Structural and Functional Studies with Native Mass Spectrometry — Insights into Membranes and Glycoproteins



Siyun Chen

Pembroke College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

in Physical and Theoretical Chemistry

2022

Acknowledgements

First of all, I would like to express my immense appreciation for my supervisor, Prof Carol Robinson. I am grateful that she gave me such a precious opportunity as a visiting research student to learn native mass spectrometry. I would have never known this could change my life and restart my PhD journey in Oxford. Carol also always provided many inspiring ideas for my research and enough freedom to explore my research projects. I genuinely thank Carol for her trust, encouragement and support, including funding when I overcame hardship during my PhD and difficulties in the project.

For the period of my PhD, I am also grateful for the practical training from Dror, Weston, and Di. They are great scientists. Sincere appreciation for Weston helping me transit from chemistry to molecular biology on my arrival in Oxford. I am grateful for his expertise and supervision on glycoprotein projects. I feel sincerely grateful for learning from and working with them. Dror is always kind and caring and patiently provides native mass spectrometry training and others in the laboratory. Last but not least, I would like to express a tremendous amount of appreciation for Di. It was my pleasure to learn all omics know-how from Di. In addition, Di always generously and selfless offers to help me.

I would like to express my gratitude to Vicky and Andy. They are the best as they are always responsible for keeping our working environment safe, tidy and organised. Thanks to their hard

work, my research can be conducted conveniently and smoothly, which helps improve efficiency and effectiveness during my PhD research.

I feel honoured and pleased to get to know and learn from my friends in the Robinson group and Benesch group. I always excellent excited and free to discuss various scientific questions in such a good scientific community. Truly appreciate their help and advice over the past three years. Thanks a lot, Tom, Corinne, Tarick, Phil, Hugh, Francesco, Denis, Jani, Haiping, Anna. H, Zihao and Dom.

The next ‘thank you goes for my precious friends in Oxford, Fangfang, Hang, Sushanta and Yujia (now Professor Dr Qing) and Chan Cao in Switzerland. I truly appreciate the interdisciplinary discussion among us, support and company in Oxford, which was an unforgettable time for me.

I feel especially thankful for Dr Xinhao Zhang and Prof Yun-Dong Wu, my former professors at Peking University. I could not have gone to the University of Oxford without the opportunity of visiting Oxford and generous support for me to pursue the scientific interests that excite me the most.

I am deeply grateful for Fabian being my teammate all the time (and my life partner from now on). Thanks for witnessing and closely supporting me throughout the process of my application for Oxford, my struggles during my PhD and to the end of Oxford time. Thanks for keeping me optimistic and faithful when I am down and being excited and happy about what I achieve.

Your encouragement is polaris in the dark, guiding me through a hard time. You are a great role model for my whole life.

Endless gratefulness and love for my beloved mum. Thanks for giving me so much freedom and lots of support to have my independence and bravery developed. You are the bravest and the most persistent woman I know. I feel proud and fortunate of being your daughter. There are also mountains of gratefulness for the rest of my families, my wise papa, my selfless stepfather, and my brother. Thank you all for staying next to me to reach my PhD.

Abstract

The membrane is a boundary between extracellular and intracellular environments or separates intracellular spaces. It also plays a vital and irreplaceable role in hosting membrane protein and is a matrix or media to exchange material and information in and out of the cell. Hence understanding the protein structure, tracing the signalling and investigating active ligands are essential. This thesis studies the structure details of glycoproteins, molecular events of rhodopsin in disc membrane, and the endogenous ligands of voltage-dependent potassium ion channel beta subunits. Native mass spectrometry (native MS) monitors the photoactivation and regeneration of rhodopsin in membrane in Chapter 3. The retinal isomerisation and hydrolysis process are slower than in the detergent environment. It is also discovered that the membrane component, N-ret-PE is a source of cis-retinal for rhodopsin regeneration. This pathway is independent of other cells and does not require conversion between retinal and retinol, which may improve the efficiency of vision persistence. Later in Chapter 4, the phototransduction of rhodopsin in membrane is further studied with native MS. The GDP release from the transducin is directly observed on transducin when it couples with rhodopsin. Hence the activated Gtα-PDE6 hydrolyses and releases cGMP. This detection of photoactivation and transduction in the membrane can be used for rhodopsin-targeting compounds. These provide insights into the regulation of rhodopsin-base active ligands for rhodopsin signalling. Apart from the static observation of protein structure, employment of native MS reveals occupancy of N- and O-glycans, LacNAcs, GlcNAc branches on N-glycans and linkage of glycans applied in

transferrin, fetuin, epoetin and alpha 1-acid glycoprotein in Chapter 5. In Chapter 6, native MS detects cofactor and active ligands that monitor Kvβ2, both aldo-keto reductase and potassium ion channel beta subunit. An endogenous inhibitory ligand is likely to be an inhibitor of Kvβ2, maintaining an increasing redox state in the ratio of NADPH/NADP+. The finding of a small ligand may result in a possible drug prototype for sleep rhythm. This thesis shows that the new application of native MS providing detailed information in structural, functional and metabolomic biology studies. Together these studies by the means of native MS advance the understanding in protein structure and functions.

Declaration

This thesis describes research carried out between January 2019 and August 2021. The discussion is in the form of an integrated thesis (as approved in the sub-department of Physical and theoretical Chemistry from Hilary Term 2021/2022). The thesis is a hybrid of conventional chapters (Chapter 1, 2, 6, 7) and scientific papers (Chapter 3, 4, 5) that have been published in scientific journals. The presented data were acquired and analysed by myself unless explicitly stated; each chapter summarises the contributions to the described work. For stylistic reasons and since the work described in this thesis involved collaboration of varying degrees, I have used the third-person pronoun ‘we’ throughout the thesis.

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

Siyun Chen

October 2022

Contents

Chapter 1 Recent Advances in Biological Mass Spectrometry 1

1.1 Overview 1

1.2 Mass Spectrometry 10

1.2.1 Ion Sources 10

1.2.2 ESI Process Mechanisms 14

1.2.3 Mass Analysers 17

1.2.4 Detectors 31

1.2.5 Instrumentation Used for This Dissertation 32

1.3 Native Mass Spectrometry for the Elucidation of Structure and Function of Protein Assemblies 35

1.3.1 Advances in Membrane Protein Studies 37

1.3.2 The Evolution and Progress of Membrane Mimetics for Membrane Proteins 38

1.4 The Aim of the Dissertation 43

Chapter 2 Rhodopsin and Its Disc Membrane 46

2.1 Renewal of Rod Discs: Formation and Shedding 46

2.2 Organization of Rhodopsin in A Disc 48

2.3 Rhodopsin Photoactivation 53

2.4 Rhodopsin Downstream Signalling 56

2.5 Regeneration of Rhodopsin Chromophore 57

2.6 Rhodopsin and Lipids 60

2.7 Rhodopsin, An Important Target for GPCR Studies 63

Chapter 3 Monitoring Photoconversion and Regeneration of Rhodopsin in Disc Membranes 69

3.1 Introduction 72

3.2 Methods 72

3.3 Results and Discussion 81

3.4 Conclusion 89

3.5 Supplementary Information 91

Chapter 4 Capturing Rhodopsin, A G-Protein Coupled Receptor, Signalling Cascade across A Native Membrane 104

4.1 Introduction 107

4.2 Methods 107

4.3 Results and Discussion 112

4.4 Conclusion 122

4.5 Supplementary Information 124

Chapter 5 Native Mass Spectrometry Meets Glycomics – Resolving Structural Detail and Occupancy of Glycans on Intact Glycoproteins 129

5.1 Introduction 132

5.2 Methods 135

5.3 Results and Discussion 138

5.4 Conclusion 155

5.5 Supplementary Information 157

Chapter 6 Probing Endogenous Ligands of Potassium Ion Channel Beta Subunit Potential for Sleep Regulation 162

6.1 Introduction 162

6.2 Methods 168

6.3 Results and Discussion 175

6.4 Conclusion 182

6.5 Supplementary Information 184

Chapter 7 Concluding Remarks and Outlook 187

References 193

Chapter 1 Recent Advances in Biological Mass Spectrometry

1.1 Overview

The cell is the basic unit of life. A cell is composed of nucleotides, proteins, and other sub-cellar structures. Nucleotides serve as the database and intermediate templates to initiate the physiological events and various sub-cellar structures are the organs of the cell providing space to carry out specific tasks. Proteins carry out the functions in the cell, encoded by the genome. These proteins form higher order structure and shuttle between organelles to execute their roles. Nowadays, 20,000 to 25,000 proteins are expected to be encoded by the genome and there are more considering beyond estimating the number of protein-coding gene[1-4]. These proteins play a variety of roles including among many others: catalysis of substrates, signalling, transporting materials, playing an important role in communicating between intracellular and extracellular environments.

Approximately 30% of all genes are known to encode membrane proteins[3]. Membrane proteins are embedded in or associated with the phospholipid layer and together with the lipid membrane compartmentalize cells and organelles. Certain membrane proteins provide specific communication channels between cells, compartments, and the surrounding environment. The specificity and the particular composition of membrane proteins are fundamental for specialized cells and their particular cellar function[5].

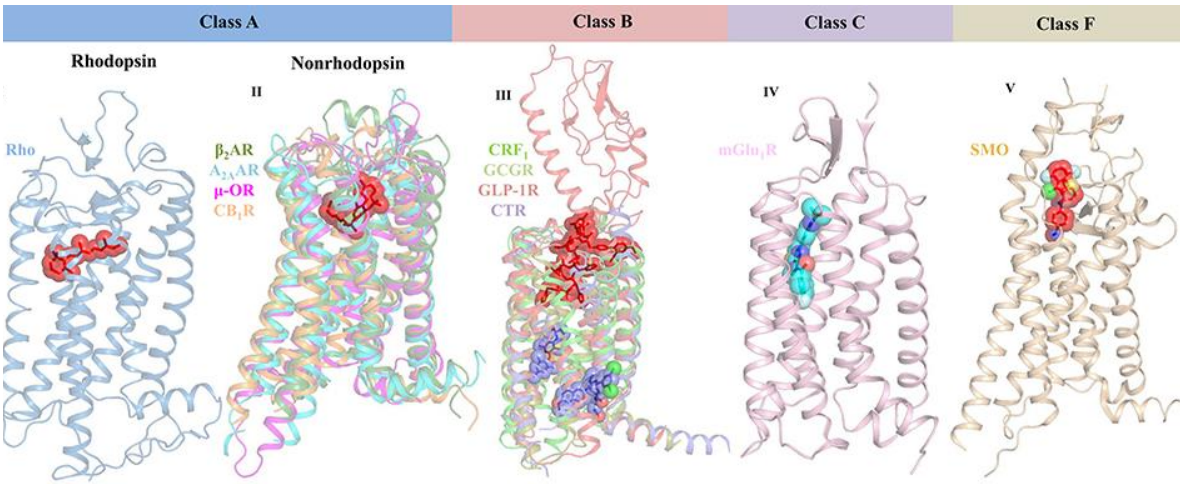


Figure 1.1 Crystal structures of representative GPCR-ligand complexes from classes A, B, C, and F presenting diverse ligand-binding sites. Class A GPCRs are classified into rhodopsin (bRho, PDB ID: 2HPY) and nonrhodopsin GPCRs. The representative structures of class A nonrhodopsin GPCRs which are further subdivided into aminergic-like (β2AR, PDB ID: 3P0G), nucleotide-like (A2AAR, PDB ID: 3QAK), peptide-like (μ-OR, PDB ID: 5C1M), and lipid-like receptors (CB1R, PDB ID: 5XRA) along with their bound ligands are shown. Similarly, representative structures for class B (CRF1 PDB ID: 4K5Y, GCGR PDB ID: 5EE7, full-length GLP-1R PDB ID: 5NX2, and CTR 5UZ7), class C (mGlu1R PDB ID: 4OR2), and class F (SMO PDB ID: 4QIN bound to negative allosteric modulator) are shown. Receptors are shown in cartoon representation and the ligands are shown as stick models with transparent surfaces. Bound ligands are represented as red sticks, antagonists are shown as purple sticks, and negative allosteric modulator is shown as blue stick model. Figure from Basith, et al..

Membrane proteins consist of two groups, integral membrane proteins and peripheral membrane proteins. Integral membrane proteins have one or more segments embedded in a phospholipid bilayer. It is known that integral membrane proteins can contain one or more α helix[6-9] spanning into the membrane; or multiple β strands can form membrane-spanning barrels[10-12] or anchoring to membrane can occur with covalently attached hydrocarbon chains[5,13-15] such as G protein[16] or PDE6[17]. Unlike integral membrane protein, peripheral membrane proteins often do not interact directly with the hydrophobic fatty acids of the lipid bilayer, and instead, they usually interact with integral membrane proteins or polar head groups of lipids, like phospholipase A2[18,19]. Membrane proteins have very important structural and functional roles including energy metabolism, signal transduction, transport and immune response. Over 50% of marketed small-molecule drugs[20] target human membrane proteins even though membrane proteins account for only 30% of the human genome[3]. The G-protein coupled receptor (GPCR) superfamily is the most targeted followed by ion channels; however, recent efforts to find new drug targets have led to a focus on solute carriers[21,22]

G protein-coupled receptors are the largest family membrane proteins and mediate most of the physiological responses to hormones, neurotransmitters, ions, photons and other stimuli as well as being responsible for vision[23], olfaction[24] and sweet, umami, and bitter taste[25]. GPCRs in vertebrates are commonly divided into five families on the basis of their sequence and structural similarity[26,27]: rhodopsin (family A), secretin (family B), glutamate (family C), adhesion and Frizzled/Taste2. At the most basic level, all GPCRs are characterized by the

presence of seven membrane-spanning α-helical segments separated by alternating intracellular and extracellular loop regions. The classical role of GPCRs is to couple the binding of agonists to the activation of specific heterotrimeric G proteins, leading to the modulation of downstream effector proteins via allosteric regulation. Looking into the downstream effects, GPCRs have much more complex downstream behaviours. Individual GPCRs have their preference for specific G proteins that constitute activity[28,29] and unique combinations for signal transduction. Similarly, GPCR desensitisation includes receptor phosphorylation events, arrestin-mediated internalization into endosomes, receptor recycling and lysosomal degradation[30,31]. Oligomerization, localization to specific organelle membrane and unique regulation by particular lipids further complicate GPCR molecular activities.

The structure and activation of GPCRs are generally conserved and it is well-known to signal via G proteins uniformly. But their heterogeneous and complex downstream regulation as discussed establish the bases for multifaceted functions in vertebrates and prove their significance in various physiological functions and so have great potential as therapeutic targets for a broad spectrum of diseases[32].

Another type of membrane protein, transporters allow specific materials to pass through the cellar membrane. They control the cellular nutrient levels and get rid of the waste from the cell and regular the cell volume. The transporter family can be subdivided into four groups based

on the vast majority of compounds: (1) ATP-binding cassette transporter; (2) ATPases (3) ion channels (4) solute carrier proteins (SLC).

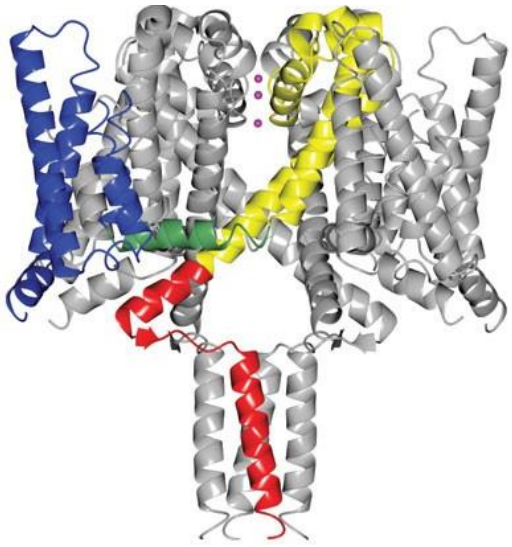


Figure 1.2 Structure representation of the NavMs tetrameric channel. One of the monomers is coloured as follows: voltage sensor (blue), S4–S5 linker (green), pore helices (yellow), C-terminal domain (red), and sodium ions (purple). The other three monomers are depicted in grey, for ease of viewing. The figure is adapted from Sula et al, (PDB Code: 6YZ2).

Mammalian ATP-binding cassette (ABC) transporters are involved in the cellular export of several groups of molecules, including cholesterol and sterols, lipids, retinoic acid derivatives, bile acid, iron, nucleosides, and peptides. These pumps use energy released by ATP hydrolysis to move substrates across membranes in or out of cells or into cellular vesicles, against their electrochemical gradient[33]. ATPases catalyzes ATP into ADP to release energy and import metabolites through the membrane. All ATPases share a common basic structure and are

composed of two major components: F0/A0/V0 and F1/A1/V1. The F1 catalytic domain is located on the N-side of the membrane and is involved in the synthesis and degradation of ATP and in oxidative phosphorylation. The F0 transmembrane domain plays a role in the movement of ions across the membrane[34].

Ion channels are ion-selective macromolecular protein pores that ions traverse the cell membrane through the pores and may therefore affect the membrane potential for controlling cellular excitability[35]. Ion channels respond to mechanical, electrical (voltage-dependent ion channels), or chemical stimuli. Several different ion channels contain pore loops, short polypeptide segments that loop into the ion conduction pore from one side of the membrane (**Figure 1.2**).

The other kind of transporter worth mentioning is the solute carrier (SLC) superfamily, which recently is reputed to be potential drug targets. The SLC superfamily currently includes 458 transport proteins in 65 families that carry a wide variety of substances across cellular membranes. So far, the SLC family is found to act as passive facilitative transporters or secondary active transporter, in which the substance is transported in a thermodynamically favourable way in high-to-low gradient or one substance passes through SLC at the cost of the free energy generated by the other transported substance goes down its electrochemical gradient. SLC proteins are relevant to a variety of cellular and physiological processes[36]. For instance, SLC6 and SLC18 regulate the concentration of neurotransmitters in synapes[37] and

also many of them participate in transporting nutrients between tissues like the blood-brain barrier[37,38].

Glycosylation of mammalian proteins

Glycosylation is a post-translation modification (PTM) responsible for attaching sugar moieties to specific amino acids of the proteins or lipids. For example, most secreted proteins and membrane-bound proteins expressed in the endoplasmic reticulum are glycosylated. Protein glycosylation encompasses N-glycans, O-glycans, and glycosaminoglycans (frequently termed proteoglycans). N-glycans are linked to asparagine residues of proteins, specifically a subset residing in the Asn-X-Ser/Thr motif, whereas O-glycans are attached to a subset of serines and threonines[39].

Glycosylation happens via biosynthetic-secretory process in the endoplasmic reticulum (ER) and Golgi apparatus. In ER, it helps monitor the status of protein folding to ensure only the properly folded proteins are trafficked to Golgi. Eventually, they are transferred bound mostly to the cell surface and extracellular compartments[40,41]. Different from encoding proteins based on gene, glycans are constructed without a template. Glycans are produced sequentially involving the distinct substrate specificities of enzymes, glycosyltransferase[42] and glycosidase[43] (**Figure 1.3**). Glycosyltransferases synthesize glycan chains, whereas glycosidases hydrolyze specific glycan linkages. Such a biosynthetic pathway controls the production and endogenous functions of different glycan structures. The structural diversity of

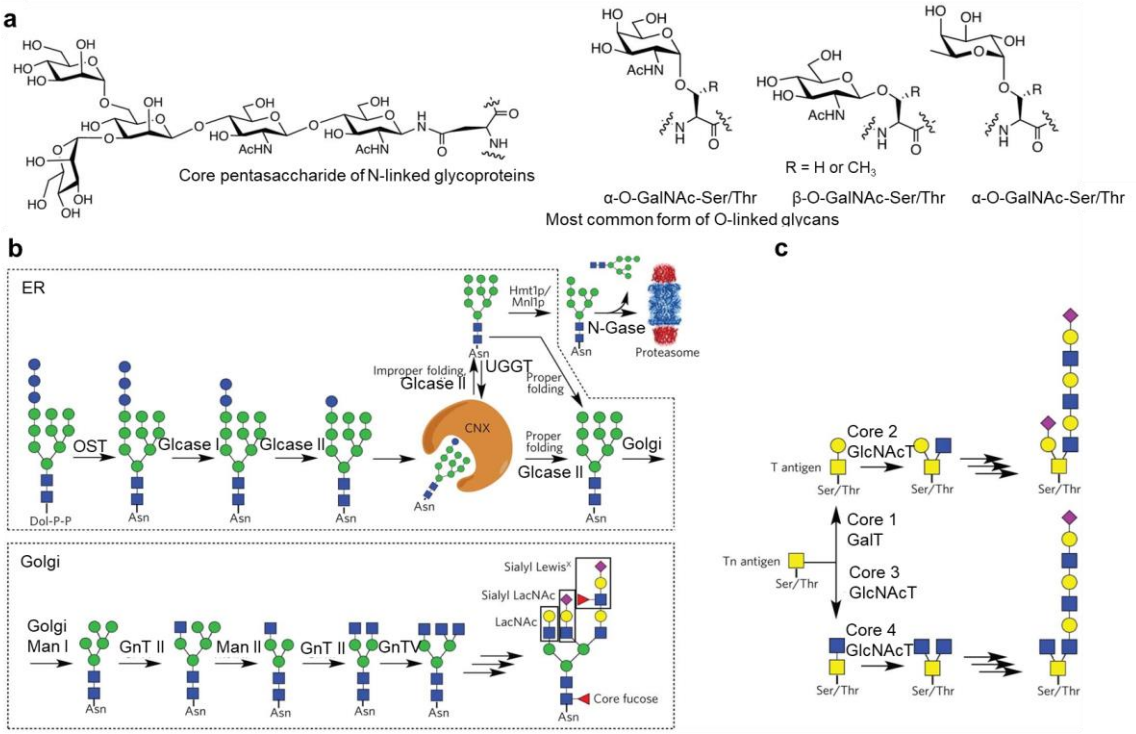


Figure 1.3 (**a**) Common types of protein glycosylation: core pentasaccharide common to all N-linked glycans and various types of O-glycosylation. (**b**) Biosynthesis of N-glycans. The biosynthesis of N-glycans is initiated in the ER by the transfer of a dolichol-linked glycan to an asparagine moiety of an Asn-Xaa-Ser/Thr sequon of a polypeptide. OST, oligosaccharyltransferase; glcase, α-glucosidase; UGGT, UDP–glucose-glycoprotein glucosyltransferase; Hmt1p, HnRNP methyltransferase 1; Mnl1p, mannosidase-like protein 1; N-gase, N-glucosidase; Man, α-mannosidase; GnT, N-acetyl glucosaminyltransferase. (**c**) Biosynthesis of O-glycans. The biosynthesis of O-glycans that start the structure with an O-GalNAc moiety begins in the Golgi, where various polypeptide polypeptide GalNAc transferases (GalNacT) attach a GalNAc residue. The resulting glycoprotein can be converted into various core structures by glycosyltransferases. Figure adapted from Wolfert and Boons.

the mammalian glycome is produced predominantly in a dynamic secretory pathway of the cell[44]. Additionally, glycans become increasingly oligomeric and branched in the Golgi apparatus[45] as they transit through this latter portion of the secretory system bound mostly for the cell surface and extracellular compartment.

Mammalian glycans participate in multiple cellular mechanisms that contribute to health and disease[46,47]. These basic structural paradigms have been found to regulate protein folding[48], cell adhesion[49], molecular trafficking and clearance[50], receptor activation[51,52], signal transduction[53], and endocytosis[54], evolution of multicellular and vertebrate organisms[55], cell differentiation[56], and disease[47]. For instance, species that are not yet properly folded are prevented from exit to the Golgi apparatus and are transported to the cytosol and degraded in the proteasomes if permanently misfolded. N-glycosylated in the ER, by which the quality control of folding is performed. Protein-linked monoglucosylated glycans, formed by glucosidase-dependent partial deglucosylation of the oligosaccharides transferred to N-glycosylation $Glc_3Man_9GlcNAc_2$, mediate glycoprotein recognition by two ER-resident lectins, membrane-bound calnexin and calreticulin[48,57]. In another example of glycan repertoire representing different cancer biomarkers, sialylated N-glycan isomers with α-2,3 and α-2,6 linkages play crucial and distinctive roles in diverse physiological and pathological processes such as alteration of human chondrocytes phenotype[58] and signs of tumor progression i.e. the development of cancer phenotypes[59]

There is a growing awareness that determining structures of glycans is vital to understanding their functions. The further advances in glycomics-based initiatives can assist by facilitating the development of techniques to produce complex glycan structures, detect glycan interactions with other molecules, and enable glycan-based methodologies to modulate cell processes[46,60,61]. A solution is proposed in the Chapter 5 to characterize the heterogeneity of glycans with native mass spectrometry[62]. The methodology focuses on structural analysis but also provide a basis for glycan interaction studies in a native environment.

1.2 Mass Spectrometry

Mass spectrometry is a technology for detecting the mass-to-charge (m/z) of ions. The molecular weight of ions is revealed by deciding the charge of the ions based on their isotopic distribution or charge distribution. The advantages of mass spectrometry are high resolution, speed, sensitivity and broad application to a wide range of analytes.

1.2.1 Ion Sources

To satisfy the fundamental condition of MS, the analyte first needs to be ionized and transferred to the gas phase in the ion source. Ions from liquid- or solid- phase are produced via injection or capture of electrons, protonation or deprotonation, by adduct formation with other species or gas-phase chemical reactions. Various ionisation techniques have been used to generate ions from a variety of different sample types, for example electron ionisation (EI), matrix-assisted

laser desorption/ionisation (MALDI), electrospray ionisation (ESI) and atmospheric pressure chemical ionization (APCI) and atmospheric pressure photoionization (APPI)[63].

Unlike traditional harsh ion sources like EI, soft ionisation techniques have been needed and developed which can ionize samples of a moderate mass range (up to 10 kDa), such as fast atom bombardment[64] and for the high mass range, up to tens of megadaltons as in the case of ESI[65,66] and MALDI[67,68]. Other soft ionisation techniques include laser-induced liquid beam ion desorption (LILBID)[69,70] and desorption electrospray ionisation (DESI)[71,72]. Considering ESI and its analogue-type nano ESI, which are widely used in the thesis, the mechanism of ESI and nESI are going to be discussed next.

