao and Siyun for inspirations and discussions. It has been a pleasure to work with them.

I want to give my huge gratitude to the contributions of my collaborators: Professor Anthony Watts from the University of Oxford provided lipodisq samples of rhodopsin; Dr. Hiroshi Hara from Saitama University provided *E. coli* strain UE54; Prof. Zihe Rao

from Tsinghua University provided protein samples of respiratory complex. Many thanks for their help and generosity.

Thanks to Xiao and Zheyu who took care of me in my illness, and my friends in the online chatting group who cheered me up through the tough times. And finally, the special appreciation is for my family and friends, especially my parents and cousin sister Xue, for their unlimited support and love all the time.

From bottom of my heart, I thank them all profusely.

Thanks to Zhu Yilong & Bai Yu

In memory of the unreproductive summer in 2018

For *Guardian*