In ESI, analyte is introduced to the mass spectrometer in solution. Charges accumulate in solution on the surface at the tip of the capillary and subsequently highly charged droplets are formed[73]. As the surface tension of the droplet breaks, the droplet shape forms the so-called Taylor cone[74]. Droplets forming from the Taylor cone first carry a charge dependent on the surface tension of the solvent and droplet diameter, as defined by the Rayleigh limit equation, using equation 1.1.

$$Z_R = \frac{8\pi}{e}(\varepsilon\gamma R^3)^{1/2} \qquad (1.1)$$

Where the charge Z_R predicted by the Rayleigh stability limit for Coulomb fissions[75], of water droplets of the same radius R as the globular proteins. where γ is the surface tension (0.072

N/m for water), ε_0 the electrical permittivity of vacuum (8.8×10^{-12} C^2/J m), e the elementary charge (1.6×10^{-19} C).

With solvent evaporating and charge density increasing, the potential further increases and the Rayleigh limit is exceeded so that increased and the solution-air interface is distorted. When the surface tension is eventually overcome, the droplet splits into smaller droplets through coulombic explosion (fission), in which the smaller droplets form with an estimated volume of 2% of the parent droplet but with 15% of their charge, thus the charge per unit of volume increases by a factor of ~7.5[76]. This process cascades and repeats promoted by further evaporation and desorption until these highly charged droplets generate into gas-phase ions.

With ESI, large protein species can be ionized into multiply charged ions allowing analysis by mass analysers and detectors[77]. The generated ions are charged consecutively as a Gaussian distribution. The mass of the analyte can be deconvolved by considering the *m/z* values of two adjacent peaks. In the case that ions generated via protonation, the mass and charge can be calculated using equation 1.2 and 1.3, where z_1 is the charge of the first peak, m_1 and m_2 are the *m/z* values of the first and second peaks respectively, m_p is the mass of a proton and *i* is the number of ions separating the first and the second peaks.

$$mass = \frac{i(m_2+m_p)(m_1+m_p)}{(m_2-m_1)} \tag{1.2}$$

$$z_1 = \frac{i(m_2+m_p)}{(m_2-m_1)} \tag{1.3}$$

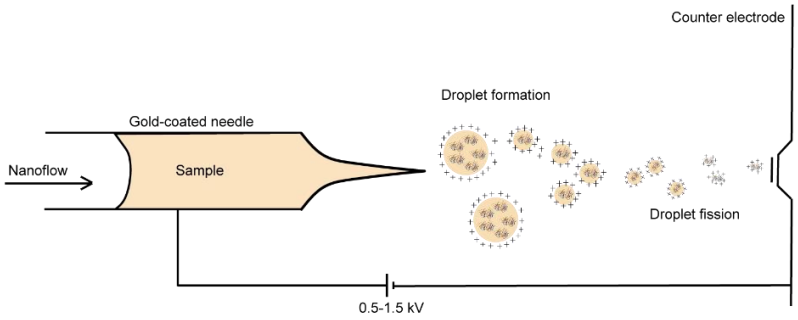


Figure 1.4 Schematic of the nano electrospray process. Sample solution is introduced into a conductive capillary (gold-coated needle) under an applied voltage relative to the downstream electrode and under the gas pressure of nanoflow. Ion accumulation leads to the formation of Taylor cone, which is the accumulation of charge causes a protrusion to appear on the end of the droplet, distorting the droplet into a conical shape known as the Taylor cone. With increasing field strength, the repulsive electrostatic force overcomes the surface tension and a charged jet of fluid is ejected from the tip of the Taylor cone when a critical value is attained. Macromolecule ions are produced and transmitted into the mass spectrometer for analysis.

Nano-electrospray ionisation (nESI) is a sub-type of ESI developed by Wilm and Mann in 1994[78]. The principles underlying nESI are the same as ESI but in a smaller volume at 1-3 µl, a smaller inner diameter of capillary needle orifice (1-5 µm in nESI) and very slow flow rates, in the range of 1-50 nL/min driven by 0.5-1.5 kV potential applied to a metallic-coating glass or quartz capillaries[79] **(Figure 1.4**). These changes enable and advance the analysis of biological samples including noncovalent protein assemblies. The reduced analyte amount for

measurement is a great help for biological analysis due to the limited quantity of biological samples. In addition, the initial smaller droplet size from nESI reduces the number and the energy of the collisions required to desolvate the macromolecules of interest. This results in more tolerance of salt contamination of the protein solution than conventional ESI, a gentler and more reliable ionisation for protein complexes, higher resolution and mass precision.

1.2.2 ESI Process Mechanisms

The mechanism in which protein desorption occurs has been discussed over decades. Three models have been proposed: ion evaporation model (IEM), charge residue model (CRM), and chain ejection model (CEM).

IEM describes that emitted droplets contain pre-charged ions and the variables involved in the process are the electric fields on the droplet surface, and the chemical properties of the ions, comprising the molecular structure[80] (**Figure 1.5a**). Following solvent evaporation, the droplet shrinks and the charge density increases until the electric field at the surface overcomes solvation forces, hence ejecting the ions from the droplet[81]. This mechanism could explain the charge state distributions of small molecules, however for larger ones (> 1000 Da) it seems to deviate from experimental evidence since it underestimates their charge state[82].

To explain the ionisation process of large molecules, the CRM is used for large macromolecules, such as folded proteins[81,83] (**Figure 1.5a**). According to this model, the final

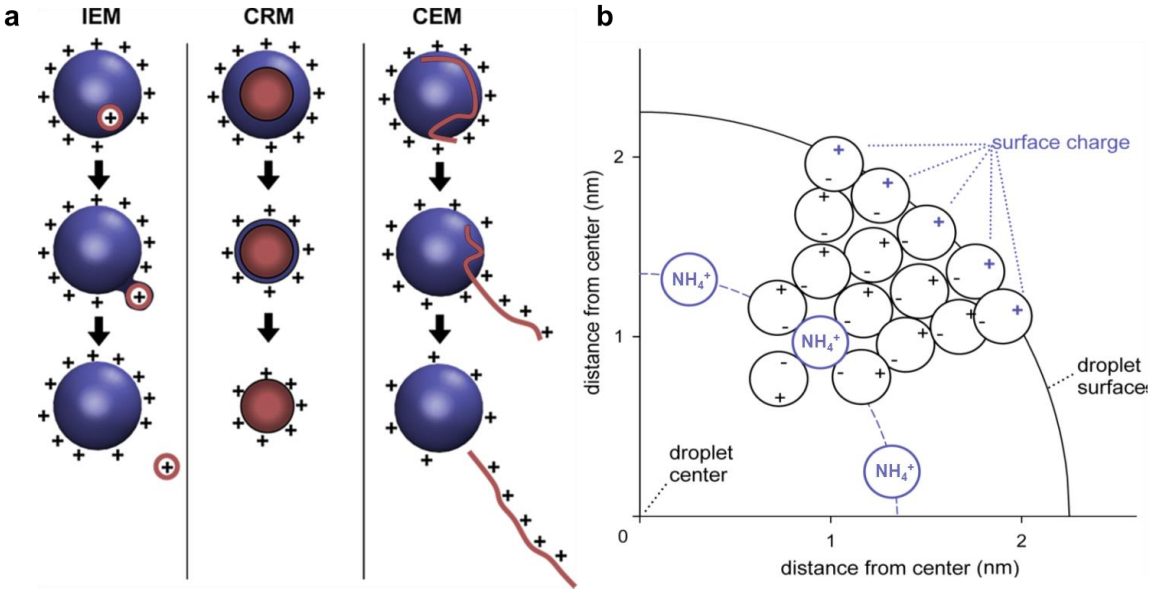


Figure 1.5 Models of electrospray ionization under positive mode, (**a**) where droplets (blue) with positive surface charge (illustrated with +) are sprayed from nESI and macromolecules are ionized (red, circle and streak represent folded and unfolded states respectively). IEM: Small ion ejection from a charged nanodroplet. CRM: Release of a globular protein into the gas phase. CEM: Ejection of an unfolded protein. (**b**) Schematic representation, illustrating the internal positioning of NH_4+ and water dipoles within a nanodroplet. In the presence of ammonium acetate that is compatible with ESI-MS, the ions within the droplet induce large-scale orientational polarization of the surrounding water. This solvation arrangement projects the excess charge to the surface. Figure is adapted from Konermann, et al..

charge state of the analyte is independent of the molecule's charge state in solution but rather on the droplet size and the mass of the molecule. This model proposes that the droplets undergo multiple evaporations and fission cycles until they approximately reach the size of the analyte. Further evaporation removed neutral solvent molecules to finally generate the gas-phase charged ion. The final charge on the molecule is closely related to the Rayleigh limit (thus to the diameter of the droplet) and the size of the molecule (which, for proteins, is roughly proportional to their mass)[84]. This model was confirmed by experimental observations as equation 1.4 for a diverse range of proteins, regardless of their surface chemistry[85].

$$Z_R = \frac{\gamma\pi}{e}\left(\frac{\rho\gamma\varepsilon 0}{4\pi NA}\right)^{1/2} M^{1/2} = 0.078 M^{1/2} \quad (1.4)$$

Where radius of the proteins is evaluated from the molecular mass M of the proteins in Daltons, by assuming that the density of the globular proteins is the same as that of water ρ = 1 g/cm, numerical factor 0.078 for M in Da.

Clean $[M + z\,H]^{z+}$ protein ions can be formed via the CRM under conditions where the excess charge of the droplet is predominantly due to protons. Under physiological conditions (aqueous solution at pH ≈ 7), ammonium acetate is used because it is compatible with ESI-MS because it consists of two volatile components, i.e., ammonia and acetic acid. Like other cations, NH_4^+ associates with protein carboxylates as the droplet dries out. The resulting adduct then loses NH_3 during collisional activation in the ion sampling region of the mass spectrometer.

protein COO^- + NH_4^+

→ protein COO^- NH_4^+

→protein COOH + NH_3

This two-step process generates a protonated carboxylic acid group, which contributes toward the formation of a clean [M + z H] $^{z+}$ signal. NH_4^+ can even displace previously bound metal cations.

The third model, CEM, better explains disordered polymers with solvent-accessible hydrophobic residues, such as unfolded proteins[83,86] (**Figure 1.5a**). Since the exposure of hydrophobic regions to charged solvent is energetically unfavourable, unfolded protein termini migrate to the surface and the protein is eventually expelled because of Coulombic repulsion. Given the higher surface exposure of unfolded proteins compared to folded ones, they carry a higher and broader charge state distribution.

1.2.3 Mass Analysers

Mass analysers play a key role in separating and analyzing ions. Once the ions being generated, they are directed to mass analysers via voltage-controlled ion lenses and skimmers. Several mass analysers have been developed to separate ions based on their kinetic energy, momentum, flight time, trajectory stability and resonance frequency[87-89] (**Table 1.1**). The performance of mass analysers is determined by their accuracy, resolution, sensitivity and other desirable properties like the high speed of analysis, unlimited mass range and sufficient ion transmission. Accuracy is the difference between an observed *m/z* value and the theoretical value. Current

state-of-the-art instruments have high accuracy, in the low parts per million (ppm), typically 1–5 ppm[90] or even less than 1 ppm at the cost of a more complex instrumental setup[91].

Resolution is defined as the peak width (Δm) at 50 percent of the peak height; the narrower the peak width the higher the resolution. Sensitivity is the intensity of signal recorded for a fixed concentration of the sample. For the aim of this dissertation, this thesis mainly describes three types of mass analysers: quadrupole, ion trap, and Orbitrap.

Mass Analyser	Measured Property	m/z Range	Resolution	Dynamic Range
magnetic sector	momentum/charge	10^5	10^5	10^7
time of flight	flight time	10^6	10^3 -10^4	10^4
ion cyclotron	cyclotron frequency	10^5	10^6	10^4
quadrupole	filters for m/z	10^3 - 10^4	10^3 - 10^4	10^5
ion trap	frequency	10^4	10^4	10^4
orbitrap	frequency	unlimited	10^5	10^4

Table 1.1 Characteristics of different mass analysers. Comparison on measured property, m/z range, resolution at 1000 Th and dynamic range of detection on magnetic sector, time of flight, ion cyclotron, quadrupole, ion trap and orbitrap.

Quadrupole

The most widely used mass analyser is quadrupole (**Figure 1.6**). It is composed of four parallel cylindrical rods and separates ions by the means of ion trajectory stability in an applied electric field[88,89]. Two consecutive rods have opposite voltage sign which then alternates as a function of a direct potential (U) and the amplitude (V) of a radio frequency (RF) voltage.

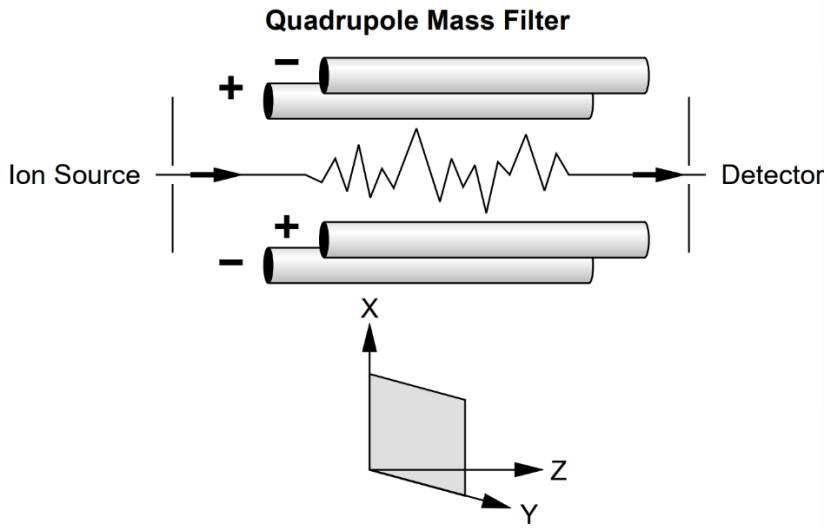


Figure 1.6 Diagram of a quadrupole mass analyser composed of four parallel cylindrical rods and. Two consecutive rods have opposite voltage sign which then alternates as a function of a direct potential and the amplitude of a RF voltage. For a given configuration of the driving voltages and quadrupole geometry, only the particles of a certain specific charge only ions with a certain m/z will be in resonance with a particular oscillating field and therefore get to the detector while non-resonant ions hit the rots and get annihilated during transmitting. Figure from Dawson, et. al..

An ion will be attracted by an oppositely charged rod until the rod switches polarity, at which point it will move in the opposite direction due to charge repulsion. If the ion reaches the rod before it switches sign it will be filtered out by the quadrupole. Hence, an oscillating ion will have a stable trajectory if the amplitude of the oscillation remains less than the distance between rods (r), otherwise, the ion will touch the rod and be discharged. For a given quadrupole with fixed r and constant angular frequency (ω) the stability area will be dependent on U and V as deduced from the Mathieu equation (Equation 1.5 and 1.6 respectively)[92], where a_u and q_u are proportionality constant of an ion in either the x or y direction.

$$U = a_u \frac{m\omega^2 r^2}{8ez} \tag{1.5}$$

$$V = q_u \frac{m\omega^2 r^2}{4ez} \tag{1.6}$$

Maintaining a constant U/V ratio allows for the scanning of ions with consecutive masses (**Figure 1.7**). In addition to this scanning mode quadrupoles can also operate in a static mode where only a particular ion is selected and allowed to pass through. Operating in this mode is desirable for the interrogation of ions from a complex mixture as the selected ion can be further fragmented later in its path through the spectrometer and mass identified. The U/V ratio will define the resolution of the quadrupole, hence maintaining a fixed U/V ratio will result in a fixed resolution through the spectrum[93]. However, this mode of operation will result in peak width broadening with increased *m/z* values. As a result, a more desirable mode of operation is varying the U/V ratio so that the peak width is constant. Establishing a high U/V ratio correlates

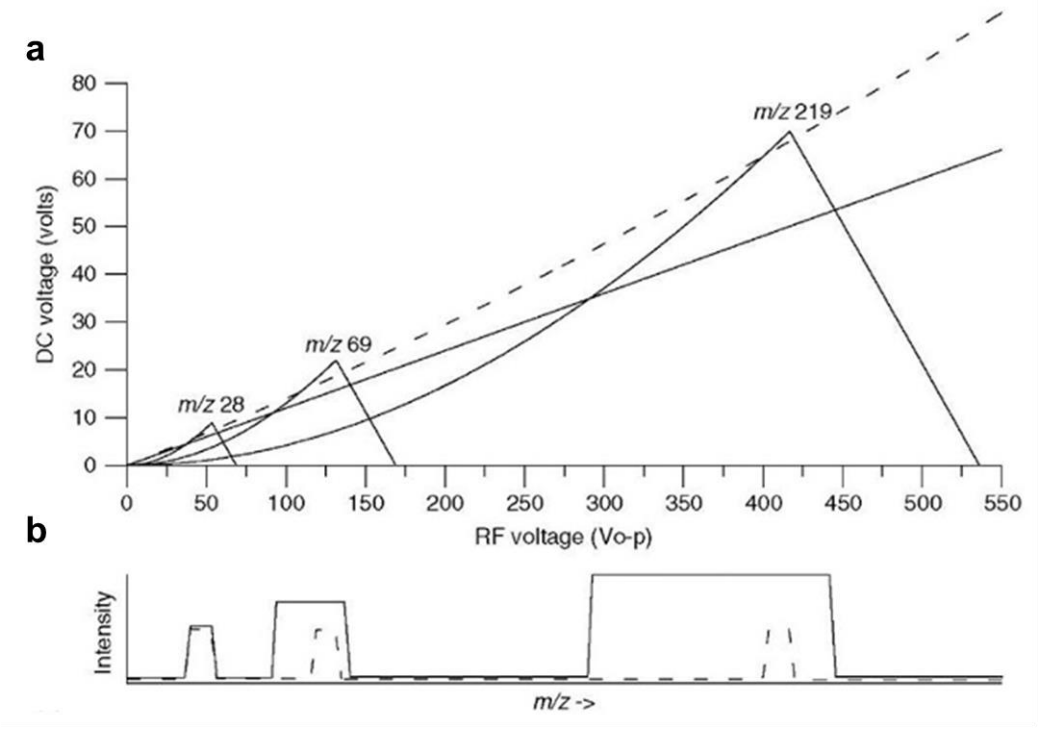


Figure 1.7 Diagrams of ion stability based on the Mathieu equation. (**a**) The dashed line and straight lines represent the mass scan line for achieving stable trajectories of ion with m/z 219 and 28 respectively. (**b**) An intensity profile of each line is shown, depicting that the dashed line will focus larger ions but meanwhile resulting in weak intensity for the smaller ions. Figure from Henchman and Steel.

---

with a high resolving power, however, when the quadrupole operates in only RF mode (U=0), the resolution becomes zero. In only-RF mode, the quadrupole no longer functions as a mass filter, nevertheless, this model is sufficient to focus ions towards the centre of the quadrupole. For a given RF voltage, ions with masses lower than the stability area will oscillate with unstable trajectories whereas ions with larger masses will be focused. However, the focusing

property is inversely proportion to *m/z* and ions with high *m/z* values with being poorly focused. The maximum *m/z* value that will be transmitted by a quadrupole is given by equation 1.7, where the operating frequency is denoted by Ω.

$$[m/z]_{\max} = a_u \frac{4eV}{0.706\Omega^2 r^2} \quad (1.7)$$

To achieve transmission and focusing of ions with large masses, quadrupoles that operate with a high RF amplitude and low frequency have been developed[94].

Several instrumental designs use quadrupoles in tandem by taking advantage of both modes of operation, as a mass filter with a non-zero DC potential or as an ion guide by applying an RF voltage only. For instance, in the Q-Exactive instrument, ions can be scanned or mass-selected in the first quadrupole then activated in a collision cell, which is a second quadrupole operating in RF mode only and filled with noble gas.

Ion trap

Ion trap has a similar mechanism as the quadrupole but with a different instrumental setup. Three electrodes are arranged in a sandwich geometry with a ring electrode in between two cap electrodes. In this device too, ions are subjected to forces applied by an RF field but the forces occur in all three, instead of just two, dimensions. The stable motion of ions in the linear quadrupole allowed ions freedom of motion in one dimension (z-direction); in the ion trap, stable motion allows no degrees of freedom. Hence, ions are trapped within the system of three

electrodes-a ring electrode and two end-cap electrodes of hyperbolic cross-section (**Figure 1.8**). The design of the ion trap principally makes it good at high sensitivity and high resolution (>106 at m/z >1000) accessible through slow scans, but unavoidably its mass measurement accuracy is relatively poor. Besides, it allows tandem mass spectrometry experiments available by performing sequential mass analysis measurements.

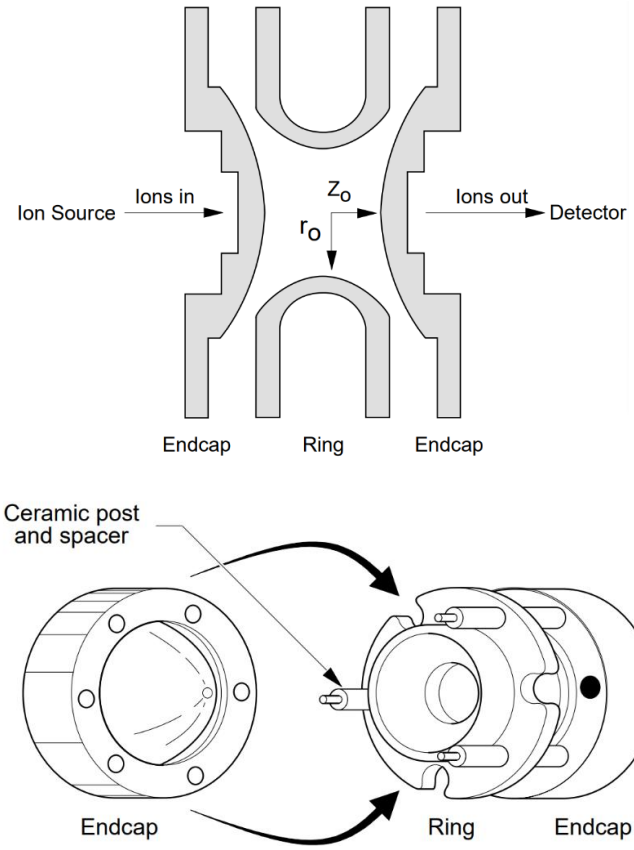


Figure 1.8 The ion trap consists of three electrodes, the central ring electrode, and two adjacent endcap electrodes. The schematic of the assembly shows how the electrodes are aligned and isolated using ceramic spacers and posts. The device is radially symmetrical, and r_0 and z_0 represent its size. Figure is adapted from Wong and Cooks.

Ion trap can carry out multiple-stage mass spectrometry (MS^n) simply by additionally operating which are performed sequentially in time. MS^n experiments can only be performed in multi-quadrupole instruments, sequentially in space but this may cause ion loss over tandem mass spectrometry.

An RF voltage is applied to generate an electric quadrupole field. Because the electric field is rotationally symmetric, it is convenient to consider only radial and axial (z) displacements. The potential ($\Phi r,z$) at any point in this field is given by equation 1.8:

$$\Phi r, z = (\mathrm{U} + \mathrm{V}\cos\omega t)\frac{r^2-2z^2+z0^2}{r0^2+2z0^2} \quad (1.8)$$

Note that r_0 is the internal radius of the ring electrode and z_0 is the closest distance from the centre to the end-cap, while U is the dc potential and V is the RF potential (zero-to-peak) applied between the ring and end-cap electrodes, ω is its angular frequency, and t is time.

Ions of a given m/z value may undergo stable motion in the trap for the reasons already given qualitatively. This second-order differential equation of the Mathieu form is divided into two independent equations describing the uncoupled motion of an ion in the r- and z-directions and finally rearranged into:

$$m/z = \frac{8V}{qz(r0^2+2z0^2)\omega^2} \quad (1.9)$$

Note that q_z = 0.908 when ion unstable so instable that their kinetic energies and z-direction excursions increase and they transit through the trap through a hole in the end-cap electrode. For traps built with the so-called ideal geometry, $r_0=\sqrt{2}\ z_0$, equation can be simplified into:

$$m/z = \frac{4V}{qzr0^2\omega^2} \tag{1.10}$$

The mass range of an ion trap can be calculated by substituting appropriate values into EQ. qz = 0.908 is the q_z value at which instability occurs in the normal mass-selective instability mode of operation. However, under resonance conditions, q_z becomes a variable provided the supplementary resonance frequency can be decreasing the size of the trap or lowering the frequency has been used to extend the mass range of the ion trap. However, the most straightforward method is to lower (q_z) eject using the dipolar electric field applied across the end caps to eject ions of a certain frequency from the trap. Conceptually, the ions exit the ion trap in a sequence of m/z values but at a lower RF amplitude.

A unique feature of the ion trap is the MS^n capability with unprecedented power in structural elucidation. The selectivity of MS^n means that a compound can be fragmented, and the resulting fragments are further isolated and analyzed to yield structural information about complex molecules in the presence of mixtures. A sequential MS^4 experiment was performed on a lipid released from octamer demonstrating dramatic control over the loss of the sugar monomers shown in **Figure 1.9**. The experiment allows simple, rapid identification of a complex molecular structure.

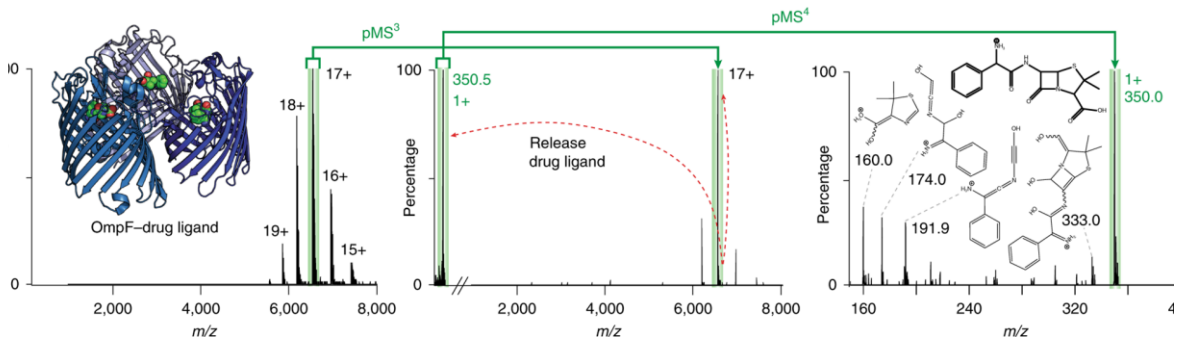


Figure 1.9 Schematic of in-built ion trap mass spectrometry to identify ligands or proteoform components of membrane protein assemblies. OmpF (PDB 4GCP) bound to drug ligand ampicillin is released from detergent micelles (pMS^2), the 17+ charge state (green) isolated for activation and the ligand released at *m/z* 350.5 (pMS^3). Isolation and fragmentation of the ligand (pMS^4) yields characteristic ions that identify ampicillin following database searching or spectral matching. Figure origins from Gault, et al..

Orbitrap

Orbitrap was invented to meet the need of high mass accuracy that is comparable to FTICR[95], instead of an instrumental architecture of high complexity. Therefore, orbitrap was developed based on the Kingdon trap[96,97].

Like Kingdon trap, orbitrap is constructed by an inner electrode in a spindle shape with an outer part shaped like a barrel that covers the entire inner electrode (**Figure 1.10**). with rims enclosing the trapping volume. When a static voltage is applied between the wire and the cylinder, the strong field attracts ions to the wire. Only ions that have enough tangential

velocity miss the wire and survive. Motion along the inner electrode is restrained by the field curvature caused by the flanges of the outer cylinder. Orbitrap analyser is described as below which employs orbital trapping in an electrostatic field with potential distribution[96]:

$$U(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}\left(R_m{}^2\right)\ln\left[\frac{r}{R_m}\right] + C \qquad (1.11)$$

where r and z are cylindrical coordinates (z = 0 being the plane of the symmetry of the field), C is a constant, k is field curvature, and Rm is the characteristic radius. This field is the sum of a quadrupole field of the ion trap and a logarithmic field of a cylindrical capacitor; therefore, it may be also called a quadro-logarithmic field.

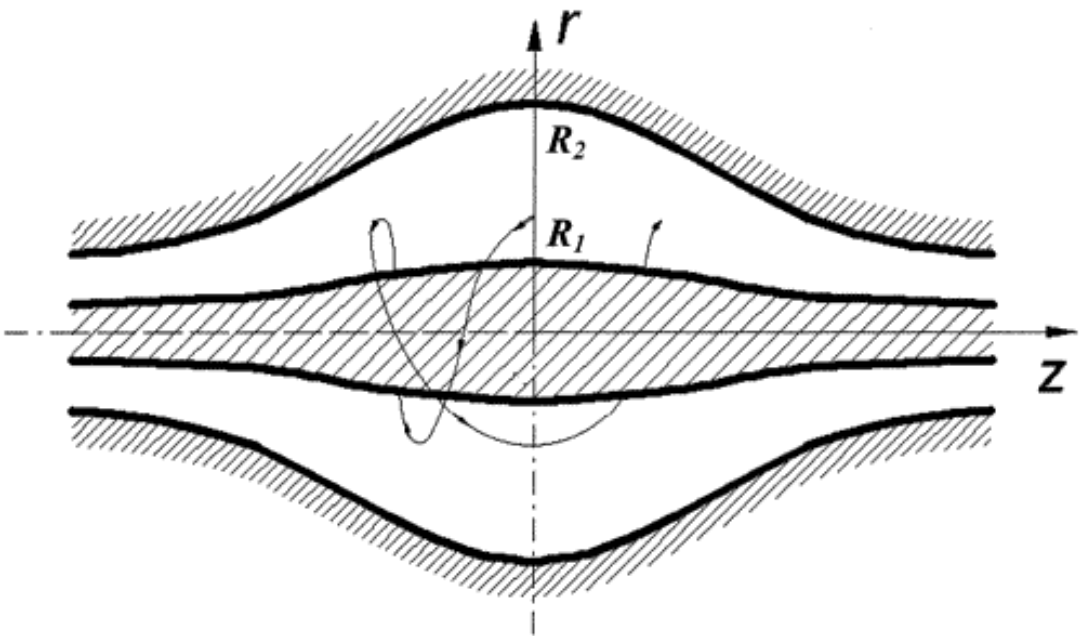


Figure 1.10 Equipotentials of the quadro-logarithmic field and an example of a stable ion trajectory. Figure is adapted from Makarov.

The trap consists of an outer barrel-like electrode and a central spindle-like electrode along the axis (**Figure 1.10**). The shape of these axially symmetrical electrodes could be deduced from equation 1.11:

$$z_{1,2} = \sqrt{\frac{r^2}{2} - \frac{R_{1,2}}{2} + {R_m}^2 \ln\left[\frac{R_{1,2}}{r}\right]} \quad (1.12)$$

, where index 1 denotes the central electrode, index 2 denotes the outer electrode, z = 0 is the plane of symmetry, and $R_{1,2}$ are the maximum radii of the corresponding electrodes.

Ion stable trajectories combine rotation around the central electrode with oscillations along the axis, resulting in an intricate spiral. Equation of motion in polar coordinates (r, φ, z) for ions with mass-to-charge ratio m/q

$$\ddot{r} - r\dot{\varphi}^2 = -\frac{q}{m}\frac{k}{2}\left[\frac{R_m^2}{r} - r\right] \quad (1.13a)$$

$$\frac{d}{d_t}(r^2\dot{\varphi}) = 0 \quad (1.13b)$$

$$\ddot{z} = -\frac{9}{m}kz \quad (1.13c)$$

with initial conditions at the moment t = 0, motion in polar coordinates $(r, \varphi, z)_0$ is (r_0, φ_0, z_0).

According to equation 1.13, motion in the polar plane (r, φ) is completely independent of the motion along z. Also, φ could be excluded from equation 1.13a using equation 1.13b. Therefore, it is convenient to introduce separate energy characteristics Er, Eφ, Ez for each direction:

$$qE_r = \left({}^{m}/_{2}\right){r_0}^2 \tag{1.14}$$

$$qE_\varphi = \left({}^{m}/_{2}\right)(r_0\varphi_0)^2 \tag{1.15}$$

$$qEz = \left({}^{m}/_{2}\right){z_0}^2 \tag{1.16}$$

and the ion initial kinetic energy is the sum of all three.

The equation of motion along z, equation 1.13c describes a simple harmonic oscillator and its exact solution is

$$z_{(t)} = z_0\, cos(wt) + \sqrt{\frac{2E_Z}{k}}\, sin(wt) \tag{1.17}$$

Ions are injected into the orbitrap at a point (arrow) offset from its equator (z = 0) and perpendicular to the z-axis, where they begin coherent axial oscillations without the need for any further excitation. The oscillating ions will induce a current which is then Fourier transformed to individual frequencies. The frequency of the harmonic oscillation (ω) along the z-axis can be related to the *m/q* value of a given ion and is independent of the kinetic energy and spatial spread of ions, as shown by equation 1.18[92], where *k* is field curvature.

$$\omega = \sqrt{(q/m)k} \tag{1.18}$$

The potential advantages of the orbitrap include the following: i, high mass resolution (up to 100 000−200 000); ii, increased space charge capacity at higher masses due to independence

of trapping potential on mass-to-charge ratio and larger trapping volume in contrast to FTICR; iii, high mass accuracy, dynamic range, and upper mass limit.

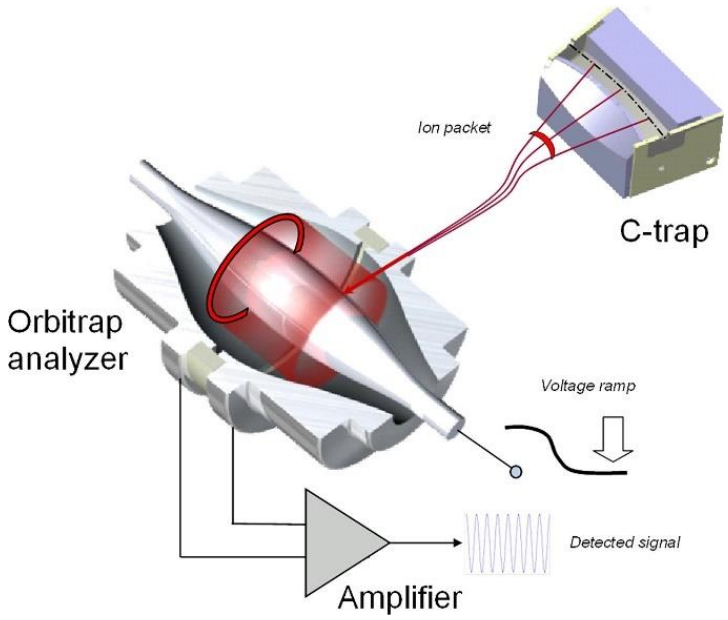


Figure 1.11 Cross-section of the C-trap and orbitrap analyser (ion optics and differential pumping not shown). Ions are stored in the rf-only bent quadrupole of the C-trap, then the RF is ramped down and a high-voltage pulse is applied across the trap, each m/z being ejected in a short packet. The packets from the C-trap enter the analyser during the voltage ramp and spread into oscillating rings that induce current detected by the differential amplifier. Figures is adapted from Thermo Fisher Scientific.

Orbitrap is usually combined with a C-trap design (**Figure 1.11**). The C-trap comprises four curved eletrodes of rectangular cross section. The rods are arranged evenly around a central bent axis. Since ion is produced by electrospray in ambient conditions and a number of this ion

species are injected into the C-trap. Ion of a wide m/z range are trapped simultaneously and accumulate by moving back and forth along the central bent axis. Once the RF is turned off and DC voltages are applied, ions fly out of the C-trap and are injected to the orbitrap together to enable mass detection.

1.2.4 Detectors

Ion current detectors play an important in measuring the intensity of ions with different m/z values that have been separated in a mass analyser. As in the case of ion sources and mass analysers, several types of detectors have been developed to improve their collection efficiency, signal-to-noise ratio, signal amplification, dynamic range and the frequency of the response. The detectors can be classified into two groups, one considering the charge of the ions whereas others are based on image charge detection or measurement of the kinetic energy that an ion transfers to an electronic plate. It is worth mentioning that orbitrap and FTICR detect ions based on the signal averaging of the oscillating ion packet.

Ions inside the orbitrap are detected via image current. As the ions oscillate back and forth according to their harmonic axial potential, they induce current on the outer electrodes. All of the ions inside the trap induce current on the electrodes simultaneously. As such, the raw image current is a sum of many different sine waves in the time domain. In order to convert this signal to the desired mass spectrum, it is broken into its various frequencies via Fourier transform, providing a frequency spectrum. Since each distinct m/z value has a distinct axial frequency,

the frequency spectrum is technically a mass spectrum. Since m/q and frequency are directly tied to each other, mass resolution (m /Δm) is directly proportional to the frequency resolution shown as:

$$\frac{\mathrm{m}}{\Delta \mathrm{m}} = \frac{\omega}{2\Delta\omega} = \frac{\sqrt{(\frac{kq}{m})}}{2\Delta\omega} \tag{1.19}$$

, where k is field curvature. The best m/q resolution is attained at low m/q ratio and then decreases with the square root of the ion's mass-to-charge ratio. This is preferable in comparison to FTICR, for which frequency drops linearly with mass-to-charge ratio[98].

1.2.5 Instrumentation Used for This Dissertation

Ultra-high mass range hybrid quadrupole-orbitrap mass spectrometer

Most of the native MS experiments performed in this dissertation were carried out on a Thermo Fisher Scientific Q-Exactive Ultra High Mass Range (UHMR) Hybrid Quadrupole-Orbitrap Mass Spectrometer (**Figure. 1.12**) equipped with an nESI source[99]. After ionisation, the analyte is directed to the cone and guided through a stacked ion guide (S-lenses) towards an injection flatapole that, compared to the previous versions of the instrument, has been upgraded to enable ion trapping. This modification converts the flatapole in a linear ion trap that allows fragmentation of trapped ions through CID[100], in which protein assemblies may be dissociated and the monomer species can be selected in the quadrupole according to their m/z[99]. In addition, this allows detergent removal before the quadrupole, thus allowing transmission of protein

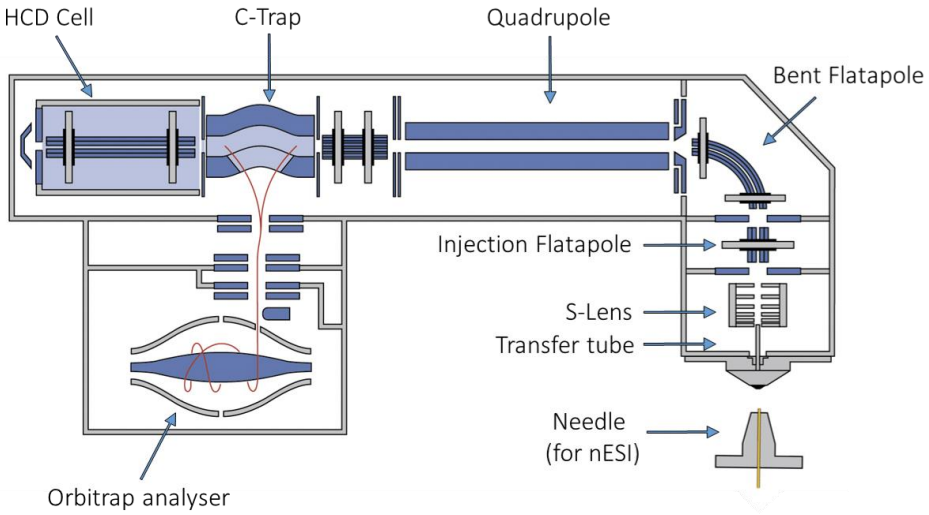


Figure 1.12 Schematic of the Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer. Ions produced from nESI are focused in the S-lens and transmitted to through the flatapole and the quadrupole which may operate in scanning or in mass-selection mode. Then ions enter the C-trap where they are focused and injected in the HCD for collisional activation, or directly in the Orbitrap. Figure adapted from Ben-Nissan, et al..

---

liberated from micelle in the quadrupole consenting selection of ligand-bound species for further HCD experiments. The injection flatapole is followed by a bent flatapole operating in RF-only mode that precludes the transmission of neutral species. The next step of the ion path is the first mass analyser represented by a quadrupole that allows the selection of ions for tandem MS analysis. Ions are then transmitted to the C-trap that allows storage before the injection in the higher-energy collisional dissociation (HCD) cell. Here, they undergo

collisional activation that, in the case of membrane protein complexes, allows the removal of detergent micelles or subunit and ligand dissociation. The current instrumentation enables ultra-high mass transmission up to 80,000 m/z[101]. Moreover, the higher pressure achievable in the HCD cell compared to previous instruments cell allows for improved ion trapping and activation[102]. The pressure of the ultra-high vacuum changes from the standard $1x10^{-10}$ mbar to ~ $1x10^{-9}$ mbar.

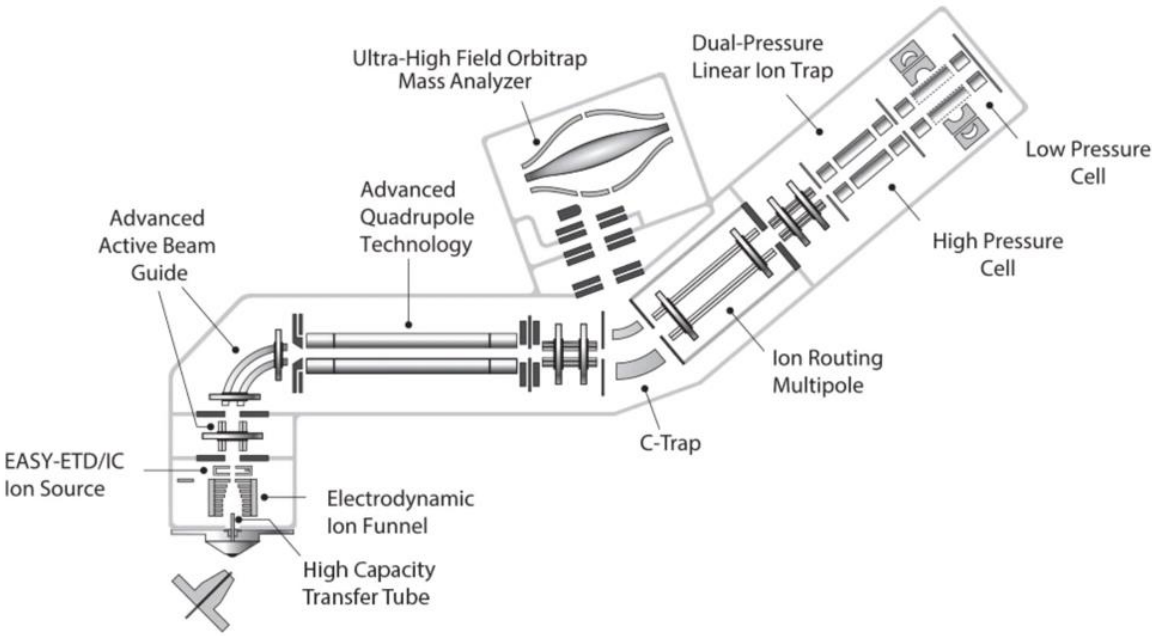


Figure 1.13 Schematic representation of the Orbitrap Eclipse Tribrid mass spectrometer, Figure adapted from Robittaille, et al.. Compared to QE, a secondary ion source set after the ion funnel and an additional linear ion trap is equipped

Orbitrap Eclipse™ Tribrid™ mass spectrometer

The Orbitrap Eclipse™ Tribrid™ Mass Spectrometer (Eclipse, **Figure 1.13**) was used for native MS^n, so-called nativeomics, proteomics and lipidomics analysis in this thesis. Eclipse is a powerful tool for studying global protein dynamics within a cell, tissue or organism. In addition to sharing a similar ion transmission route like QE, a secondary ion source set after the ion funnel delivers a regulated number of calibrant ions into the analyte ions[103]. This feature enables real-time fine adjustment of the instrument's m/z calibration, correcting otherwise uncompensated errors due to temperature fluctuations and scan-to-scan variations. An advanced linear ion trap equipped low pressure cell after HCD cell expands the mass range to 8,000 m/z realizing multiple MS/MS fragmentation in low-pressure cells.

1.3 Native Mass Spectrometry for the Elucidation of Structure and Function of Protein Assemblies

As discussed before, soft ionization of mass spectrometry opens avenue for the analysis of biological samples[74,81]. Further modifications such as elevated pressure at the front end of MS resulted in less dissociation with frequent low-energy collision[104] The use of low-frequency high-m/z quadrupoles and the high-resolution mass analyser of orbitrap enabled the detection of very large complexes[105]. Numerous methodological and instrumental advances on mass spectrometers enable further application in noncovalent macromolecule assemblies.

Around 30 years ago, the first folded protein sample were measured at physiologically favoured pH[106,107], later defined as native mass spectrometry (native MS). This offered a promising route for obtaining structural and physiological information of protein complexes. Since then, native MS has enabled the observation of large complexes up to the mega-Dalton size[108], including ribosomes[109] and intact viruses[110].

With further instrumental improvements, native MS realized the first observation of intact membrane protein aided by optimisation of solubilizing conditions for membrane protein (*E. coli* tetrameric ABC transporter BtuCD) aided by optimisation of membrane protein solubilization approaches[111]. The detergent micelle surrounding the membrane protein was gradually stripped off the protein through collision with inert gas within a mass spectrometer. The environment of membrane proteins analyzed by native MS has been expanded to nanodiscs and other membrane mimetics, described in the following subchapter (see review[112], for the interest in latest advances in native mass spectrometry of membrane proteins).

The latest advances in native MS employing orbitrap technology opens a new chapter of high sensitivity enabling single ion detection[113] and ever higher resolution[112,114-124] (see review of high-resolution mass spectrometry[125]). For instance, the advancement in resolution allows the observation of the fine details of post translational modifications on the protein[62,114,116,126] and the binding and stoichiometry of small ligands on an intact protein[114,115,117,124,127,128] as well as the influence of metal cation ions on ligand-protein interactions[112,122].

1.3.1 Advances in Membrane Protein Studies

Native MS has demonstrated its ability to determine subunit stoichiometry alongside other biological tools such as X-ray crystallography[129,130], cryo-EM[119] and mass photometry[131]. In particular, high-resolution native MS helped to solve controversies regarding membrane protein complex stoichiometry[132]. This shows its potential in assisting in solving ambiguous densities in X-ray crystallography or cryo-EM (see review by Pukala and Robinson[133]).

With the novel development of native MS methodology for membrane proteins, lipids have subsequently been defined that change protein conformation, mediate oligomeric states[134] and facilitate downstream coupling of G protein-coupled receptors[118]. Coupled with liquid chromatography-mass spectrometry (LC-MS)-based lipidomics, native MS also allows for the detailed analysis of low mass (700-1,500 Da) lipid entities[135].

Beside shedding light on stoichiometry and lipid mediation, native MS is also used in investigating molecular biological activities of protein complexes. Through exogenous addition of small molecules, native MS allows the evaluation of ligand binding on protein-protein interactions[115] or the enzymatic activity of membrane enzymes[136]. Until recently and also described in Chapters 3 and 4, native MS makes an initial step into a visualization of native biological activities in vitro[120], in which expanding the capabilities of membrane mimetics played a key role.

1.3.2 The Evolution and Progress of Membrane Mimetics for Membrane Proteins

Lipids-membrane interactions are essential for stability, structure, and function of membrane proteins[137-139]. Their predominant hydrophobic associations had long been considered unfavourable for the survival of subunit interactions. Additionally, membrane mimetics or high concentrations of detergent were overwhelming electrospray mass spectra and predicted to suppress protein signals[140]. These aspects, together with the inherently dynamic nature of protein-lipid interplay in cellular membranes[141] poses significant challenges for all structural biologists. Referring to challenges in crystallography and electron microscopy when aiming for extracting proteins from their heterogenous environment, this includes choice of detergent micelles, lipid cubic phase, or membrane mimetics[142].

Similar to crystallography and electron microscopy, selecting and optimizing an appropriate membrane mimetic is essential for native MS studies. However, native MS offers the benefit, i.e. label free, to study membrane protein−lipid interaction, while being able to distinguish lipid stoichiometry based on their masses and quantify binding of multiple species simultaneously. The ambition to further improve solubilisation conditions for membrane proteins and to enable detecting interactions has helped evolving membrane mimetics from detergent micelles, to lipid particles and to lipid vesicles over the past 15 years.

Detergent micelles

In 2008, native MS was first used to analyze membrane protein assemblies in detergent micelles[111]. The use of detergent micelles can extract and solubilise membrane proteins, which is a common approach in structural biology. For native MS, extraction and purification of membrane proteins is followed by exchanging protein−detergent complexes into volatile buffers with low amount of detergent at about two times critical micelle concentration (CMC). This allows for efficient desolvation and ionization when injecting the protein solutions into the mass spectrometer. After the entire protein−micelle complex ionized, detergent adducts need to be removed by activation inside the mass spectrometer at a good balance between adding enough activation for detergent removal and preserving membrane protein assemblies and other non-covalent interactions.

Over time, certain detergents being classical first choices, such as saccharide detergents including ndodecyl-β-D-maltopyranoside (DDM) and n-octyl-β-D-glucopyranoside (OG), which are also employed for crystallography. To improve the quality of native mass spectrum, charge-reducing detergents with polyethylene glycol hydrophilic head groups, such as tetraethylene glycol monooctyl ether (C8E4) and lauryldimethylamine-N-oxide (LDAO) are preferred as they are easier to remove. However, these detergents were not able to suit all membrane proteins as some membrane proteins are not stable when these detergents are present. To overcome this, a new library of modular oligoglycerol detergents (OGDs) for native MS of

membrane proteins were developed[143,144]. OGDs consist of a polar headgroup and a hydrophobic tail, which connect via a linker. Each section can be modified to optimize protein purification, retention of protein−lipid interactions, and charge reduction. OGDs improved purification and native MS analysis of functional neurotensin receptor type 1 (NTSR1), a GPCR instable in above mentioned "classical" detergents [145].

Recently, using detergent mixtures for native MS emerged as a promising strategy, by combining several detergent and optimizing the exact composition experimentally[146]. The presence of several detergents may allow membrane proteins to select the optimal detergent interactions within its microenvironment, which may help retain native lipids and stabilize challenging membrane proteins. Mixtures of OGD regioisomers generally improved the membrane protein isolation and quality of the mass spectra[147]. With more conventional detergents, NTSR1 was exchanged into mixed DDM, laurylmaltose neopentylglycol (LMNG), and fos-choline micelles while the adenosine A2A receptor (A2AR) and β1 adrenergic receptor (β1AR) were exchanged into mixed DDM and fos-choline micelles[115].

Nanodiscs

Nanodiscs are nanoscale lipid bilayers encircled by two copies of amphipathic membrane scaffold proteins[148].To constitute membrane proteins into nanodiscs, membrane proteins need to assemble with amphipathic membrane scaffold proteins and lipids. These complexes are purified by gel filtration to remove empty nanodiscs after solubilizing with detergent. One of

the advantages of the nanodisc is the construction of a lipid bilayer microenvironment for membranes protein to better preserve the stoichiometry and stability of membrane proteins compared with detergent micelles[149,150]. Due to their homogeneity and narrow size distribution, nanodiscs are particularly suitable for native MS[151,152] and they are resolvable in its intact form in the spectra[153].

Furthermore, nanodiscs can be engineered by using different MSP variants, to suit a wide range of membrane protein targets[149]. Moreover, the nanodisc lipid composition can be strictly controlled during assembly to investigate specific protein−lipid interactions[154,155]. Other than conventional nanodiscs described above, peptidiscs and picodiscs using amphipathic peptides and saposin A proteins as scaffolds respectively are emerging as potential membrane mimetics (see relevant review[156] and studies in peptidiscs[157] and picodiscs[158,159]).

SMALPs.

Although nanodiscs provide a more native lipid environment for membrane proteins than detergent micelles, working with nanodiscs still requires solubilising membrane proteins with detergents. Styrene maleic acid lipid particles (SMALPs) are membrane mimetics that can be used to directly solubilise membrane proteins from their native lipid membrane [160,161], resulting in membrane proteins in the natural lipid bilayer surrounded by SMA copolymer belt. However, both the SMALP and SMA-copolymer belt are highly heterogeneous, so it is difficult to obtain a resolved spectrum by activating SMALPs with conventional collision-induced dissociation.

Alternatively, membrane proteins may be resolved by being ejected from the SMALP with higher vacuum and activation energy[162] Alternatively, exchanging membrane proteins from SMALPs to the membrane mimetics acceptable for native MS such as amphipols or DDM[163], even though it is still hard to resolve the spectrum.

Sonicated lipid vesicles

Ten years after the first application of native MS in membrane proteins, a novel approach is emerging by using sonicated lipid vesicles (SoLVe), i.e. ejecting protein complexes directly

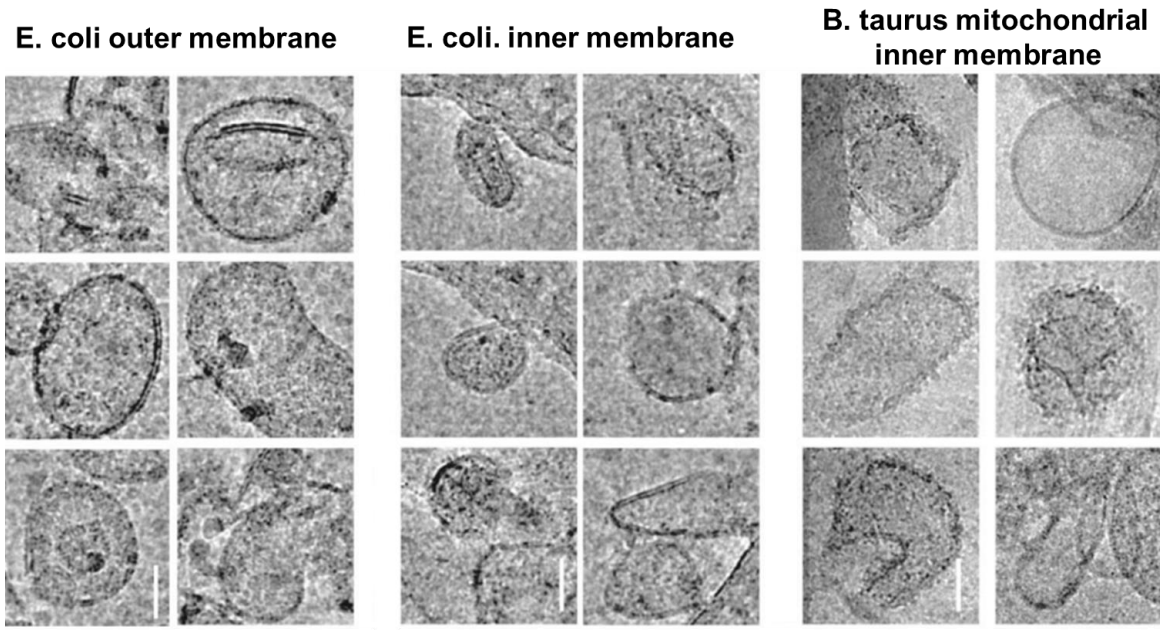


Figure 1.14 Electron Cryo-microscopy of vesicles before and after sonication. Inner and outer E. coli membranes and B. taurus mitochondrial inner membranes were sonicated and visualized by EM. Scale bar is 50 nm. Figure is from bv Chorev, et al..

from lipid membranes vesicles into mass spectrometers[117,128]. Following the collection of native membrane, membrane are sonicated into lipid vesicles, which is assumed to generate into lipid vesicles in which membrane proteins are embedded[117] **(Figure 1.14).** These lipid vesicles are then transferred to the mass spectrometer for analysis. In this process, the membrane proteins are radically ejected from membrane vesicles yielding resolved spectra which indicate detection of folded membrane proteins[128]. This approach is free of detergent, and therefore is likely to better preserve lipid-protein interactions which may help tackle challenging membrane proteins that are incompatible with detergent. This method was first applied to bacterial and mitochondrial complexes[117]. Very recently, this approach was further applied to allow real-time monitoring of the activation and transduction of rhodopsin[120]. Based on the insights obtained in these studies, SoLVe-native MS even allows for the investigation of multi-protein systems, when proteins of interest are present in sufficient abundance. Otherwise, the diversity and heterogeneity of protein components in a native membrane can generate data of high complexity. These challenging samples indicate further avenues to explore, such as spiking, i.e. addition of defined amounts of components to the sample, as well as data analysis supported by user-friendly software[164], modelling and artificial intelligence.

1.4 The Aim of the Dissertation

This dissertation mainly focuses on the application of native MS in structural and functional studies of membrane proteins and glycoproteins. The recent advances in biotechnologies like proteomics, lipidomics, glycomics, molecule simulation, crystallography and cryo-EM and a

variety of biochemical techniques have generated a large amount of discovery in biology and life science. With the advantage of native MS, we further probe the structure details of glycoproteins, the activation and signalling of a class A GPCR in the membrane as well as the structure and endogenous ligands of voltage-dependent potassium ion channel beta subunit. This dissertation pursues preservation of the native states of the targets and discovery of the structural and functional information of physiologically important proteins.

In Chapter 1, the principle of mass spectrometry is largely focus as well as the recent advances in biology using native mass spectrometry in the past decades. As the study of rhodopsin disc membranes is important in this thesis, a specific introduction is given in Chapter 2 — covering the special cell structure of rod and the formation of discs, rhodopsin structure and functions, the relevant studies of lipids along with its implication for GPCR studies.

In Chapter 3, we aim for establishing a platform with native MS for class A GPCR, rhodopsin in a native disc membrane and tracing its change in mass from dark to illuminated states. This finally shed the light on the significance of the interaction between rhodopsin and the membrane.

Later in Chapter 4, to further investigate the signal amplification of rhodopsin in such a native microenvironment and apply this established method in evaluating the insights of rhodopsin-target compounds on photoconversion and phototransduction.

Structure is a basis of studying functional activities. In Chapter 5, the aim is to develop a methodology specialised for glycoprotein characterization. It is a goal that the microheterogeneity and macroheterogeneity can be revealed in a native state. Therefore, this can be fundamental and find out the specific glycan structure vital for the interactions.

In addition, native MS has shown its advantages in studying small-molecule ligand-protein interactions. In Chapter 6, we hope to find out the endogenous ligands of Kvβ2, a voltage-dependent potassium channel beta subunit. Kvβ2 is an intermediate protein to communicate between mitochondrial peroxidation chain and potassium ion channel for sleep regulation. So the possible existence of endogenous substrates or inhibitors could inspire the drug discovery for sleep.

Lastly, Chapter 7 concludes the accomplishments of all research chapters, presents current challenges which are not only for native MS studies but also remain in the fields and provides outlook to future native MS studies.

Altogether, the native mass spectrum conserves the structural information of glycoproteins and protein assemblies, stimulus-responding activation and signalling of rhodopsin disc membrane as well as the identification of endogenous ligand. These present the latest advances in native MS application.

Chapter 2 Rhodopsin and Its Disc Membrane

2.1 Renewal of Rod Discs: Formation and Shedding

The rhodopsin-containing disc membrane is a specialized subcellular structure that stacks at outer segment of the vertebrate rod cells. This multilayered disc membrane is densely filled with visual pigment molecules that are essential to maximize light absorption. Rod outer segments connect to inner segments via a narrow cilium. This cilium separates rhodopsin disc membranes from the nucleus, mitochondria, endoplasmic reticulum (ER) and Golgi which locate in the inner segment.

Rhodopsin is synthesized in the endoplasmic reticulum and transported to the Golgi membranes where it is glycosylated. Then rhodopsin-containing vesicles move from the Golgi to the outer segment. But how these vesicles move from the inner segment to the outer segment and how multilayered rhodopsin discs are formed has remained controversial. There are two models under debate. One was proposed by Steinberg and his colleagues in 1980[165] wherein the rhodopsin vesicle fuses into the plasma membrane and new discs of all photoreceptors are formed from evaginations (open discs) of the plasma membrane at the outer segment base. The primary evidence for this proposal is the findings of open discs at the base of the outer segment reported by a series of electron microscope (EM) studies[166-168].

In contrast to the evagination model previously purposed by Steinberg[165], the other model of disc morphogenesis in rods purposed by Chuang et al.[169] In this model discs are formed upon vesicular fusion and transported via a ciliary process to the outer segment base. This was supported by EM images of outer segments completely enclosed by the plasma membrane but rarely in open discs. Lots of fusion proteins, most notably syntaxin 3[169], were particularly enriched in the basal outer segment region supporting this mechanism. Considering that discs have distinct protein and lipids contents from plasma membranes, it is difficult to rule out the possibility that vesicular trafficking (at least in part) participated in disc incorporation of rhodopsin as well as disc assembly.

These competing hypotheses were addressed in 2015[170-172]. It was demonstrated by Burgoyne et al.[170] that >90% of rhodopsin molecules were transported to the OS via the ciliary plasma membrane followed by fusion between the leading edges of neighboring evaginations. Before fusion, the leading edges of evaginations make PCDH21-containing contacts with the inner segment plasma membrane that may stabilize and control the evagination. This is in agreement with the finding by Volland et al[171] wherein rhodopsin is transported via the plasma membrane of the connecting cilium, but differently this process may be initiated by an invagination of the plasma membrane.

The renewal of rhodopsin discs requires shedding of old discs and formation of new discs. The exquisite studies into the renewal of rod discs were first carried out in 1960s by Young and colleagues[173,174]. In pulse-chase experiments with radioactive amino acids, a discrete band of

radioactive material that moved over time from the basal end of the outer segment. Eventually it was detached from the distal end of the outer segment. The detached outer segment fragments were found to become deformed, compacted, undergo chemical changes, and were displaced within retinal pigment epithelium (RPE). This intracellular process is known as photoreceptor phagocytosis[4]. This needs a variety of machinery components in the steps of recognition, engulfment and degradation[175]. Studies in mammals, fish, and amphibians demonstrated a highly conserved mechanism for ROS renewal[176-181]. It is also worth noting that disc membrane formation is a continuous process[173], but that detachment of groups of discs occurs intermittently. Moreover the occurrence of disc shedding was found to be modulated by light and temperature[181]. These discoveries suggested a possible circadian regulation of outer segment disc shedding. Such regular renewal of rhodopsin discs of outer segment is of profound significance to shed aged proteins, lipids and fluorophores for retinal health.

2.2 Organization of Rhodopsin in A Disc

The structure of a disc

Mature discs are a discal shaped round lumen acquiring a hairpin-like rim at the edge **(Figure 2.1**), which was observed by EM images[172] and atomic microscopy imaging (AFM)[182,183]. AFM and quantitative analysis of single ROS disc membranes revealed that the diameter of discs in a ROS can range from 0.7 - 1.9 μm in mice and humans[182,183]. Interestingly, as shown by antibody labeling[184], rhodopsin is distributed widely on the disc with an exception of the

hairpin-like edge. In contrast, the large intrinsic membrane protein, ATP-binding cassette, sub-family A, member 4 (ABCA4), peripherin and ROS membrane protein 1 (ROM1) are exclusively localized only to the rims and incisures of discs[167,184,185]. Thus the interior of the disc is specialized for rhodopsin transduction. Confined to the rim, peripherin and ROM1 are essential for maintenance of the curved bulbous rim of ROS discs[186-190]. Besides, ABCA4 residing in rims assists in retinaldehyde clearance from the ROS disc through its N-retinylidene-phosphatidylethanolamine flippase activity[191-193].

In the interior of the disc, rhodopsin is densely packed. Such a high density of receptor distribution is unique among GPCRs. The average density of rhodopsin is estimated to be about 20,000 molecules/μm^2 [194], which is orders of magnitude higher than that of other GPCRs, with estimates ranging from 5 to 30 molecules/μm^2 [195]. The organisation of rhodopsin was first unveiled by AFM images of an intact single bilayer disc membrane, which is a densely packed paracrystalline lattice of visual pigment[196]. The images have attracted the most attention to date and have provided the highest resolution of rhodopsin in native membranes. Subsequent AFM images consistently revealed a nanodomain organization of rhodopsin[197-200]. In all cases, rhodopsin forms an oligomeric complex composed of rows of dimeric receptor[38]. This high density of rhodopsin can be both beneficial and detrimental in facilitating the exquisite sensitivity exhibited by rod photoreceptor cells, which can respond to a single photon of light[201].

Rhodopsin oligomerization

Whether rhodopsin is monomer or oligomer (or dimer) is a long-lasting debate in the field. Rhodopsin dimerization or oligomerization has been proposed with respect to the significant amount of AFM images displaying rhodopsin organization in disc membrane. However this fails to address whether the observed oligomerization is due to receptor crowding, i.e. concentration-dependent association like other GPCRs[202] or artifactual association at below the physiological temperature. To look into the possibility of rhodopsin association at low concentration, subsequent fluorescence resonance energy transfer (FRET) measurements with fluorescently labeled rhodopsins using CY3-maleimide as the donor and CY5–maleimide as the acceptor showed that rhodopsin was able to self-associate into dimers even at a low receptor density and that they were functional for G protein transduction[203]. Another study further verified the dimerization of opsin in rods by forming dimers in the endoplasmic reticulum in transgenic mice[204]. Moreover recently, a structural study by cryo-EM revealed rhodopsin dimer in nanodiscs if crosslinking was applied[205]. This data provides new insights into dimer interface at helix 8 between two visual pigments distinct from the helix 4 - 5 model of rhodopsin oligomers deposited at the Protein Data Bank (code: 1N3M) based on data obtained from AFM images[197] and the crystal structure of bovine rhodopsin[206]. These *in vivo* and *in vitro* studies demonstrate the possibility that rhodopsin is able to self-associates at a specific interface, probably at helix 8.

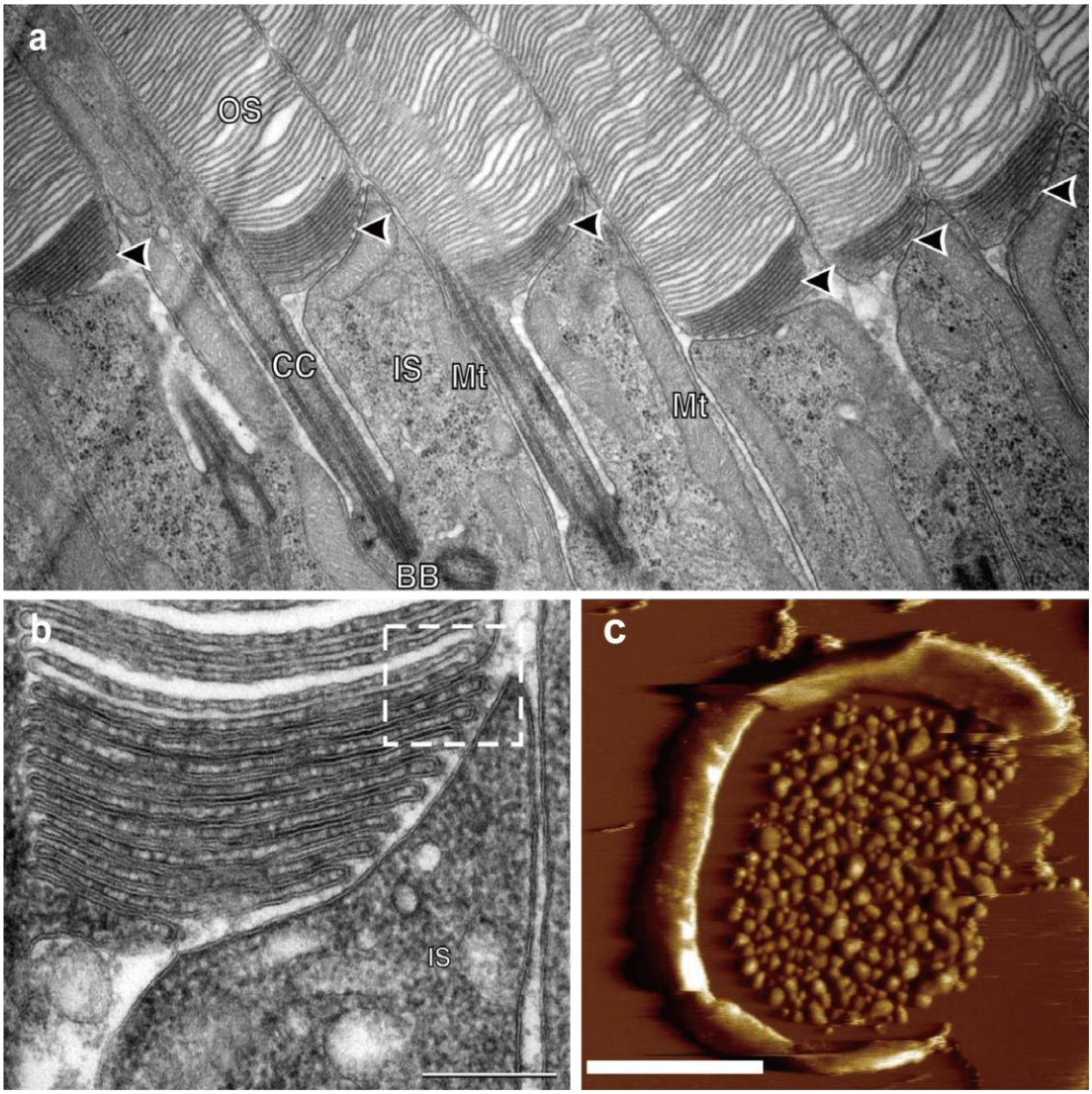


Figure 2.1 Discs at the rod outer segment. (**a**) An EM image of a mouse retina section at the interface between rod outer and inner segments stained with tannic acid–uranyl acetate. Arrowheads point at densely stained, tightly packed new discs at the outer segment base; note that mature discs are less densely stained and slightly swollen. (**b**) Transmission electron micrographs of photoreceptors and their cilia of mouse rods. (**c**) AFM imaging of ROS disc membranes. A deflection image of the same ROS disc membrane. Scale bar, 500 nm. Image **a** and **b** are adapted from Ding et al; **c** is from Whited et al

Rhodopsin functions as a monomer

Nevertheless, regarding downstream interactions, it is well demonstrated with numerous biophysical and biochemical studies[41,196,207-215] that monomeric rhodopsin is the minimal functional unit required for binding with transducin, arrestin, phosphodiesterase, G-protein-coupled receptor kinase 1 (GRK1). These research studies, led by cryo-EM, provide huge quantities of mechanistic details of rhodopsin interacting with these downstream auxiliary proteins at a stoichiometry 1:1, e.g. transducin[216,217], rhodopsin kinase GRK1[218], arrestin[219-221], phosphodiesterase[222]. All of these investigations came to a similar conclusion that monomeric rhodopsin is sufficient to couple with these downstream associated proteins and has an even higher affinity for downstream protein binding compared to oligomeric rhodopsin.

Functional monomeric rhodopsin and its oligomerization in discs enable the high sensitivity for photon detection

It is widely accepted that rhodopsin, like most class A GPCRs, can function as a monomer. In the meantime, it is also difficult to deny the tight arrangement of rhodopsin in a disc membrane and hard to rule out weak interactions between rhodopsin molecules[223]. However, research to date has tended to focus on the significance of either one rather than the connection between both. One possibility is that rhodopsin oligomerization in discs and its functioning as monomer contribute together to high sensitivity to signal light and to respond at speed[224]. First, in an area of less the 6 μm^2 [182,183], self-association of rhodopsin to high orders increases the concentration

up to 20,000 receptors/μm^2 [194]. This concentration is vital to capture photons in dim light and increases the possibility of transducin coupling with respect to the fusion of transducin on the membrane. Second, monomeric rhodopsin as a minimal functional unit, maximizes the number of signalling copies, and maybe frequency, in a certain surface area with regard to the given rho-G protein binding affinity and concentration of transducin in ROS. Evidence is reviewed from a wide range of studies relevant to sensitivity and the speed of rod phototransduction[207,224,225]. The special architecture of ROS and density of rhodopsin are essential for dim vision in respect of sensitivity of light absorption, photo-response speed and temporal resolution of the visual stimulus.

Furthermore, as the majority of studies pay high attention to understanding vision in the dim light environment, there is limited systematic research into visual response in ambient light — the relationship of desensitization after phototransduction and excessive receptor crowding. It would be interesting to see whether rhodopsin oligomerization plays a role in response to ambient light.

2.3 Rhodopsin Photoactivation

Rhodopsin is a seven transmembrane (7TM) protein. Rhodopsin is covalently bound with its chromophore, 11-*cis* retinal via a protonated Schiff base to the polypeptide chains of each opsin, which is exclusive compared to other GPCRs. It is unique among GPCRs in that its intermediates during photoactivation process can be captured and identified by spectroscopic

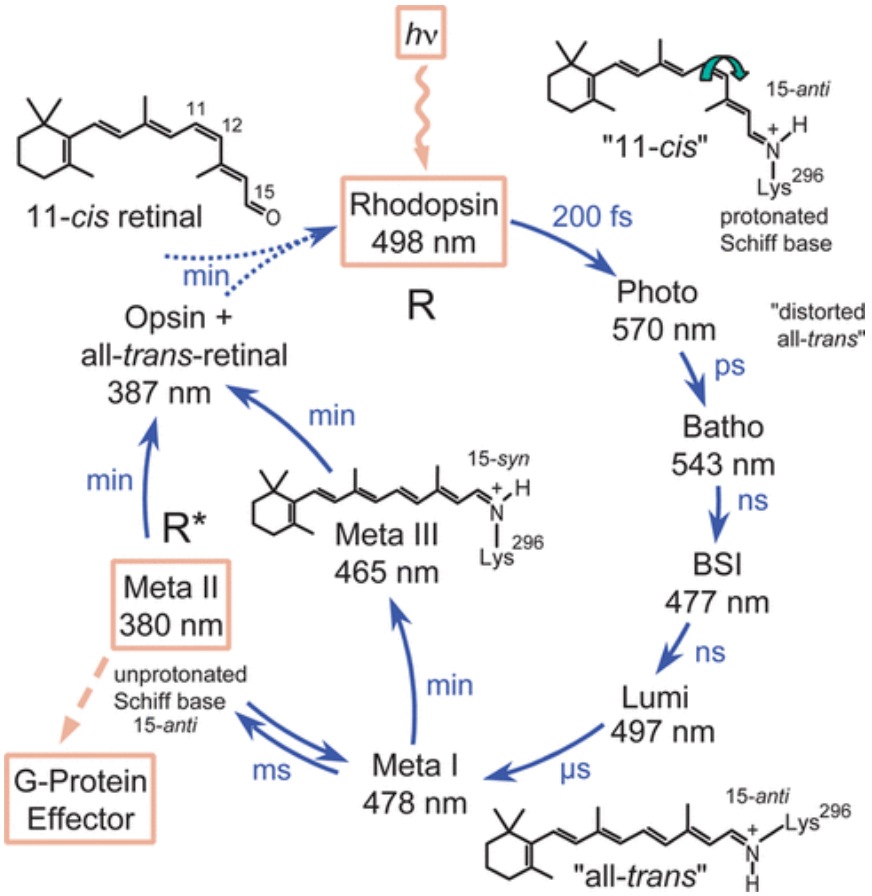


Figure 2.2 Spectroscopically detected intermediates of photoactivated bovine rhodopsin. Photoactivation of the rhodopsin with 11-*cis* retinal leads within femtoseconds to photorhodopsin with a highly distorted 11-*trans* bond. Via thermal relaxation, several intermediates form with distinct λ_{max} values. Gradual release of the strain in the chromophore leads through Batho and Lumi to Meta I, as seen by the different absorption maxima arising. Larger protein conformational changes lead to Meta II (λ_{max}=380 nm) which is in equilibrium with its predecessor Meta I. Meta I can also form Meta II. As a result of retinal Schiff base hydrolysis Meta II and Meta III decay to the apoprotein opsin and all-*trans*-retinal. Figure is adapted from a review written by O. P. Ernst, et al..

analyses. Within femtoseconds, photoisomerization of the retinal 11-cis double bond produces photorhodopsin, which has a severely distorted 11-trans bond[226]. Thermal relaxation causes a number of intermediates with different λmax values (**Figure 2.2**). These intermediates can be distinguished by low-temperature[227] or time-resolved spectroscopy[226,228]. The changing absorption maxima that result from variations in the chromophore/protein interaction show that the strain in the chromophore gradually releases and passes through Batho in 1 picosecond[226] and Lumi to reach Meta I. At low temperature, a transitory blue-shifted intermediate cannot be trapped. Time-resolved UV-vis measurements identified other transitory forms of Lumi[229,230] and Meta I[231-233]. The retinal Schiff base stays protonated through Meta I, most likely because of the low pKa of the stabilizing counterion Glu113[234].

In milliseconds, Metarhodopsin II (Meta II) generates from Meta I with larger protein conformational changes, which consists of the substates, Meta IIa and Meta IIb. Meta II is in equilibrium with Meta I (**Figure 2.2**). Meta II also refers to the active receptor state which is agonist-bound and able to catalyze the nucleotide exchange of GDP to GTP in the G protein transducin. Deprotonated retinal Schiff base in Meta II provide a significant blue-shifted value for λ_{max} at 380 nm (**Figure 2.2**). Later Meta II degrades to the apoprotein opsin and all-trans-retinal as a result of Schiff base hydrolysis.

Meta I can also transform into Meta III in minutes by the heat isomerization of the retinal Schiff base from all-trans,15-anti to all-trans,15-syn (λmax = 465 nm)[235]. Meta III can then decay to opsin and all-*trans* retinal, like Meta II, and can also photoconverted to Meta I and Meta II[236]

(not indicated in **Figure 2.2**). Since the reaction from Meta I to Meta II is much faster than the one from Meta I to Meta III, Meta II is considered to be the main state from Meta I to opsin[235].

To recapitulate, dark-state rhodopsin is inactive with an 11-*cis* retinal via protonated Schiff base. Once exposed to signalling photons, its chromophore 11-*cis* retinal immediately isomerises to all-*trans* retinal in the first microseconds while the retinal Schiff base still stays protonated, corresponding to a series of intermediates covering Photo, Batho, BSI, Lumi and Meta I. Then rhodopsin undergoes a large conformational change with all-*trans* retinal Schiff base deprotonated in milliseconds. It is capable of activating transducin (heterotrimeric G protein, G_t) and is recognized as Meta II. Afterwards, the deprotonated all-*trans* retinal of Meta II is hydrolysed into opsin and free all-*trans* retinal.

2.4 Rhodopsin Downstream Signalling

During photoactivation all-*trans*-retinal rho* interacts with GDP-bound transducin (G_{tGDP}) leading to GDP-GTP exchange, subsequent dissociation of G_t and formation of GTP-bound transducin α subunit ($G_t\alpha_{GTP}$)[237,238].

$G_t\alpha_{GTP}$ then interacts with phosphodiesterase 6 (PDE6) displacing the PDE6 γ subunit, releasing inhibition of catalysis and resulting in rapid hydrolysis of cGMP. The decreased level of cGMP causes the closure of the cation ion channel, which blocks the current of Ca^{2+} and Na^+ into the rod cell[239]. This amplifies the visual signal cascade that culminates in an electrical impulse to the visual cortex of the brain. The visual signal transduction in rods is terminated

by a series of molecular events that are initiated by the phosphorylation of rho* by GRK1 followed by binding of a regulatory protein, arrestin.

Activated rhodopsin (Meta II) is phosphorylated by GRK1 at multiple sites in its C-terminal region[240]. Each phosphorylation site provides an independent step in rhodopsin deactivation and collectively these steps precisely control rhodopsin's active lifetime[241]. Arrestin exists in a basal state that transitions into a binding-competent conformation by encounters with phosphorylated Meta II[242]. Arrestin in the basal state competes with rhodopsin kinase GRK1 and slows Meta II deactivation[241], where arrestin has a lower binding affinity to phosphorylated rhodopsin and arrestin binding is a rate-limiting step[243]. The mechanism of how arrestin is regulated by Calcium ions to terminate phosphorylated rhodopsin is well summarised in a review written by Zang et al[244] as well as a recent report[245]. The phosphorylation by GRK1 and arrestin binding takes place with Meta II. The phosphorylated Meta II will hydrolyse into phosphorylated opsin followed by dephosphorylation of the receptor and unbinding of arrestin[243,246]. It is important to note that phosphorylated receptors have a reduced catalytic activity of G_t.

2.5 Regeneration of Rhodopsin Chromophore

To sustain vision, opsin subsequently recruits the regenerative chromophore to reform rhodopsin. It is clearly established that at least two physiologic pathways[247] are involved in visual chromophore production — indirect or direct isomerisation of released all-*trans*-retinal

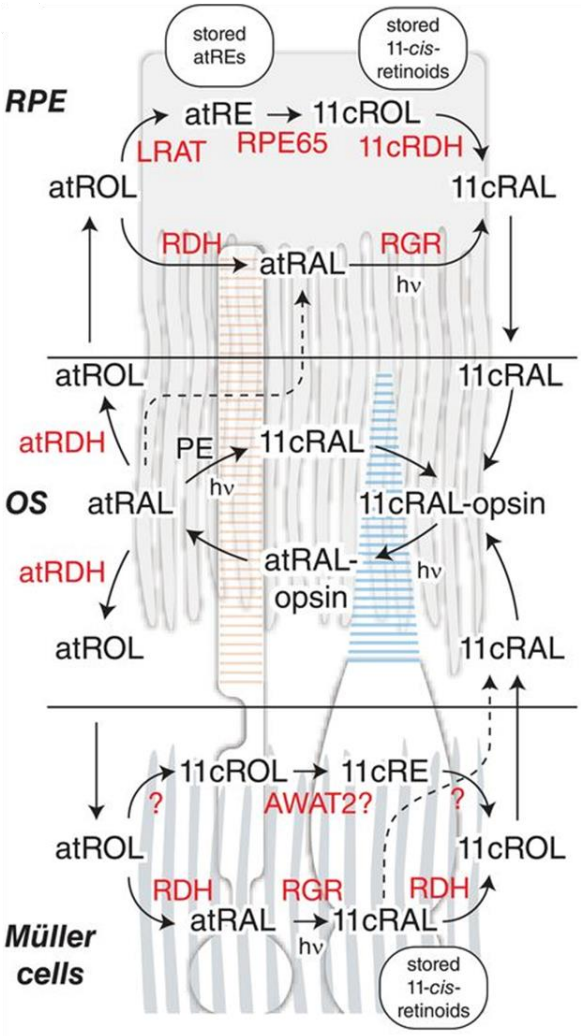


Figure 2.3 Pathways and enzymes for production of 11-*cis*-retinal in the vertebrate retina showing the location outer segments (OS) of rod (orange) and cone (blue), RPE cells (top), and Müller cells (bottom). Pathways for 11-*cis*-retinal production are indicated by arrows and the localization of enzymes in red text are shown. atRE and 11cRE represent all-*trans*- and 11-*cis*-retinyl esters; atROL and 11cROL are for all-*trans*- and 11-*cis*-retinol; atRAL and11cRAL stand for all-*trans*- and 11-*cis*-retinal, respectively. Figure is adapted from a review by Palczewski et al..

from visual pigments (**Figure 2.3**): one that involves the classical visual cycle with a key isomerase, RPE65, that uses retinyl esters to generate 11-*cis*-retinol followed by the transformation to 11-*cis*-retinal; and the second pathway involves a retinal G protein-coupled receptor photoisomerase (RGR) that uses all-*trans*-retinal as a substrate to directly generate *cis*-retinal. The 11-*cis*-retinal can regenerate in the RPE, Müller cells, or potentially within photoreceptor outer segment, through light-dependent and light-independent mechanisms.

The classical visual cycle occurs mainly in the RPE through an enzymatic pathway involving lecithin retinol acyltransferase (LRAT), retinoid isomerohydrolase (RPE65), and 11-*cis*-retinol dehydrogenases, most notably RDH5, are responsible for the consistent replenishment of the visual chromophore residing in the RPE[248]. Alternatively, 11-*cis*-retinal also forms directly in the RPE through a light-dependent pathway involving RGR[248-250]. RGR is a potential source of long wavelength-induced isomerase activity[251]. After forming all-*trans*-retinylidene of RGR, in contrast to opsin, all-*trans*-retinylidene is stable whereas 11-*cis*-retinylidene is hydrolysed and released from the active site. RGR is also expressed in the Müller cells where it carries out a similar function, but with the generated chromophore being selectivity channeled to cones[252]. Within the photoreceptor OS, all-*trans*-retinol dehydrogenases such as RDH8 reduce all-*trans*-retinal formed during photobleaching to all-*trans*-retinol, which is shuttled to adjacent cell layers[253,254].

However, alternate to these established mechanisms, a further pathway has been suggested - the so-called photic 11-*cis* retinal synthesis via N-ret-PE lipid conjugates[255-257]. In this pathway

it was proposed that blue light converts all-trans-retinal specifically to 11-*cis*-retinal through a nonenzymatic process utilizing a phospholipid intermediate in photoreceptor membranes. In these extensive studies, it was shown that blue (450 nm) light converts all-*trans*-ret-PE with high specificity to 11-*cis*-retinal through a retinyl-phospholipid (N-ret-PE) intermediate in the ROS photoreceptor membrane[257]. Interestingly, the quantum yield of this light-dependent conversion at light 450 nm was reported to be similar to the one of rhodopsin photoisomerization. But the quantum yield of the rhodopsin to bathorhodopsin transformation is largely wavelength-independent[258] and the photoisomerization of N-ret-PE was more effective using white light than light at the maximum absorption of this adduct at 450 nm[257]. These findings are interesting but were not well correlated to address how photoisomerisation of N-ret-PE plays a role in sustaining vision with regard to transient rhodopsin photoisomerisation. To clarify their physiological relevance, the production efficiency of *cis*-retinal via N-ret-PE requires more investigation under physiological conditions, e.g. white light. More chemical analysis in the formation, photoisomerisation and dissociation N-ret-PE would also help understand the physiological mechanism of N-ret-PE isomerisation.

2.6 Rhodopsin and Lipids

Lipid composition and distribution on a disc

The disc membrane has a different composition of lipids from rod plasma membranes. Rod disc membranes exhibited approximately equal amounts of PE and PC (42% and 45%,

respectively) while rod plasma membrane was much lower in PE (10%) and much higher in PC (65%). PS was also lower in discs (14%) compared with the plasma membrane (24%)[259,260]. The fatty acyl chains of the phospholipids likewise have a remarkable difference especially with respect to the enrichment of 22:6 (DHA) in disc membranes relative to plasma membranes[261-263].

These lipids are not distributed homogenously on a disc. Instead there are critical differences in lipid content between the centre and the rim[264]. Recently the native lipids from these two regions were identified by using co-immunopurification with styrene maleic anhydride (SMA) and following untargeted lipidomic analysis of membrane proteins from the rim and the centre. It was found that a lower PC to PE ratio and increased long chain- (LC-), i.e. more than 18 carbons, and very long chain- polyunsaturated fatty acids (VLC-PUFAs), i.e. more than 22 carbons with more than 2 carbon-carbon double bonds, appeared in the centre (rhodopsin) relative to the rim.

Lipid composition of disc membranes varies in age. The newly formed disc membranes evaginating from plasma membranes have 6-fold more cholesterol than aged discs at the apical tip of the ROS. The partitioning of cholesterol out of disc membranes as they age, and are apically replaced, is consistent with the high PE content of disc membranes relative to the plasma membrane[265,266].

The influence of lipids on rhodopsin signalling states

The lipid composition of membranes has a profound outcome on rhodopsin properties. It was found in model lipid membrane that different rhodopsin states are favored in different lipid compositions around rhodopsin[267-270]. It has been shown in several studies that Meta I−Meta II conformational change of rhodopsin is tightly controlled by negative-curvature elastic stress in membranes introduced by non-lamellar phase forming lipid components, DHA[269,270] and PE[270].

Lipid compositions can alter the signaling cascade by changing protein-protein interactions, e.g. cholesterol's inhibition of the photocycle[271-274]. In particular, cholesterol can selectively stabilize the dark state and in turn slow the kinetics of the photocycle as a result from cholesterol's stiffening effects on the membrane[273]. Complemented with experimental findings, a molecular-level mechanism based on molecular dynamic studies proposed that cholesterol stabilized the native dark-adapted state of rhodopsin without binding directly to the protein[274]. These studies in model membranes have demonstrated that cholesterol can modulate the activity of rhodopsin by altering the membrane hydrocarbon environment[275]. These studies also suggest that mature disc membranes initiate the visual signal cascade more effectively than the newly synthesized, cholesterol-rich discs.

In addition to lipid composition, lipid bilayer thickness also has an effect on rhodopsin states and its optimal function. The increase of bilayer thickness and helicity of rhodopsin are accompanied by higher Metarhodopsin II/Metarhodopsin I ratios and faster rates of Metarhodopsin II formation, which implied the possible adaptability of this G protein-coupled membrane receptor to properties of the lipid matrix[276]. Accordingly, DHA was proposed to

enhance rhodopsin's activity[277-280] and facilitate the conformational transitions that lead to the Meta II state[277,281].

It is also worth to note the allosteric interaction between lipids and zinc binding on rhodopsin. A recent study revealed simultaneous binding of lipids and zinc to rhodopsin[112]. Lipid binding of rhodopsin can be strongly and differentially influenced by cations. This suggests that the effects of metals and other cations may represent important modulators of lipid interactions.

Overall, the effects exerted by lipid on rhodopsin function as a whole are needed and have far-reaching implications for the study of rhodopsin. The function of a visual receptor can be modulated by lipid bilayer in its change of composition, thickness, curvature stress which could extend to other GPCR studies.

2.7 Rhodopsin, An Important Target for GPCR Studies

The studies of rhodopsin and the visual signal transduction system are important but not only for inherited retinal diseases including Leber congenital amaurosis, congenital night blindness, and retinitis pigmentosa caused by over 150 different mutations in rhodopsin[282]. Over and above that, these studies also contribute a wealth of information on the structure and function of GPCRs[23]. Visual signal transduction in rod cells has been studied extensively by biochemical and biophysical means[283-286], because of the relative ease of obtaining pure preparations of protein from bovine eye tissue and being the first expressed eukaryotic GPCR gene in mammalian cells. Therefore, bovine rhodopsin was the first GPCR to be conventionally

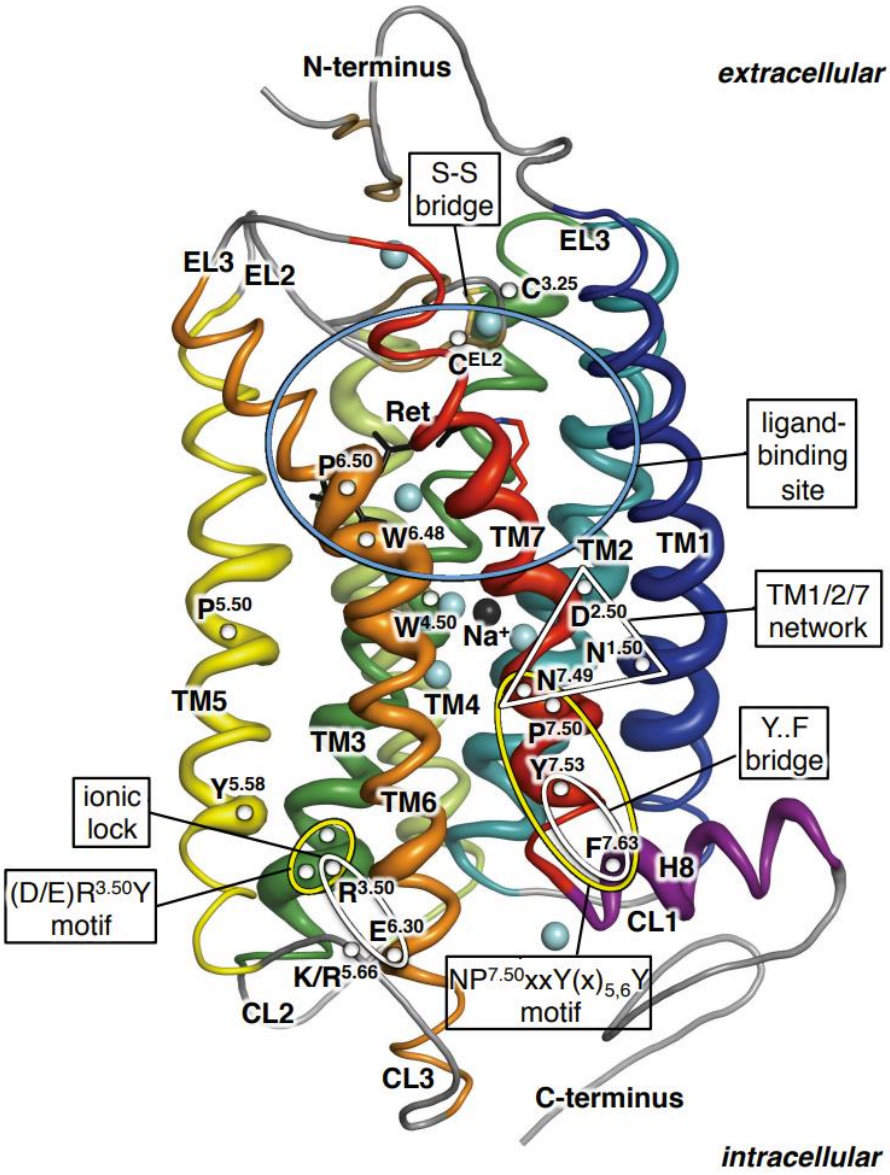


Figure 2.4 Sequence and motif conservation and ordered water molecules in GPCRs mapped on bovine rho (PDB ID:1U19) structure. Greater "Tube" thickness indicates greater residue conservation at that position. Ordered water molecules (shown in light blue) found by superposition of antagonist bound GPCR structures. Figure is adapted from a review written by O. P. Ernst, et al.

sequenced[287,288] and cloned[289]. Its sequence homology to the β2AR was observed in 1986[290], making rhodopsin the eponym of the class A GPCRs which comprises much of the ~800 GPCRs in the human genome[291].

A milestone in both rhodopsin and GPCR research was the elucidation of the crystal structure of bovine rhodopsin by Palczewski and colleagues in 2000[206] (**Figure 2.4**). This first GPCR structure served, and continues to serve, as a model to interpret much of the biochemical and biophysical studies of GPCRs. Comparison of inactive and active rho and GPCR structures indicate general functional roles for microdomains comprising residues of motifs conserved among rhodopsin-like GPCRs (**Figure 2.4**). Particularly, a central salt bridge between Arg135 of the (D/E)RY motif in TM3 and Glu247 on TM6 is known as the "ionic lock", and was initially proposed for the β2AR as a feature that stabilizes the inactive receptor conformation[292]. The model of rhodopsin also reflects results from GPCR structure analyses, where structurally bound waters are observed in "homologous" positions. This homology of water positions couples to the high degree of conservation seen in residues interacting with these waters, suggesting a functional role for water in the activation process[293,294]. The presence of a G-protein-interacting conformation in opsin or Metarhodopsin II structures was established by co-crystallization experiments with synthetic peptides representing one of the key interaction sites on rhodopsin's cognate heterotrimeric G protein, transducin (Gt)134-136. The human β2-adrenergic receptor (β2AR) in an active agonist-bound form, was also crystallized in complex

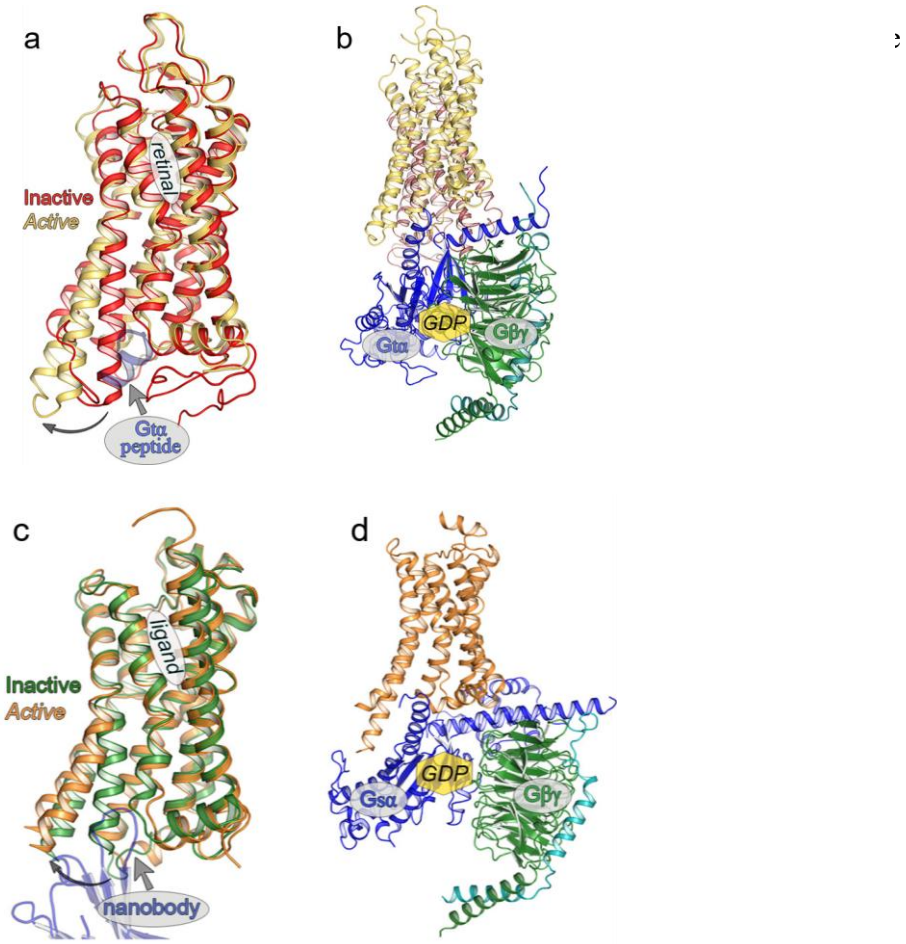


Figure 2.5 Structure changes of rhodopsin and other GPCRs upon attainment of the active state. (**a**) Structural rearrangement of inactive (red; PDB ID: 1U19) and G-protein-interacting active state (Meta II, yellow; PDB ID: 3PQR) of bovine rho. (**b**) Model of the complex between a rhodopsin dimer and transducin. (**c**) Structural superposition of antagonist bound (inactive state, green) and agonist bound (active state, orange) β2-adrenergic receptors (β2AR) with structural displacement of TM5 and TM6 similar to the ones in part A. (**d**) Crystal structure of an agonist bound β2AR-T4L fusion (T4L not shown) in complex with its cognate heterotrimeric G protein, Gsαβγ, and a stabilizing nanobody (not shown) (PDB ID: 3SN6). Figure is adapted from a review written by O. P. Ernst, et al.

with a nanobody binding in the cytoplasmic opening for the G protein because similar to these peptides derived from the distal C-terminus of the Gtα-subunit, the active conformation of structure of the complex between the β2AR and the G protein Gs, the binding mode for the Gsα C-terminus was similar to that of the Gtα-derived peptide bound to Metarhodopsin II[295]. The β2AR–Gs complex structure provided important insight into the interface and the relative orientation of the receptor and G protein and confirmed the role of the Gα C-terminal helix in coupling the Gα nucleotide binding site to the GPCR. GPCR-catalyzed nucleotide exchange in heterotrimeric G proteins is a complex multistep process which is not fully understood yet. For the rho/Gt pair, a high-resolution structure of the nucleotide-free complex remains elusive, but a model similar to the β2AR-Gs complex was built into the 3D envelope calculated from single particle analysis of negative-stained electron microscopy images of bovine rho–Gt complexes purified directly from the native source (**Figure 2.5b,d**)[296]. The proposed rho–Gt complex model, however, differs from the β2AR–Gs complex crystal structure by comprising one G protein and two rho molecules, with one protomer interacting with the active rho* conformation, although binding to Gt occurs in a similar orientation[296-298]. Dimerization of rhodopsin and opsin as well as of other GPCRs[299] can occur and affect receptor function and signaling allosterically[196,197,300].

Vertebrate rhodopsins thus can be seen as both highly specialized and photo-activated GPCRs which differ from diffusible ligand-activated GPCRs. Due to the common architecture found for all GPCR structures determined to date and the common gross conformational changes

observed in bovine rho and other GPCR structures upon activation[301], the knowledge of lipid-mediated rhodopsin signalling states from its conformational change to the G protein coupling form are likely to be directly applicable to the understanding of lipid-mediated activation and G-protein coupling of other GPCRs.

Chapter 3 Monitoring Photoconversion and Regeneration of Rhodopsin in Disc Membranes

Adapted from a publication in Nature, 2022

Siyun Chen[1,2], Tamar Getter[3], David Salom[3], Di Wu[1,2], Daniel Quetschlich[1,2], Dror S. Chorev[1*], Krzysztof Palczewski[3,4*], Carol V. Robinson[1,2*].

1. Chemical Research Laboratory, University of Oxford, South Path Parks Road Oxford, UK.
2. Kavli Institute for Nanoscience Discovery, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK.
3. Gavin Herbert Eye Institute, Department of Ophthalmology, University of California, Irvine, Irvine, CA, USA
4. Departments of Physiology and Biophysics, Chemistry, Molecular Biology and Biochemistry, University of California, Irvine, Irvine, CA, USA.

* Joint corresponding authors

Acknowledgements

All mass spectrometry research was funded by a Wellcome Trust Investigator Award (221795/Z/20/Z) and an ERC Advanced Grant ENABLE (695511) to C.V.R. A grant from National Institute of Health (R01EY030912) and unrestricted grants from Research to Prevent Blindness to the Departments of Ophthalmology at UCI are gratefully acknowledged by the K.P. laboratory. The authors acknowledge helpful discussions with Martin Galpin and Tarick El-Baba on the interpretation of kinetic data. For the purpose of Open Access, the authors have applied for a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Contributions

C.V.R. and K.P. conceived the idea through the CIFAR network and provided the overall supervision for the study. S.C. performed all the experiments described in this paper under the supervision of D.S.C. and D.W. for the SoLVe and lipidomics experiments respectively. D.Q. provided computational support for analysis and T.G., D.S. and K.P. prepared and provided ROS Disc membranes, identified rho targeting compounds, and prepared PDE6 and transducin complexes. C.V.R., S.C. and K.P. wrote the paper with contributions from other authors.

Abstract

Rhodopsin, a best-known G protein-coupled receptor (GPCR), is a photon sensor responsible for dim-light detection in mammalian visual system. Numerous studies show the effects exerted by lipids on rhodopsin function and suggest type of lipid types, chain length and the thickness of lipid bilayer are influential on rhodopsin activation, but influence of lipids on rhodopsin has not been investigated in native membrane. Here we use native mass spectrometry to probe rhodopsin, directly from the microenvironment of native disc membranes. We monitor real-time photoconversion of dark-adapted rhodopsin to opsin, delineating the steps of retinal isomerization from *cis* to *trans* and hydrolysis, further showing that the reaction is significantly slower in its native membrane than in detergent micelles. We also observe that opsin can regenerate rhodopsin in membranes through photoisomerised retinal-lipid conjugates from the native membrane. Our results not only reveal the impact of the native lipid environment on rhodopsin activation and regeneration but also establish a novel approach for other retinal-bound opsin activation studies in native lipid membrane environment.

3.1 Introduction

Rhodopsin is a visual receptor in eye retina and responsible for dim-light detection. Rhodopsin is densely packed and embedded in rod outer segment disc membrane (ROS). Each mammalian ROS consists of a 'disc-like' stack of 1000–2000 distinct disc membranes enclosed by the plasma membrane. Rhodopsin occupies approximately 50% of the disc membrane area, whereas the remaining space is filled with phospholipids and cholesterol, where lipid matrix and possible specific interactions with lipids stabilize activated rhodopsin. The photo-response of rhodopsin is monitored via metarhodopsin II decay by fluorescence studies[302] or absorbance spectroscopy[303-305] when solubilizing with detergent. In addition, non-typical rhodopsin regeneration pathway within rod cells is supported by a series of in vivo studies but the in vitro evidence is lacking to reveal the relationship between disc membrane and the cycle of rhodopsin. In this Chapter, we use native MS to enable the study of rhodopsin disc membrane focusing on the photoconversion of rhodopsin to opsin in native disc membrane as well as look into light- and disc membrane- dependent rhodopsin regeneration depending on the membrane.

3.2 Methods

Purification of bovine rod disc membrane and rhodopsin in LMNG

Rho-containing disc membranes were purified from dark-adapted retinas as described previously[188,306]. ROS washed with isotonic and hypotonic buffers were solubilized in 5 mM LMNG and rho was purified with immobilized 1D4 antibody as previously described[307], with

the modification that LMNG was used as detergent. Finally, the purified rho sample was dialysed (10 kDa cut-off) to remove the nonapeptide TETSQVAPA. All procedures were carried out in the dark under dim red light (>670 nm).

Preparation of disc membrane vesicles and detergent extraction of rho for native mass spectrometry

Collected purified disc membranes containing 81 µg rho were suspended in 250 µL of 200 mM ammonium acetate. To prepare vesicles we adapted an earlier protocol[117] and homogenized disc membranes using a probe sonicator with a stepped tip microtip (2 mm) (Vibra-Cell VCX-500 Watt, Sonics) using maximal amplitude at 40% with a circle of 1 s on and 2 s off applying 2 J per cycle for 1.5 min. Disc membrane vesicles containing rho (~9 µM) were ready for study via native MS either directly, or alternatively following hydroxylamine addition to a final concentration 5 mM. For detergent extraction, ROS was incubated in a hypotonic buffer with 20 mM LMNG overnight. The solution was then diluted into 50 µL 200 mM ammonium acetate with 0.004% LMNG and buffer exchanged 3 times with 200 mM ammonium acetate and 0.002% LMNG using a micro bio-spin® column (6,000 Da MW limit) prior to native MS. Rho (~9 µM) in the LMNG ammonium acetate solution was then subjected to MS, operated under the same instrument settings as for disc membrane vesicles.

Native mass spectrometry set-up and illumination conditions

The native membrane preparation (2-3 µl) was placed directly in a gold-coated nano-ES capillary prior to the source of the mass spectrometer (**Supplementary Figure 3.1a**). The cold LED light source (120 µW at 10 cm distance with a surface area of illumination 3 mm^2) was assembled via a bracket to illuminate the tip of the capillary for a defined interval in a dim red-light environment. For a real-time dynamic process, the membrane was first recorded in the dark and then illuminated with light pulses, or continuously, depending on the experiment. Native mass spectrometry experiments were carried out on a Q-Exactive adapted for membrane proteins [124]. The following parameters were used typically and adjusted to release the membrane protein from the detergents and membrane lipids: capillary voltage: 0.8-1.4 kV, HCD energy: 125-200 V, desolvation voltage: 0 V, source fragmentation: 100 V, HCD pressure: 3-8x10^{-10} mbar. C-trap entrance lens tune offset was set to 2, injection flatapole was set to 7 V, inter flatapole lens was at 0 V, and bent flatapole at 1 V. Threshold was set to 3. These MS conditions were modified to focus on the individual components of the spectrum. For example, spectra of rho/opsin (**Figure 3.2**) were collected at a capillary voltage: 1.1 kV, desolvation voltage: 0 V, source fragmentation: 100 V, HCD energy: 200 V. HCD pressure was typically 5.5x10^{-10} mbar.

Polyunsaturated PC lipids surrounding rho identified with multistage native MS^n

The mixture of rho, Gt and PDE6 at a molar ratio of 65:8:2 (described above) was used to identify the polyunsaturated PC surrounding rho during light induced signalling. PC lipids are released from rho and detected in positive mode using an Orbitrap Eclipse Tribrid Mass

Spectrometer (Thermo Fisher Scientific)[127]. In general, in-source activation (100 V) was applied to rho in membranes and the resulting ions transferred to the ion routing multipole (IRM) for activation (HCD NCE 100%, IRM pressure 8mM) to promote dissociation into lipids. AGC target values (100-150%) and maximum injection time (100 ms) were adjusted manually to maximize the normalized level. Detection at this stage was typically performed in the Orbitrap at a high m/z range of 500-8000. High resolution at 60,000 can be employed to enhance the intensity of lipid signals at low m/z. Data were analyzed using the Xcalibur software package v.2.2 (Thermo Fisher Scientific).

Native mass spectrum analysis and data processing

Raw data were first analysed manually using Xcalibur 2.2 (Thermo Fisher Scientific). The relative abundance of each species in a real-time measurement was quantified using Lig2Apo, a simple jupyter notebook. A folder of text files (format m/z versus intensity) is exported from Xcalibur 2.2 (Thermo Fischer Scientific) and two series of m/z values are defined corresponding to the molecular species of interest. The program then reads all txt files and calculates the relative intensity of the series and the intensity ratio of two species (see below for code availability). The zero charge spectra presented in **Figure 3.2** were analyzed by Unidec software 2.7.3[308].

Kinetic study of rhodopsin photoisomerization and hydrolysis

According to the kinetic model for the photoconversion and regeneration of rho in disc membranes (**Supplementary Figure 3.2**), a natural logarithm plot of rho abundance against time fits to a straight line in the period of 18 s to 300 s in disc membrane vesicles and 18 s to 180 s for the faster reaction in LMNG micelles. Because at the beginning of hydrolysis, the concentration of all-trans-retinal remains low, and the contribution from regeneration > 18 s is negligible in the system studied here, hydrolysis is the rate-limiting, and dominant reaction, such that it can be considered as irreversible. *k1* and *k2* are referred to as *kiso* and *k'hyd* in this phase respectively.

Experimentally, we first independently obtained the rate constant *kiso*. This rate constant is derived from the experiment in which *cis*-retinal rho isomerisation is measured after the addition of hydroxylamine and is represented in our kinetic model as the isomerization of *cis*-retinal rho to rho* at a rate *kiso* including a minimal contribution of regeneration.

Since this is a first-order reaction, the rate can be expressed as follows:

$$\frac{d[cis-\text{retinal rho}]}{dt} = -kiso \times [cis-\text{retinal rho}] \quad (3.1)$$

where [*cis*-retinal rho] is the concentration of *cis*-retinal rho and t is the illumination time. The solution can be written as follows.

$$[cis-\text{retinal rho}] = [cis-\text{retinal rho}]_1 \, e^{-kiso\times (t-t_1)} \quad (3.2)$$

where t_1 is the initial time point of this data subset and [*cis*-retinal rho]$_1$ is the concentration of *cis*-retinal rho at t=t_1.

Equation 3.2 can be used to fit the experimental data of the rho photoconversion in the presence of hydroxylamine. In this case rho only consists of *cis*-retinal rho, and excludes rho*. Plotting the abundance of *cis*-retinal rho as a function of time gives the fitting parameter in equation 3.2: *kiso*.

Next, under normal photoconversion conditions, the intermediate rho* is isomerized from cis-retinal rho at a rate kiso and slowly hydrolyses (k'hyd) into opsin and all-trans-retinal at a rate expressed as:

$$\frac{d[\text{rho}*]}{dt} = kiso \times [cis - \text{retinal rho}] - k'hyd \times [\text{rho}*] \quad (3.3)$$

where [rho*] is the abundance of all-*trans*-retinylidene-opsin, t is the illumination time. substituting equation 3.2 into 3.3, yields a differential equation with the solution:

$$[\text{rho}^*] = \frac{kiso}{k'hyd - kiso} \times [e^{-kiso \times (t - t_1)} - e^{-k'hyd \times (t - t_1)}] \times [cis - \text{retinal rho}]_1 + [\text{rho}^*]_1 \quad (3.4)$$

,where t_1 is the initial time point of this data subset and [rho*]$_1$ is the abundance of all-*trans*-retinal rho at t=t_1. [rho*] is obtained from the difference between the measured abundance of rho under normal conditions and the abundance of *cis*-retinal rho from equation 3.2. The abundance of rho* as a function of time. *kiso* can be determined independently using equation 3.2. The fitting parameters in equation 3.4 gives *k'hyd*.

Finally, in this kinetic model, [rho] can be expressed as

$$[\text{rho}] = \frac{k'hyd}{k'hyd - kiso} \times e^{-kiso \times t} - [cis - \text{retinal rho}]_0 \times \frac{kiso}{k'hyd - kiso} e^{-k'hyd \times t} \quad (3.5)$$

Since at all times the total abundance of *cis*-retinal rho, rho* and opsin as [cis-retinal rho]+[rho*]+[ops]=[tot], [ops] can be expressed as

$$[\text{ops}] = \frac{kiso \times e^{-k'hyd \times t} - k'hyd \times e^{-kiso \times t}}{k'hyd - kiso} \times [cis - \text{retinal rho}]_0 + [\text{ops}]_0 \quad (3.6)$$

Since *kiso* >> *k'hyd*, equation 3.5 and equation 3.6 can be simplified as follows,

$$[\text{rho}] = [cis - \text{retinal rho}]_0 \times e^{-khyd \times \text{t}} \quad (3.7)$$

$$[\text{ops}] = [1 - \text{e}^{-k'hyd \times \text{t}}][cis - \text{retinal rho}]_0 + [\text{tot}] - [cis - \text{retinal rho}]_0 \quad (3.8)$$

t is illumination time. [rho] is the measured abundance of rho under normal experimental conditions (without the addition of hydroxylamine). [ops] is the measured abundance of opsin under normal conditions. [*cis*-retinal rho]$_0$ is the abundance of *cis*-retinal rho from the dark-adapted membrane and regenerated from opsin before this phase. Total abundance of *cis*-retinal rho, rho* and opsin as [tot] =100% where relative abundance is used throughout. Since [*cis*-retinal rho]$_0$ is difficult to measure or estimate in a single measurement we use [rho] as a function of time and the fitting parameter is *khyd* along with [*cis*-retinal rho]$_0$.

Thus, under normal conditions, *khyd* can be determined first from this phase using equation 3.7 or equation 3.8. With the additional independent data set in presence of hydroxylamine, *kiso*

can be determined independently using equation 3.2. Furthermore, the rate constant of hydrolysis literally from rho*, *k'hyd* is refined using equation 3.4.

Notes: isomerization of cis-retinal rho to rho* goes to completion as hydrolysis takes place and the concentration of all-*trans*- retinal increases. In our kinetic model (**Supplementary Figure 2d** above), hydrolysis of the rho* Schiff base is a reversible second order reaction.

Additionally, from about 420 s onwards, the concentration of opsin was not observed to reach equilibrium but rather continues to rise slowly. A plausible explanation is that as hydrolysis proceeds slowly, all-*trans*-retinal is slowly consumed via regenerative isomerization to cis-retinal and combination with opsin.

Lipidomics and N-Ret-PE identification

N-ret-PEs and lipids were extracted by 90% isopropanol with 0.1% formic acid from the disc membrane vesicles in dark, illuminated for 1 s and 4 s, respectively. The supernatant was transferred into a new glass tube and dried using a SpeedVac vacuum concentrator (Thermo Fisher Scientific). The evaporated lipid mixture was dissolved in a 50 µL buffer and sonicated for 10 min. For LC-MS/MS analysis, lipids were loaded onto a C8 column (Acclaim PepMap 100, C8, inner diameter 75 µm, particle size 3 µm, length 150 mm, Thermo Fisher Scientific) using a Dionex UltiMate 3000 RSLC nano System connected to an Eclipse Tribrid Orbitrap mass spectrometer (Thermo Scientific Scientific)[103]. A binary buffer system was used with buffer A of acetonitrile: H_2O (60:40), 10 mM ammonium formate, 0.1% formic acid and buffer

B of isopropanol: acetonitrile (90:10), 10 mM ammonium formate, 0.1% formic acid. Lipids were separated at 40 °C with a gradient from 30% to 99% buffer B at a flow rate of 300 nl/min over 30 min. The electrospray voltage was set to 2.2 kV with funnel RF level at 40 and heated capillary temperature at 320 °C. For data-dependent acquisition, full MS mass range was set to 300 to 2000 with a resolution of 120000 and AGC target of 100%. Fragment spectra were acquired in the Orbitrap with a resolution of 15000 using higher-energy collisional dissociation (HCD) with stepped collision energy of 25, 30 and 35%. Phospholipids were detected in negative ion mode. The raw data were processed by LipiDex[309] and MZmine 2[310] for phospholipid identification and quantification). N-Ret-PE identification was processed manually. The extracted ion chromatogram of ret-PE was integrated using Xcalibur 4.1 (Thermo Fisher Scientific) and the area under the curve was used for quantification.

Data availability

The data that support the findings of this study are available and deposited to Figshare as follows: native MS and proteomics data: doi.org/10.25446/oxford.16901326 and lipidomics data: doi.org/10.6084/m9.figshare.18319361.v1, respectively.

Code Availability

The script for calculating the ratio of opsin to rhodopsin ratio in mass spectra can be accessed under https://github.com/d-que/Lig2Apo.

3.3 Results and Discussion

To establish MS and light conditions that enable monitoring of rhodopsin signalling, and before instituting our approach in native membranes, we used membrane solubilization and detergents, to extract rhodopsin from dark-adapted ROS membrane (**Figure 3.1**), as described previously for analysis via MS[311-313]. Following extraction into lauryl maltose neopentyl glycol (LMNG) we identified illumination conditions (LED emitting cold white light, **Supplementary Figure 3.1**) and MS conditions that enable us to release rho from membrane lipids and detergents (**Methods**) and monitor rho* signalling through the mass change associated with conversion of rho to opsin (Δ mass = 266 Da). Prior to illumination, 73 ± 1% rho (11-*cis*- or *trans*-retinal) and 27 ± 1% opsin are present. Recording of spectra in real-time was synchronized with illumination; equal population of both species was achieved after 3 min (**Figure 3.2a**) reaching a plateau with 27 ± 1% rho remaining at 20 min in this detergent micelle preparation (**Figure 3.2c**). We next established MS conditions that enable ejection of rho/opsin directly from disc membrane fragments while activated with light. To do this we adapted and applied a sonicated lipid vesicle approach[117] to purified ROS disc membranes forming vesicles for direct introduction into the MS. Under these conditions monomeric rho was ejected. No dimeric population was observed, attributed to the relatively small dimer interface observed in cryo-EM in nanodiscs[205] that dissociates under these MS conditions. Comparing the post-translational modification (PTM) status of dark-adapted and light-exposed rho/opsin extracted

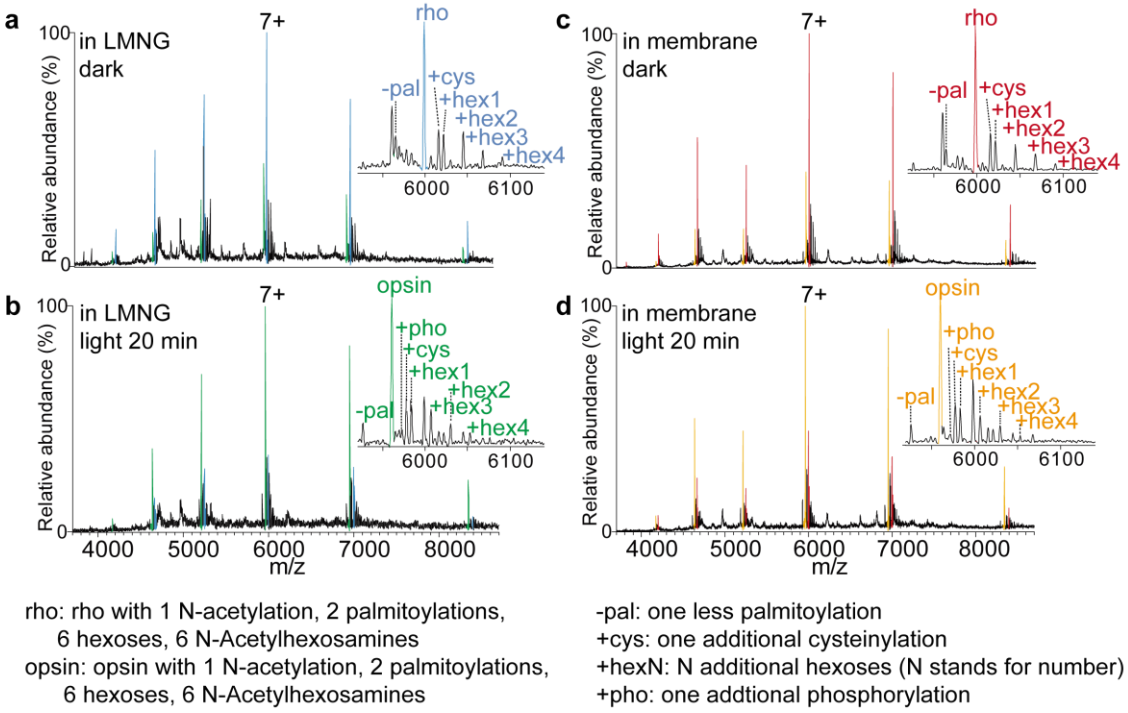


Figure 3.1 Native mass spectra of rho extracted in detergent or ejected directly from membranes, either dark-adapted or following 20 min of illumination. (**a**) Extracted from dark-adapted membranes into LMNG 73% rho (blue) and 27% opsin (green) are observed initially consistent with a small population of opsin in dark-adapted bovine retinae. Inset: Expansion of the 7 + charge state reveals glycans (hexoses: hex1, hex2, hex3 and hex4) cysteinylation (cys) and palmitoylation (pal) occupancy. ~95% rho /opsin has two pal sites occupied, while 5% have only one pal. (**b**) After 20 min of exposure to light the population of opsin has increased and ~5% of opsin is observed with at least one phosphorylation (pho). All other PTMs remain unchanged, expansion of the 7+ charge state (inset). (**c**) Rho/opsin expelled from membranes reveals 73% rho and 27% opsin, as in the detergent extraction, and an identical pattern of PTMs is observed for the 7+ charge state (inset). (**d**) After 20 min exposure to light in the membrane the majority of the rho population has converted to opsin. 5% of opsin is

phosphorylated while all other PTMs remain unchanged (7+ charge state, inset). The experiment was repeated three times independently (relevant statistics see **Supplementary Table 3.1**).

---

in detergent or ejected from membranes we find that glycosylation, palmitoylation and cysteinylation status are largely indistinguishable (**Supplementary Table 3.1**). An increase in phosphorylation of opsin was observed 20 min after signalling was initiated, both in detergent extracted and membrane ejected rho/opsin, consistent with progress towards termination of this signalling pathway through eventual interaction with arrestin[219].

Monitoring rho/opsin populations during continuous light exposure in membrane vesicle fragments we found that after 3 min rho remained predominant over opsin (**Figure 3.2b**). The reaction in membranes proceeds until a plateau is reached with ~32% rho remaining after 20 min (c.f. ~27% in detergent micelles after the same time period in **Figure 3.2c**). We next deduced and compared the rates of chromophore hydrolysis (k_{hyd} at pH 7.0 and 28 °C) in LMNG micelles and from ROS membrane vesicles according to a kinetic model (**Figure 3.2c** and **Methods** equation 3.7). Since the hydrolysis of all-*trans*-retinal from photoactivated rho is a slower step than photoisomerization of 11 *cis*-retinal to *trans*-retinal in rho, our results allow us to conclude that hydrolysis from rho to opsin is significantly slower in vesicles than in detergent micelles. Considering the isomerisation reaction, loss of all-*trans*-retinal from the binding pocket of rho is monitored through a change in mass following hydrolysis of the Schiff

base. The extent of *cis* or *trans* retinal bound to opsin prior to hydrolysis, however, is not apparent via mass alone. To delineate these isomeric populations, we adopted a well-established protocol in which ROS membrane preparations are pre-incubated in the dark with 5 mM hydroxylamine (**Figure 3.2e**). Rho in its ground-state is insensitive to hydroxylamine but, upon photoisomerization of *cis*-retinal to all-*trans*-retinal, hydroxylamine is able to attack and cleave the Schiff base in the photoactivated rhodopsin to form all-*trans*-retinyl oxime[314]. This hydroxylamine experiment, together with the rho/opsin photoconversion of untreated membranes above (k_{hyd}), enable us to subtract the two decay curves and consequently obtain the rates of retinal isomerisation (k_{iso}) and hydrolysis without isomerisation (k'_{hyd}) in native membranes (**Figure 3.2e** and **Supplementary Figure 3.3**). When photoactivated, rho* in membranes treated with hydroxylamine decays significantly faster than in the absence of pre-treatment. Careful examination of the data reveals an initial phase in which opsin regenerates rho (<18 s) in competition with the overall rho/opsin conversion. The increase in *cis*-retinal rho at 2.7 ± 1 s is coincident with the lowest level of opsin at 3.6 ± 1 s (Figure 2e insert). This initial reaction, in the absence of hydroxylamine, must therefore arise primarily from *cis*-retinal binding to opsin. To determine if this observation was a feature of the membrane, we applied short-term illumination on disc membrane preparation and to rho extracted into LMNG detergent micelles. Upon each transient illumination of rho from ROS, a rapid rise of rho was observed together with the general trend of rho transforming to opsin. Some evidence for regeneration of rho, when extracted into micelles, was also apparent although we were unable

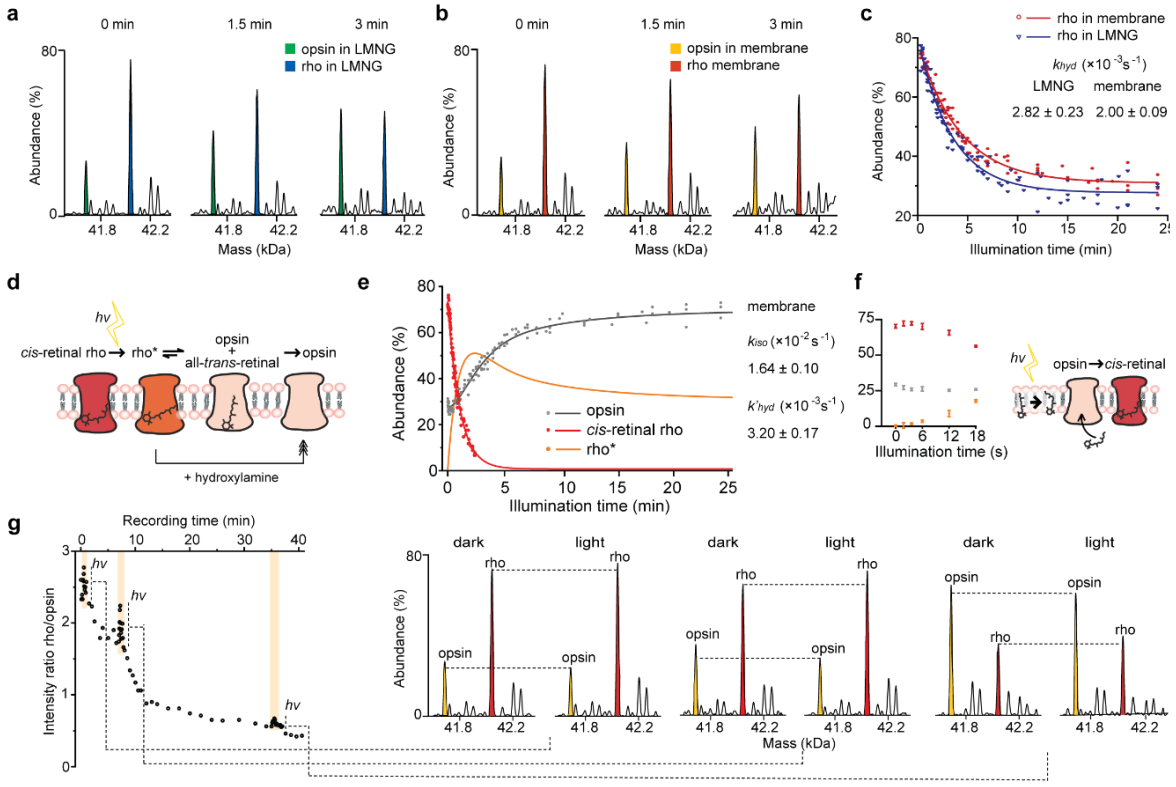


Figure 3.2. Monitoring the real-time conversion of rho to opsin in detergent micelles and native membranes, probing transient regeneration of rhodopsin at the cost of opsin. (**a**) Changes in the population of rho and opsin in LMNG blue/green and (**b**) in native membranes red/yellow respectively. Individual spectra are shown as zero charge plots at the times stated. (**c**) Plot of the relative abundance of rho during illumination, either in LMNG detergent micelles (blue) or native membranes (red) monitored as a change in mass as a function of time. The rate constant for hydrolysis is shown for the reaction in detergent micelles or in membranes. Points show data from three independent experiments. Data are fitted according to a kinetic model (**Supplementary Figure 3.2 and Methods**, R^2=0.94 ,0.98, 0.98 from three repeats in disc membrane vesicles and R^2=0.95, 0.96, 0.99 in LMNG micelles) and presented

as mean ± s.e. ($n = 3$). P = 0.0294, P<0.05; two-tailed unpaired Student's t-test is applied. (**d**) Schematic of the dark-adapted state undergoing light-activated conversion of *cis*- to all-*trans*-retinal, hydrolysis of the Schiff base, and dissociation from rho to form opsin. (**e**) Monitoring photoconversion of rho, rho* to opsin in normal membrane (grey) and in hydroxylamine pre-treated membranes (red). (**f**) Early time points (< 18 s) an increase in rho abundance is observed at the expense of opsins. The schematic depicts a possible regeneration mechanism through isomerisation of *trans*-retinal. Data of three repeats are ploted and presented as mean ± s.e. ($n = 3$). (**g**) monitoring the conversion of rho to opsin and the regenerative reaction of opsin to rho following pulses of light at three discrete time points 0.51 min (i), 7 min (ii) and 35 min (iii) for a time period of 3, 12 and 32 s, respectively. Zero-charge mass spectra are shown at time points taken during this conversion reaction, prior to illumination (dark) and at the peak of the three light pulses. The experiments were repeated three times independently.

---

to observe this phenomenon to the same extent as in the membrane and in green light (**Supplementary Figure 3.4 and Supplementary Figure 3.5**). The ability to regenerate is a known property of the visual system which requires the conversion of all-*trans*-retinal to 11-*cis*-retinal and has been shown to be carried out by enzyme pathways in neighboring cells[247].

Our proteomics experiments (**Supplementary Table 3.2**) were unable to detect key enzymes from the regeneration cycle, such as another retinal-binding GPCR (RGR) or the retinoid isomerase RPE65[315], prompting us to consider the possibility that this rapid regeneration (< 18

s), maybe a feature of the membrane itself in the light of preferred wavelength range of light . Previous reports have implicated an N-all-*trans*-retinyl-PE conjugate that can undergo photoisomerisation to form *cis*-retinals (primarily 11 *cis*-retinal (85-86%) with small contributions from 9 *cis*-retinal (12-13%) and 13 *cis*-retinal (1-2%) in blue light[257]. While our MS approach cannot readily distinguish these structural isomers our lipidomics data reveal the presence of high concentrations of polyunsaturated fatty acids (PUFA) including PE (**Figure 3.3a-c**). To explore formation of retinal conjugated to PE we extracted lipids into isopropanol directly from dark and controlled-light exposed ROS disc membranes. We selected three representative lipids (PE 18:0/18:1, 18:0/20:4 and 18:0/22:5) with minimal overlap in spectra, and searched for their respective retinyl conjugation. We found direct evidence for the formation of N-retinyl-PE in all cases and monitored its change as a function of exposure to light. After 1 s of photo-activation a significant decrease in conjugation was observed (>2.5-fold for PE 18:0/18:1), consistent with light stimulated retinal dissociation from PE (**Figure 3.2g** and **Figure 3.3d-f**). Peaks assigned to N-retinyl-PE increased after 4 s of light exposure to similar levels as the dark-adapted membrane suggesting further conjugation events. A significantly reduced initial phase was observed in the presence of micelles, wherein a lack of the conjugation system for retinal and PE would be anticipated (**Supplementary Figure 3.4**). Moreover, regeneration might be expected to be limited by the availability of *cis*-retinal, to replenish the binding pocket in opsin. To investigate this possibility, we incubated intact ROS disc membranes with a ten-fold excess of all-*trans*-retinal. Following illumination,

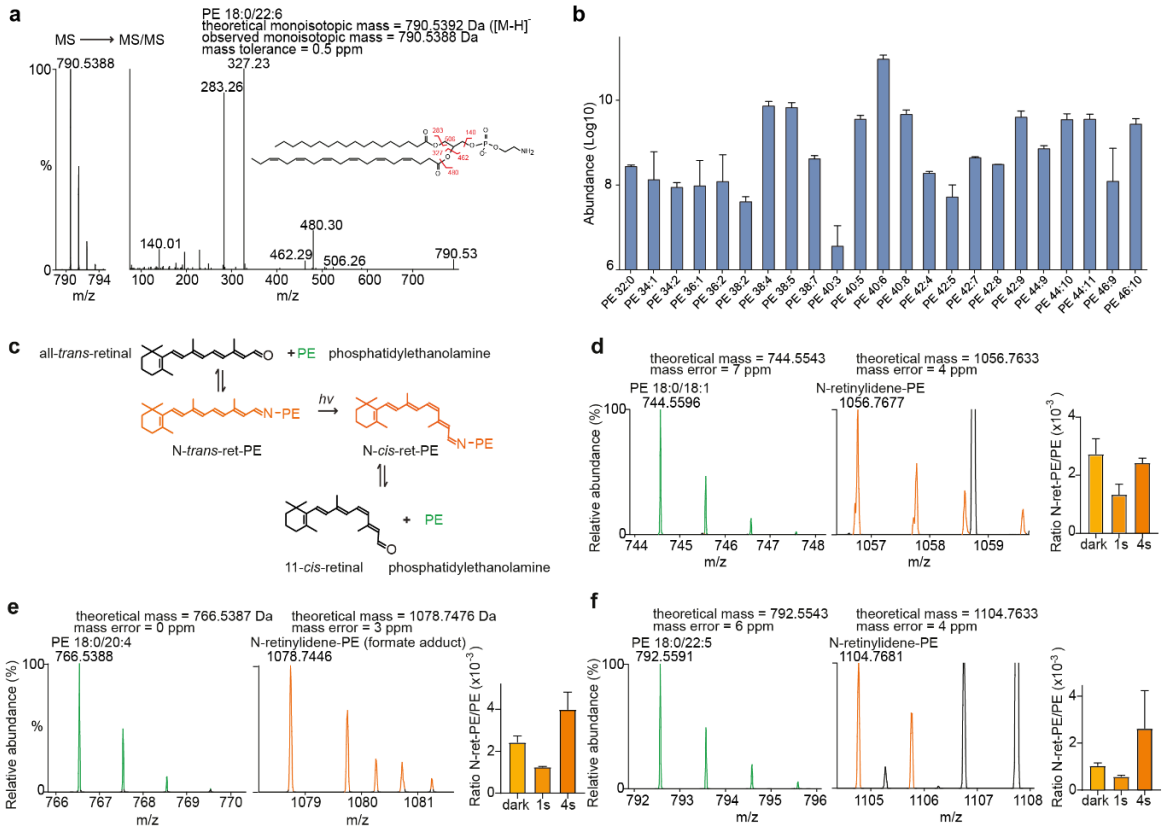


Figure 3.3 Extraction of lipids from the ROS membrane, identification of phosphatidylethanolamines and the formation of *N*-retinylidene-PE under controlled illumination conditions. (**a**) Analysis of lipids through tandem MS in the negative ion mode and assignment (**Methods**). Shown is the fragmentation pattern of PE 40:6 confirming unsaturation in one chain. (**b**) Profile and abundance of phosphatidylethanolamine in ROS membranes. Error bars represent the standard deviation from triplicate experiments. A high degree of unsaturation is observed with the most abundant lipid assigned to PE 40:6. (**c**) The reaction scheme of the formation of N-ret-PE. (**d**) Detection of the formation of *N*-retinylidene-PE 36:1. (**e**) *N*-retinylidene-38:4 and (**f**) *N*-retinylidene-40:5 under controlled illumination conditions and extraction into organic solvent for lipidomics. Negative ion mass

spectra, monitoring the formate ion of the conjugated lipid, reveals a low-intensity peak assigned through accurate mass to the N-retinylidene conjugate. Bar charts (**d**, **e**, **f** left): absolute abundance of *N*-ret-PE was monitored as a function of the illumination time and shows a clear decrease at 1 s for all three *N*-ret-PE lipids. Data are presented as mean ± s.d. (n = 3) from triplicates experiments. Between in dark and 1 s-illumination, P=0.020, P < 0.05 (**d**), 0.0030, P<0.01 (**e**), 0.0045, P<0.01 (**f**) respectively; two-tailed unpaired Student's t-test.

---

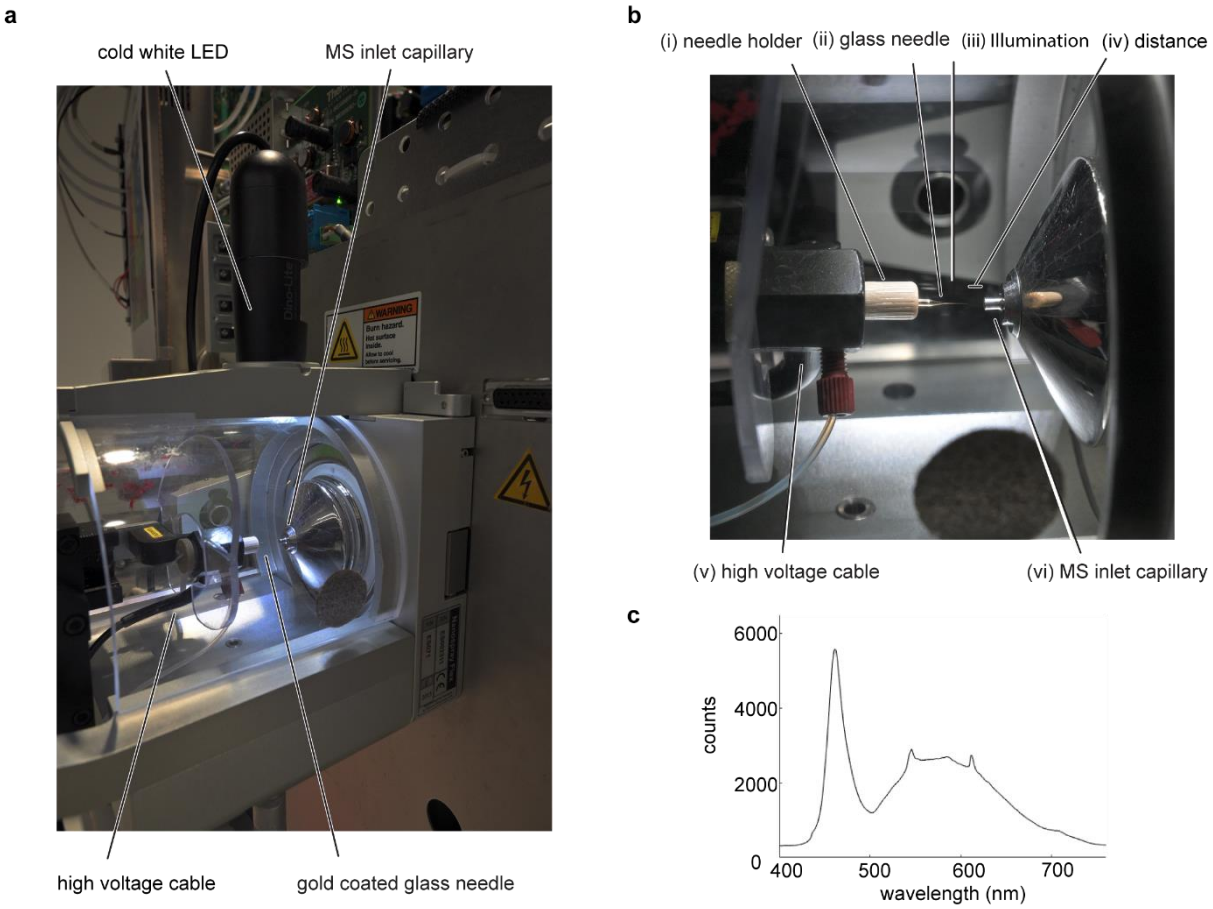
regeneration from 79% to 91% rho was observed after 3.8 min (**Supplementary Figure 3.6**). Together these results imply that conversion of opsin to rho in membranes can be supplemented by addition and photo-activation of all-*trans*-retinal and that PE conjugation provides a source of retinal for regeneration in membranes, a mechanism suggested for sustained vision in daylight[257].

3.4 Conclusion

Taken together our native mass spectrometry data document that we have profiled photoactivation and regeneration of rhodopsin in its native membrane. we have captured the conversion of rho to opsin in real-time and shown that it is significantly slower in the membrane compared to the detergent micelle. Complemented with the hydroxylamine experiments, we provide the insight into active rhodopsin with real-time measurement. We observed rapid

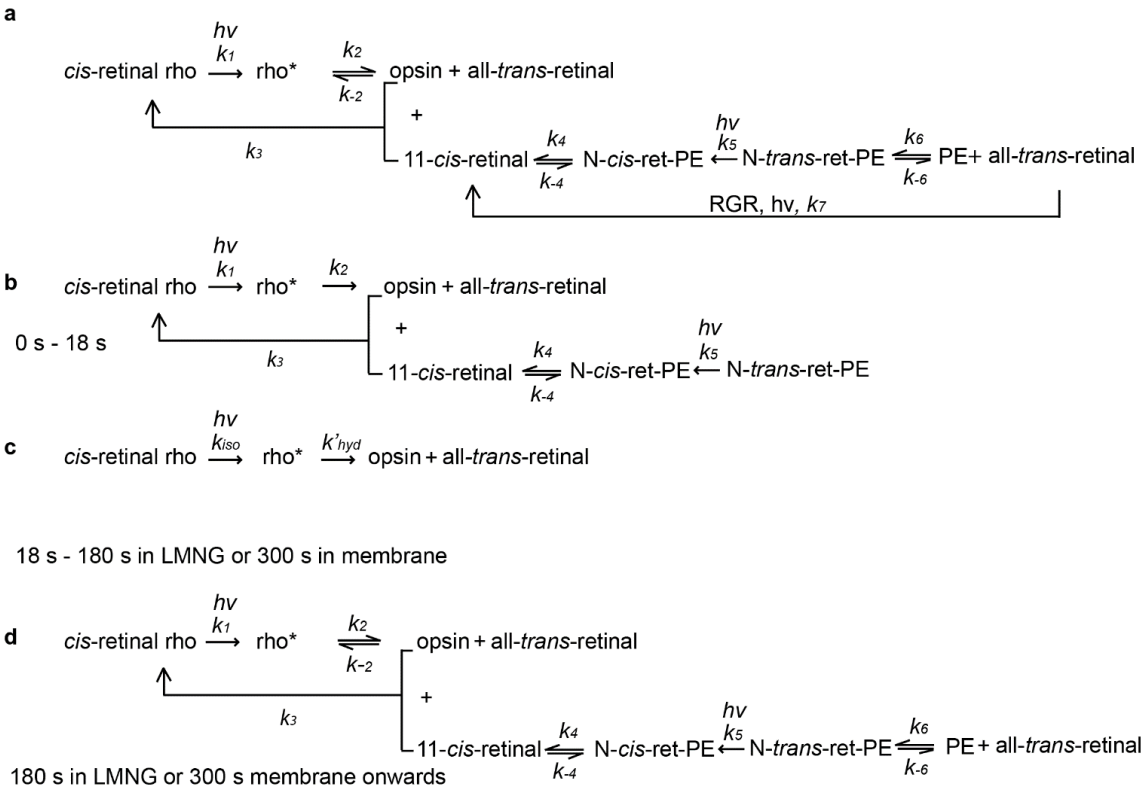
regeneration of rho following exposure to light and detected changes in the level of N-ret-PE and depletion of the conjugated chromophore following illumination. We anticipate that similar approaches will be applicable to other cell surface receptors (e.g. opsin family and olfactory receptors) present at high density enabling kinetic studies of molecular events, and identification of endogenous ligands that are vital to receptor functions. Overall we have highlighted a new approach potential for rhodopsin-like drug target investigation and drug discovery in a native environment.

3.5 Supplementary Information



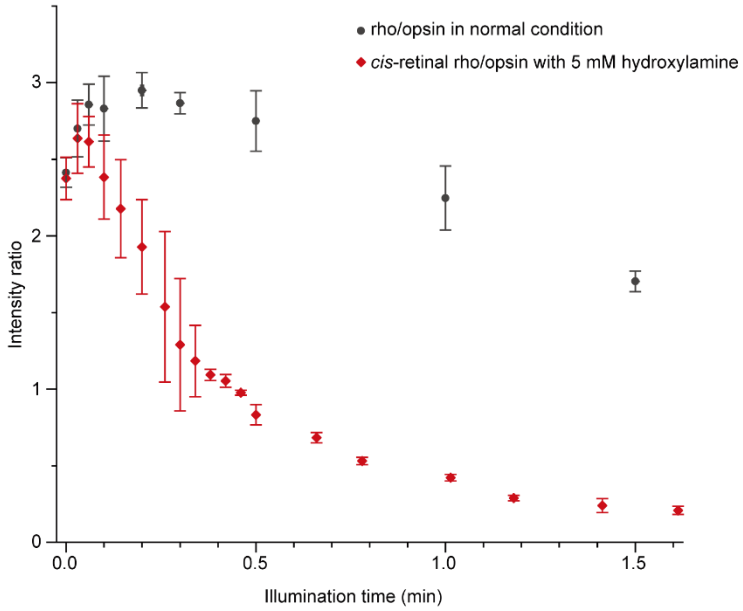
Supplementary Figure 3.1 Photographic images of the electrospray set-up showing the relationship between the light source the electrospray capillary and MS inlet and UV visible emission spectrum of the LED light source used here. (**a**) The relationship between the electrospray interface, the cold white LED light source and the mass spectrometer. The glass capillary was displayed on the same axis of the source aperture. The light stimulation set up was set orthogonally onto the capillary and the sample was at the illuminating centre. (**b**) (i) and (ii) membrane fragments are placed in the glass needle prepared in-house and secured in the needle holder. (iii) shows the orientation of the light source orthogonal to the electrospray

plume. (iv) to (vi) high voltage is applied to the needle holder and the distance between the needle tip and the MS inlet is adjusted to produce a stable electrospray ion current. Once this is achieved the electrospray interface is illuminated (t=0) and spectra are recorded in real-time either with continuous illumination or as a function of the post illumination period. (**c**) UV-visible emission spectrum of the cold white light source used here. The spectrum was recorded on an Andor Shamrock 303i Spectrograph equipped with an Andor Newton 920 camera. The wavelength maximum is 460 nm.



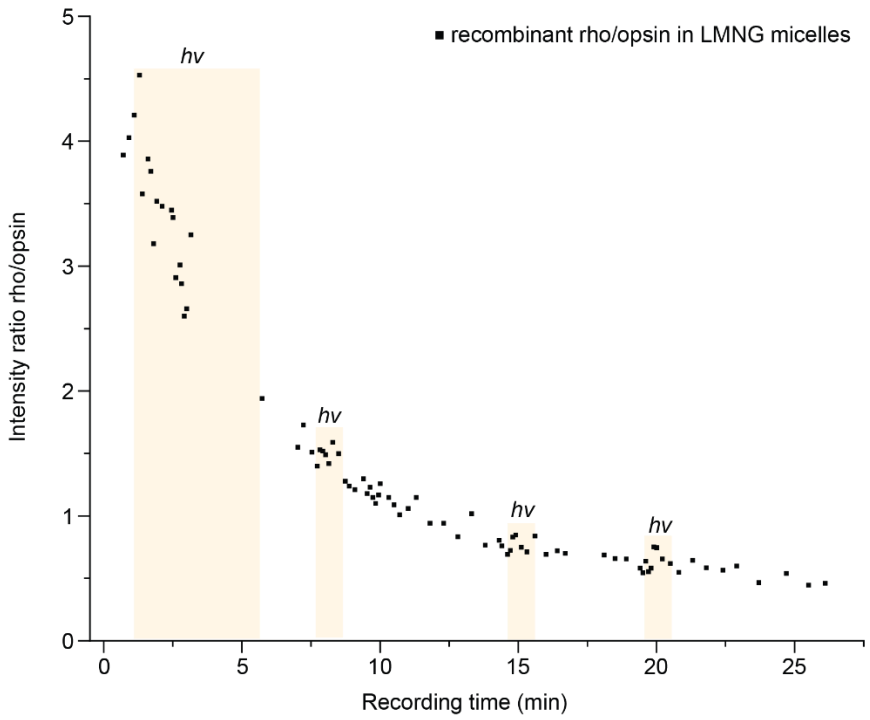
Supplementary Figure 3.2 Kinetic model for the photoconversion and regeneration of rho in disc membranes. (**a**) The photoconversion and regeneration processes involve a number of reactions outlined above. To simplify our kinetic model, we have divided the light response process into three phases. First, < 18 s when a rapid increase of rho is observed, prior to a steady decrease. (**b**) This initial fast reaction 0 - 18 s is attributed to the rapid regeneration of *cis*- retinal rho competing with photoconversion (rho -> ops). Dark-adapted membranes possess N-ret-PE which responses rapidly to light such N-all-*trans*-ret-PE is activated and dissociates into *cis*-retinal and PE. Regenerated *cis*-retinal is then able to interact with opsin to form *cis*-retinal rho. Concurrently, existing *cis*-retinal rho photoconverts to active rho* and then deactivated opsin. (**c**) During the second phase (> 18 s), after the original N-ret-PE is

consumed, isomerization of *cis*-retinal rho to rho* continues with minimal regeneration of *cis*-retinal rho. rho* hydrolyses into opsin and all-*trans*-retinal.

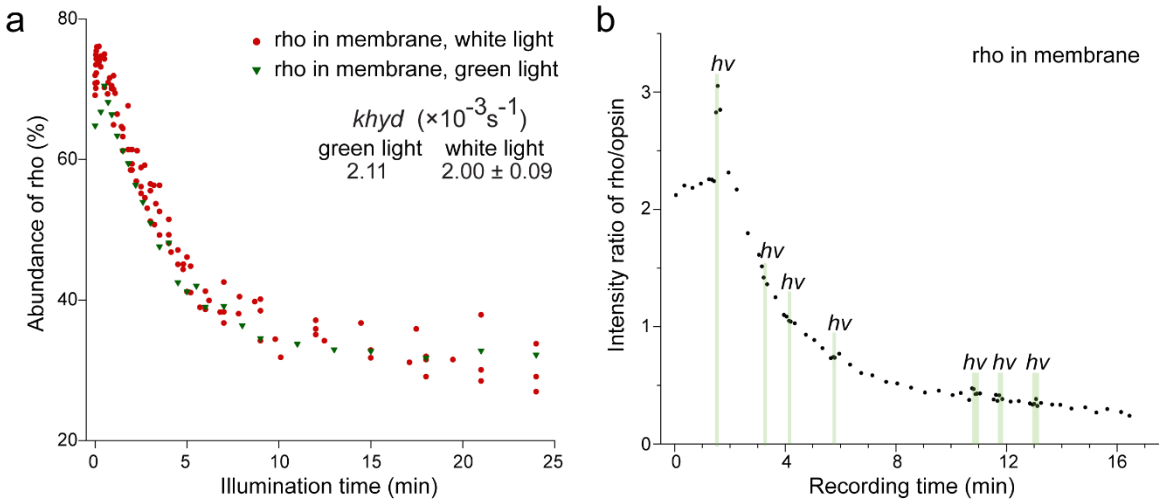


Supplementary Figure 3.3 Decay curve of rho in normal condition and cis-retinal rho following treatment with hydroxylamine measured by native MS. The plot of the photoconversion of 11-*cis*-retinal rho to opsin in membranes either pre-treated with 5 mM hydroxylamine (red) or rho to opsin in native membranes (grey) monitored by mass spectrometry under the conditions described for **Figure 2** (main text). The experiments were performed three times independently under continuous illumination conditions and data are presented as mean ± s.e. (n = 3).

.

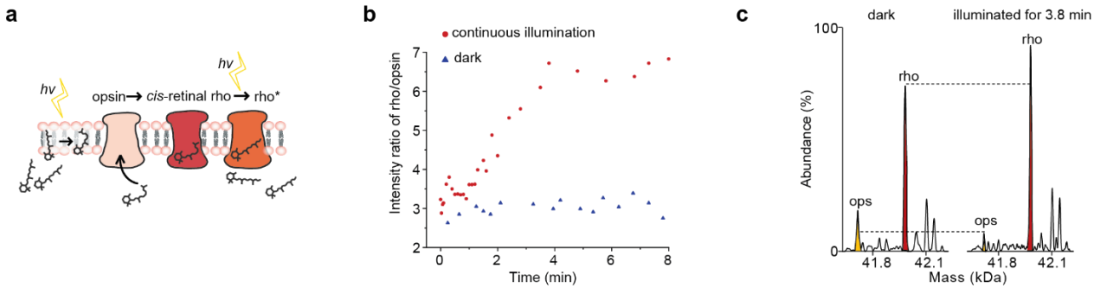


Supplementary Figure 3.4 Monitoring the conversion of detergent extracted bovine rho to opsin through its mass change in LMNG micelles during applied illumination pulses (shaded orange) at pH 7.0 and 28 °C. Pulses were applied at 1.1 min, 7.6 min, 15 min and 20 min for 4.6 min, 18 s, 40 s and 35 s individually. The experimental noise makes it challenging to measure the increase in these illumination pulses in micelles compared to the experiments in the ROS membrane. However, a clear trend is observed after *circa* 20 min. The experiments were repeated twice.



Supplementary Figure 3.5 Investigation of photoconversion under green light conditions. (**a**) The same experimental set-up as described in the main paper was used to measure the conversion of rho to opsin in native membrane fragments under either white or green light conditions. Phototransformation of rhodopsin in green light exhibits similar trends as in white light. The decay is marginally faster under green light conditions probably due to the absence of cis-retinal generation from retinal-lipid adducts. We performed an analogous experiment (**Figure 3.2g**) monitoring rho-opsin conversion following pulses of green light at discrete time points. Except for the first pulse, we did not observe increased population of rho at every pulse of green light (c.f. **Figure 3.2g** for white light comparison). We attribute the rise of rho during the first pulse to low populations of RGR, which is hard to be detected. The lack of increased rho at subsequent light pulses is attributed to the absence of the cis-retinal generation from retinal-lipid adducts under green light. (**b**) Monitoring the rho-opsin photo conversion

following pulses of green light at discrete time points. Data are presented as mean ± s.e. (*n*=3) except an experiment under continuous illumination of green light.



Supplementary Figure 3.6 Rho regeneration in the presence of excess all-*trans*-retinal. (**a**) Schematic process of ROS opsin regenerated into rho in 10-fold excess of all-*trans*-retinal under continuous illumination. (**b**) The intensity ratio of rho to opsin increases until it reaches a plateau at 3.8 min. Comparatively, rho remains constant in the presence of an exogenous all-*trans*-retinal in the dark. (**c**) Zero-charge spectra of rho and opsin with 10-fold excessive all-*trans*-retinal present before and following 3.8 min illumination. This experiment was repeated twice.

	isomers	calculated mass (Da)	LMNG				membrane			
			mass (Da)	standard error (Da)	abundance	standard error (Da)	mass (Da)	standard error (Da)	abundance	standard error (Da)
dark	rhodopsin,1N-acetyl, 1pal,6hex,6hexNAc	41744	41743.3	0.1	5.1%	0.2%	41742.6	0.3	3.0%	1.2%
	rhodopsin,1N-acetyl,2pal,6hex,6hexNAc	41982	41981.6	0.1	35.8%	0.9%	41981.1	0.1	36.8%	0.4%
	rhodopsin,1N-acetyl,2pal,6hex,6hexNAc,1cys	42103	42103.0	0.2	10.0%	0.2%	42103.3	0.3	10.4%	0.3%
	rhodopsin,1N-acetyl,2pal,7hex,6hexNAc	42144	42144.0	0.2	8.2%	0.3%	42144.1	0.2	7.8%	0.2%
	rhodopsin,1N-acetyl,2pal,8hex,6hexNAc	42306	42306.2	0.1	8.2%	0.6%	42306.2	0.2	6.9%	0.1%
	rhodopsin,1N-acetyl,2pal,9hex,6hexNAc	42468	42468.5	0.2	4.7%	0.3%	42468.1	0.2	4.8%	0.4%
	rhodopsin,1-Nacetyl,2pal,10hex,6hexNAc	42630	42629.8	0.4	1.5%	0.1%	42629.5	0.2	1.4%	0.2%
	opsin,1-Nacetyl,1pal,6hex,6hexNAc	41478	41477.3	0.0	1.0%	0.5%	41475.9	0.4	1.3%	0.4%
	opsin,1-Nacetyl,2pal,6hex,6hexNAc	41716	41715.6	0.1	13.1%	0.5%	41715.2	0.0	14.8%	0.7%
	opsin,1-Nacetyl,2pal,6hex,6hexNAc,1pho	41796	41796.8	0.5	1.3%	0.1%	41797.6	0.4	1.2%	0.2%
	opsin,1-Nacetyl,2pal,6hex,6hexNAc,1cys	41837	41837.0	0.4	3.7%	0.1%	41836.4	0.4	3.6%	0.2%
	opsin,1-Nacetyl,2pal,7hex,6hexNAc	41878	41877.9	0.7	2.4%	0.2%	41878.2	0.1	3.2%	0.4%
	opsin,1-Nacetyl,2pal,8hex,6hexNAc	42040	42039.4	0.2	2.8%	0.4%	42040.4	0.3	2.9%	0.6%
	opsin,1-Nacetyl,2pal,9hex,6hexNAc	42202	42202.6	0.3	1.4%	0.1%	42202.0	0.6	1.2%	0.1%
	opsin,1-Nacetyl,2pal,10hex,6hexNAc	42364	42364.0	1.0	0.8%	0.0%	42362.1	0.9	0.7%	0.3%
light	rhodopsin,1N-acetyl, 1pal,6hex,6hexNAc	41744	41743.7	0.1	2.6%	0.6%	41741.3	0.7	2.3%	0.5%
	rhodopsin,1N-acetyl,2pal,6hex,6hexNAc	41982	41981.7	0.3	12.1%	1.5%	41981.2	0.2	14.5%	0.6%
	rhodopsin,1N-acetyl,2pal,6hex,6hexNAc,1cys	42103	42103.7	1.3	3.1%	0.4%	42104.7	0.3	4.5%	0.6%
	rhodopsin,1N-acetyl,2pal,7hex,6hexNAc	42144	42143.8	0.7	3.1%	0.2%	42143.8	0.1	3.8%	0.6%
	rhodopsin,1N-acetyl,2pal,8hex,6hexNAc	42306	42307.0	1.1	2.9%	0.6%	42306.0	1.2	2.8%	0.3%
	rhodopsin,1N-acetyl,2pal,9hex,6hexNAc	42468	42464.9	1.7	1.6%	0.3%	42468.4	0.2	2.3%	0.4%
	rhodopsin,1-Nacetyl,2pal,10hex,6hexNAc	42630	42635.7	3.6	1.1%	0.3%	42632.3	1.5	1.6%	0.3%
	opsin,1-Nacetyl,1pal,6hex,6hexNAc	41478	41477.2	0.7	5.3%	0.2%	41446.9	0.6	1.2%	0.3%
	opsin,1-Nacetyl,2pal,6hex,6hexNAc	41716	41714.9	0.3	32.8%	0.7%	41714.8	0.2	34.9%	2.2%
	opsin,1-Nacetyl,2pal,6hex,6hexNAc,1pho	41796	41798.6	0.1	2.2%	0.3%	41795.7	0.4	2.2%	0.3%
	opsin,1-Nacetyl,2pal,6hex,6hexNAc,1cys	41837	41836.4	0.1	10.2%	0.1%	41836.1	0.1	9.4%	0.2%
	opsin,1-Nacetyl,2pal,7hex,6hexNAc	41878	41878.4	0.4	8.4%	0.4%	41878.0	0.4	7.2%	0.5%
	opsin,1-Nacetyl,2pal,8hex,6hexNAc	42040	42039.3	0.1	7.6%	1.0%	42039.6	0.7	6.9%	0.7%
	opsin,1-Nacetyl,2pal,9hex,6hexNAc	42202	42201.3	0.7	4.7%	0.1%	42200.5	0.4	4.2%	0.4%
	opsin,1-Nacetyl,2pal,10hex,6hexNAc	42364	42364.6	1.3	2.2%	0.3%	42364.2	1.1	2.2%	0.3%

Supplementary Table 3.1 Measured mass with standard error and abundance of ROS

rho/opsin ejected from LMNG micelles and disc membranes from either dark-adapted membranes (dark) or following 20 min of illumination (light). The experiment was performed under continuous illumination conditions and the standard error represents the variation across three replicate experiments.

Accession	Mass	Num. of matches	Num. of significant matches	Num. of sequences	Num. of significant sequences	Description
tr|F1MWM0|F1MWM0_BOVIN	259606	23	15	13	8	ATP binding cassette subfamily A member 4 OS=Bos taurus OX=9913 GN=ABCA4 PE=4 SV=2
sp|P00829|ATPB_BOVIN	56249	20	17	15	12	ATP synthase subunit beta, mitochondrial OS=Bos taurus OX=9913 GN=ATP5F1B PE=1 SV=2
sp|P02699|OPSD_BOVIN	39552	15	10	6	5	Rhodopsin OS=Bos taurus OX=9913 GN=RHO PE=1 SV=1
tr|F1MLW4|F1MLW4_BOVIN	140140	15	13	11	9	Retinol-binding protein 3 OS=Bos taurus OX=9913 GN=RBP3 PE=4 SV=2
tr|E1BJB1|E1BJB1_BOVIN	50274	7	5	4	3	Tubulin beta chain OS=Bos taurus OX=9913 GN=TUBB2A PE=1 SV=1
tr|F1MZV1|F1MZV1_BOVIN	103425	8	7	6	5	Hexokinase-1 OS=Bos taurus OX=9913 GN=HK1 PE=1 SV=2
sp|Q2HJ86|TBA1D_BOVIN	50935	8	7	5	4	Tubulin alpha-1D chain OS=Bos taurus OX=9913 GN=TUBA1D PE=1 SV=1
sp|P52205|ROM1_BOVIN	37965	12	10	9	7	Rod outer segment membrane protein 1 OS=Bos taurus OX=9913 GN=ROM1 PE=1 SV=2
sp|Q9BGI7|OPSR_BOVIN	41185	5	4	4	3	Long-wave-sensitive opsin 1 OS=Bos taurus OX=9913 GN=OPN1LW PE=2 SV=1
sp|Q28139-2|NCKX1_BOVIN	130174	8	6	4	3	Isoform 2 of Sodium/potassium/calcium exchanger 1 OS=Bos taurus OX=9913 GN=SLC24A1
sp|P0C0S9|H2A1_BOVIN	14083	5	5	1	1	Histone H2A type 1 OS=Bos taurus OX=9913 PE=1 SV=2
sp|Q3T100|MGST3_BOVIN	17101	5	3	2	2	Microsomal glutathione S-transferase 3 OS=Bos taurus OX=9913 GN=MGST3 PE=2 SV=1
sp|P08168|ARRS_BOVIN	45418	11	9	9	7	S-arrestin OS=Bos taurus OX=9913 GN=SAG PE=1 SV=1
tr|F1MLB8|F1MLB8_BOVIN	59767	7	5	6	4	ATP synthase subunit alpha OS=Bos taurus OX=9913 GN=ATP5F1A PE=1 SV=1
sp|P80724|BASP1_BOVIN	22997	4	4	3	3	Brain acid soluble protein 1 OS=Bos taurus OX=9913 GN=BASP1 PE=1 SV=3
sp|Q2HJ97|PHB2_BOVIN	33337	3	2	2	1	Prohibitin-2 OS=Bos taurus OX=9913 GN=PHB2 PE=2 SV=1
tr|F1MRD0|F1MRD0_BOVIN	42167	3	2	2	1	Actin, cytoplasmic 1 OS=Bos taurus OX=9913 GN=ACTB PE=3 SV=2
tr|E1BME9|E1BME9_BOVIN	51664	2	2	1	1	NADH dehydrogenase [ubiquinone] flavoprotein 3, mitochondrial OS=Bos taurus OX=9913 GN=NDUFV3 PE=1 SV=2
tr|F1N0X5|F1N0X5_BOVIN	39839	6	5	6	5	Peripherin-2 OS=Bos taurus OX=9913 GN=PRPH2 PE=4 SV=1
sp|P04695|GNAT1_BOVIN	40396	5	4	4	3	Guanine nucleotide-binding protein G(t) subunit alpha-1 OS=Bos taurus OX=9913 GN=GNAT1 PE=1 SV=3
sp|Q00194|CNGA1_BOVIN	79951	5	2	3	1	cGMP-gated cation channel alpha-1 OS=Bos taurus OX=9913 GN=CNGA1 PE=1 SV=1
sp|Q2VDL6|AT1A2_BOVIN	113419	5	3	4	2	Sodium/potassium-transporting ATPase subunit alpha-2 OS=Bos taurus OX=9913 GN=ATP1A2 PE=1 SV=1
tr|F1MS25|F1MS25_BOVIN	27999	2	2	1	1	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial OS=Bos taurus OX=9913 GN=NDUFS3 PE=3 SV=2
tr|A6H783|A6H783_BOVIN	31193	2	2	1	1	VDAC5P protein OS=Bos taurus OX=9913 GN=VDAC5P PE=2 SV=1
sp|P21457|RECO_BOVIN	23376	4	3	2	2	Recoverin OS=Bos taurus OX=9913 GN=RCVRN PE=1 SV=3
tr|F1MBC7|F1MBC7_BOVIN	11975	2	2	1	1	Cytochrome c oxidase subunit 6A, mitochondrial OS=Bos taurus OX=9913 PE=3 SV=1
sp|Q3ZBU2|CISD1_BOVIN	12146	2	2	1	1	CDGSH iron-sulfur domain-containing protein 1 OS=Bos taurus OX=9913 GN=CISD1 PE=1 SV=1
sp|P13620|ATP5H_BOVIN	18738	8	4	4	2	ATP synthase subunit d, mitochondrial OS=Bos taurus OX=9913 GN=ATP5PD PE=1 SV=2
sp|Q08DK4|GHC1_BOVIN	34965	3	1	1	1	Mitochondrial glutamate carrier 1 OS=Bos taurus OX=9913 GN=SLC25A22 PE=2 SV=1
sp|P13619|AT5F1_BOVIN	28860	4	4	3	3	ATP synthase F(0) complex subunit B1, mitochondrial OS=Bos taurus OX=9913 GN=ATP5PB PE=1 SV=2
tr|F1N2J9|F1N2J9_BOVIN	23842	4	2	4	2	Ras-related protein Rab-8B OS=Bos taurus OX=9913 GN=RAB8B PE=4 SV=1
sp|P02698|GBG1_BOVIN	8709	2	2	1	1	Guanine nucleotide-binding protein G(T) subunit gamma-T1 OS=Bos taurus OX=9913 GN=GNGT1 PE=1 SV=2
sp|Q3T165|PHB_BOVIN	29843	3	2	2	2	Prohibitin OS=Bos taurus OX=9913 GN=PHB PE=2 SV=1
sp|P27674|GTR1_BOVIN	54439	2	2	2	2	Solute carrier family 2, facilitated glucose transporter member 1 OS=Bos taurus OX=9913 GN=SLC2A1 PE=1 SV=1
sp|P34943|NDUA9_BOVIN	42879	2	1	2	1	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial OS=Bos taurus OX=9913 GN=NDUFA9 PE=1 SV=1
sp|P32007|ADT3_BOVIN	33084	9	1	6	1	ADP/ATP translocase 3 OS=Bos taurus OX=9913 GN=SLC25A6 PE=1 SV=3
sp|P00423|COX41_BOVIN	19616	4	3	2	2	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial OS=Bos taurus OX=9913 GN=COX4I1 PE=1 SV=1
sp|P55203|GUC2D_BOVIN	121433	6	3	4	2	Retinal guanylyl cyclase 1 OS=Bos taurus OX=9913 GN=GUCY2D PE=1 SV=1
sp|P00426|COX5A_BOVIN	16896	5	4	3	2	Cytochrome c oxidase subunit 5A, mitochondrial OS=Bos taurus OX=9913 GN=COX5A PE=1 SV=2
tr|F1MXX0|F1MXX0_BOVIN	83392	4	2	3	1	MICOS complex subunit MIC60 OS=Bos taurus OX=9913 GN=IMMT PE=1 SV=1
sp|P02721|ATP5J_BOVIN	12524	4	3	2	2	ATP synthase-coupling factor 6, mitochondrial OS=Bos taurus OX=9913 GN=ATP5PF PE=1 SV=2
sp|Q8MJG0|R9BP_BOVIN	25978	1	1	1	1	Regulator of G-protein signaling 9-binding protein OS=Bos taurus OX=9913 GN=RGS9BP PE=1 SV=2
tr|A0A0N4STN0|A0A0N4STN0_BOVIN	11409	2	2	1	1	Uncharacterized protein OS=Bos taurus OX=9913 GN=ATP5MG PE=4 SV=1
sp|P04394|NDUV2_BOVIN	27575	1	1	1	1	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial OS=Bos taurus OX=9913 GN=NDUFV2 PE=1 SV=3
tr|E1BJA2|E1BJA2_BOVIN	67115	3	1	3	1	Apoptosis inducing factor mitochondria associated 1 OS=Bos taurus OX=9913 GN=AIFM1 PE=1 SV=1
tr|A7Z066|A7Z066_BOVIN	68189	3	2	2	1	Calnexin OS=Bos taurus OX=9913 GN=CANX PE=1 SV=1

Supplementary Table 3.2 Protein identification of purified bovine ROS disc membrane. RGR was not found by proteomics. This excludes the possibility of RGR based regeneration in this system. Experimental details see **Methods**. Protein IDs presented here are reproducible from triplicate proteomics experiments.

Chapter 4 Capturing Rhodopsin, A G-Protein Coupled Receptor, Signalling Cascade across A Native Membrane

Adapted from a publication in Nature, 2022

Siyun Chen[1,2], Tamar Getter[3], David Salom[3], Di Wu[1,2], Daniel Quetschlich[1,2], Dror S. Chorev[1*], Krzysztof Palczewski[3,4*], Carol V. Robinson[1,2*].

1. Chemistry research Laboratory, University of Oxford, South Path Parks Road Oxford, UK.
2. Kavli Institute for Nanoscience Discovery, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK.
3. Gavin Herbert Eye Institute, Department of Ophthalmology, University of California, Irvine, Irvine, CA, USA
4. Departments of Physiology and Biophysics, Chemistry, Molecular Biology and Biochemistry, University of California, Irvine, Irvine, CA, USA.

*Joint corresponding authors

Acknowledgements

All mass spectrometry research was funded by a Welcome Trust Investigator Award (221795/Z/20/Z) and an ERC Advanced Grant ENABLE (695511) to C.V.R. A grant from National Institute of Health (R01EY030912) and unrestricted grants from Research to Prevent Blindness to the Departments of Ophthalmology at UCI are gratefully acknowledged by the K.P. laboratory. The authors acknowledge helpful discussions with Martin Galpin and Tarick El-Baba on the interpretation of kinetic data. For the purpose of Open Access, the authors have applied for a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Contributions

C.V.R. and K.P. conceived the idea through the CIFAR network and provided the overall supervision for the study. S.C. performed all of the experiments described in this paper under the supervision of D.S.C. and D.W. for the SoLVe and lipidomics experiments respectively. D.Q. provided computational support for analysis and T.G., D.S. and K.P. prepared and provided ROS Disc membranes, identified rho targeting compounds, and prepared PDE6 and transducin complexes. C.V.R., S.C. and K.P. wrote the paper with contributions from other authors.

Abstract

Rhodopsin has been an important model for the studies in G protein-coupled receptors (GPCR) which respond to a variety of stimuli to induce signalling pathways across cell membranes. Recent progress has yielded atomic structures of key intermediates[316,317] and roles for lipids in signalling[318,319]. However, capturing signalling events of a wild-type receptor in real-time, across a native membrane to its downstream effectors, has remained elusive. Here we use native mass spectrometry to probe the archetype class A GPCR, rhodopsin, directly from fragments of native disc membranes. Capturing the secondary steps of the signalling cascade, we monitor light activation of transducin (Gt) through the loss of GDP to generate intermediate *apo* trimeric G protein and observe Gt$a_{.GTP}$ subunits interacting with PDE6 to hydrolyse cyclic GMP. We then apply rhodopsin-targeting compounds to show how they either stimulate or dampen signalling via rhodopsin-opsin and transducin signalling pathways. Our results not only reveal the impact of the native lipid environment on rhodopsin signalling and regeneration but also enable us to propose a new paradigm for GPCR drug discovery in native membrane environments.

4.1 Introduction

Rhodopsin has been an important model for the studies in G protein-coupled receptors (GPCR). Molecular details of GPCRs, their G protein coupling and arrestin interactions are providing unprecedented insight into signalling cascades and are often achieved through judicious antibody stabilisation or protein engineering[316]. Meanwhile, recognition of the importance of lipids in mediating rhodopsin-like GPCR signalling interactions has been derived from the use of lipid nanodiscs and peptidiscs to recreate membrane environments[264,320,321]. Capturing unmodified GPCRs, in native membrane environments, has so far eluded biophysical measurement. Here, selecting the best characterised GPCR rhodopsin, the dim light receptor of the mammalian visual system[23], light exposure prompts to form activated all-*trans*-retinylidene (rho*). All-*trans*-retinal rho* interacts with GDP-bound transducin (heterotrimeric G protein, Gt) leading to GDP-GTP exchange and subsequent dissociation and formation of G$t\alpha_{.GTP}$ which then interacts with phosphodiesterase 6 (PDE6) displacing the PDE6γ subunit, releasing inhibition of catalysis, resulting in rapid hydrolysis of cytoplasmic cGMP[239]. Using native mass spectrometry, we discover capture signalling cascade of rhodopsin covering the interaction details with lipid and their influence in facilitating G protein-coupled rhodopsin. Moreover, this platform provides insights in how small molecules mediate rhodopsin kinetics and functions.

4.2 Methods

Purification of bovine rod disc membrane

Rho-containing disc membranes were purified from dark-adapted retinas as described previously [188,306]. ROS washed with isotonic and hypotonic buffers were used for rho purification (in dark under a dim red light with > 670 nm filter) as previously described [307,322].

Purification of bovine rod Gt and PDE6

Trimeric Gt and PDE6 were prepared as previously described [28]. Briefly, bovine ROS were suspended in isotonic buffer containing 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 100 mM NaCl, 1 mM DTT, and 5 mM $MgCl_2$ and centrifuged at 31,000 g for 25 min at 4 °C. The pellet was homogenized in a hypotonic buffer containing 5 mM HEPES, pH 7.5, 1 mM EDTA and 1 mM DTT and centrifuged at 40,000 *g* for 30 min at 4°C multiple times. The ratio between Gt and PDE6 was analysed by ImageJ software on Coomassie blue-staining sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). For addition of the soluble fraction to the membrane, Gt and PDE6 as purified above were first buffer exchange 3 times with 200 mM ammonium acetate using Amicon filter (10 kDa cut-off). Under dim red light, the portion of rho, Gt and PDE6 was mixed at a molar ratio of 13:4:1 for 15 min on ice to allow diffusion of Gt and PDE6 before analyses by native mass spectrometry. Rho, Gt and PDE6 at molar ratio of 65:8:2 was illuminated under cold white LED for 30 min at room temperature and centrifuged at 20,000 *g* for 20 min to remove the membrane to capture post-signalling PDE6 for native mass spectrometry (Data presented in **Supplementary Figure 4.1 g-i** and **Figure 4.2d**).

Native mass spectrometry set-up and illumination conditions

The native membrane preparation (2-3 µl) was placed directly in a gold coated nano-ES capillary prior to the source of the mass spectrometer (**Figure 4.1c**). The cold LED light source (120 µW at 10 cm distance with a surface area of illumination 3 mm^2) was assembled via a

bracket to illuminate the tip of the capillary for a defined interval in a dim red-light environment. For a real-time dynamic process, the membrane was first recorded in the dark and then illuminated with light pulses, or continuously, depending on the experiment. Native mass spectrometry experiments were carried out on a Q-Exactive adapted for membrane proteins [124]. The following parameters were used typically and adjusted to release the membrane protein from the detergents and membrane lipids: capillary voltage: 0.8-1.4 kV, HCD energy: 125-200 V, dessolvation voltage: 0 V, source fragmentation: 100 V, HCD pressure: 3-8x10^{-10} mbar. C-trap entrance lens tune offset was set to 2, injection flatapole was set to 7 V, inter flatapole lens was at 0 V, and bent flatapole at 1 V. Threshold was set to 3. These MS conditions were modified to focus on the individual components of the spectrum. For measurement of PDE6 spectra the following parameters were used: capillary voltage: 1.1 kV, dessolvation voltage: 0 V, source fragmentation: 50 V, HCD energy: 175 V. HCD pressure was typically 7.3x10^{-10} mbar.

Polyunsaturated PC lipids surrounding rho identified with multistage native MS^n

The mixture of rho, Gt and PDE6 at a molar ratio of 65:8:2 (described above) was used to identify the polyunsaturated PC surrounding rho during light induced signalling. PC lipids are released from rho and detected in positive mode using an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Fisher Scientific) [127]. In general, in-source activation (100 V) was applied to rho in membranes and the resulting ions transferred to the IRM for activation (HCD normalized collision energy 100%, IRM pressure 8 mTorr) to promote dissociation into lipids. AGC target values (100-150%) and maximum injection time (100 ms) were adjusted manually to maximize the normalized level. Detection at this stage was typically performed in the Orbitrap at a high m/z range 500-8000. High resolution at 60,000 can be employed to enhance

the intensity of lipid signals at low m/z. Data were analyzed using the Xcalibur software package 2.2 (Thermo Fisher Scientific).

Data analysis

Raw data was first analysed manually using Xcalibur 2.2 (Thermo Fisher Scientific). The relative abundance of each species in a real-time measurement was quantified using Lig2Apo, a simple jupyter notebook. A folder of text files (format m/z versus intensity) is exported from Xcalibur 2.2 (Thermo Fischer Scientific) and two series of m/z values are defined corresponding to the molecular species of interest. The program then reads all txt files and calculates the relative intensity of the series and the intensity ratio of two species (see below for code availability). The zero charge spectra presented in **Figure 4.4** were analyzed by Unidec software[308].

Proteomics and protein identification

Protein bands were excised from gels and processed as described previously[323]. Peptides generated were re-suspended in 0.1% Formic Acid and separated on an Ultimate 3000 UHPLC system (Thermo Fisher Scientific) and electrosprayed directly into a Q Exactive mass spectrometer (Thermo Fisher Scientific) through an EASY-Spray nano-electrospray ion source (Thermo Fisher Scientific). The peptides were trapped on a C18 PepMap100 pre-column (300 µm i.d. x 5 mm, 100 Å, Thermo Fisher Scientific) using solvent A (0.1% Formic Acid in water) at a pressure of 500 bar. The peptides were separated on an analytical column (75 µm i.d. packed with ReproSil-Pur 120 C18-AQ, 1.9 µm, 120 Å, by Dr Maisch GmbH) using a gradient (15% to 38% for 30 min, solvent B - 0.1% formic acid in acetonitrile, flow rate: 200 nL/min) for 15 min. The raw data was acquired in a data-dependent acquisition mode (DDA). Full scan mass spectra were acquired in the Orbitrap (scan range 350-1500 m/z, resolution 70000, AGC

target 3e6, maximum injection time 50 ms). After the MS scans, the 10 most intense peaks were selected for HCD fragmentation at 30% of the normalized collision energy. HCD spectra were also acquired in the Orbitrap (resolution 17500, AGC target 5^{e4}, maximum injection time 120 ms) with first fixed mass at 180 m/z. Charge exclusion was selected for 1+ and 2+ ions. The dynamic exclusion was set to 5 s. All peptides were manually validated. Peptide identification and data analysis was carried out using the MASCOT Daemon client program and server.

Selection and evaluation of rho-targeting molecules

Compound used in this study were selected via high-throughput screening (HTS) of a diverse library of 50,000 small molecules. The HTS approach used cells expressing tagged rho with β-galactosidase complementary and BRET detection systems. To validate the identified hits further rho photobleaching assays were carried out along with ex-vivo electrophysiology recordings[324,325].

Hit compounds F2515-3945 (#**1**), F3382-0749 (#**2**), F2502-0030 (#**3**), F3215-0002 (#**6**), F5103-0385 (#**7**), F5097-2767 (#**8**), F0834-0928 (#**9**) were obtained from Life Chemicals (ON L0S 1J0, Canada). The racemic mixture of sulconazole (#**4 and** #**5**) was purchased from MilliporeSigma (Burlington, MA, USA) and separated as described previously on an Agilent 1100 HPLC system (Agilent Technologies, Santa Clara, CA, USA) using a cellulose tris 3,5 dimethylphenyl carbamate chiralicar column (Chiral Technologies, West Chester, PA, USA)[324].

For mass spectrometry these compounds were dissolved individually in DMSO and diluted 1000-fold in disc membrane vesicle preparations in 200 mM ammonium acetate buffer to give a concentration for each compound of 90 μM, and a 10:1 final molar ratio of compound to rhodopsin. The disc membrane and compound were incubated for 30 min on ice, allowing for

complete diffusion before native MS measurement of rho/opsin. Disc membrane supplemented with Gt and PDE6 at molar ratio of 6.5:4:1 (rho:Gt:PDE6) was incubated with 90 μM compound in 200 mM ammonium acetate for 30 min on ice before native MS measurement of Gt. A control experiment in the absence of membranes was performed with Gt and PDE6 incubated with 90 μM compound at 4:1:50 for 30 min on ice before native MS measurement of Gt.

Data availability

The native MS data that support the findings of this study are available and deposited to Figshare as doi.org/10.25446/oxford.16901326

Code Availability

The script for calculating the ratio of opsin to rhodopsin ratio or Gt to $Gt._{GDP}$ in mass spectra can be accessed under https://github.com/d-que/Lig2Apo.

4.3 Results and Discussion

We first form lipid vesicles from disc membranes of dark-adapted bovine retina rod outer segments (ROS) and eject this GPCR directly from fragments of its native membrane environment into a mass spectrometer (**Figure 4.1a, b**). We monitored signalling by operating our mass spectrometer in a dark room and controlling light exposure to prompt photoisomerization of rhodopsin. Here we show that all participants of this signalling pathway (protein, downstream effectors, cofactors and lipids) can be ejected simultaneously and their response to photon activation in the native membrane captured by native MS directly and in real time (**Figure 4.1c**).

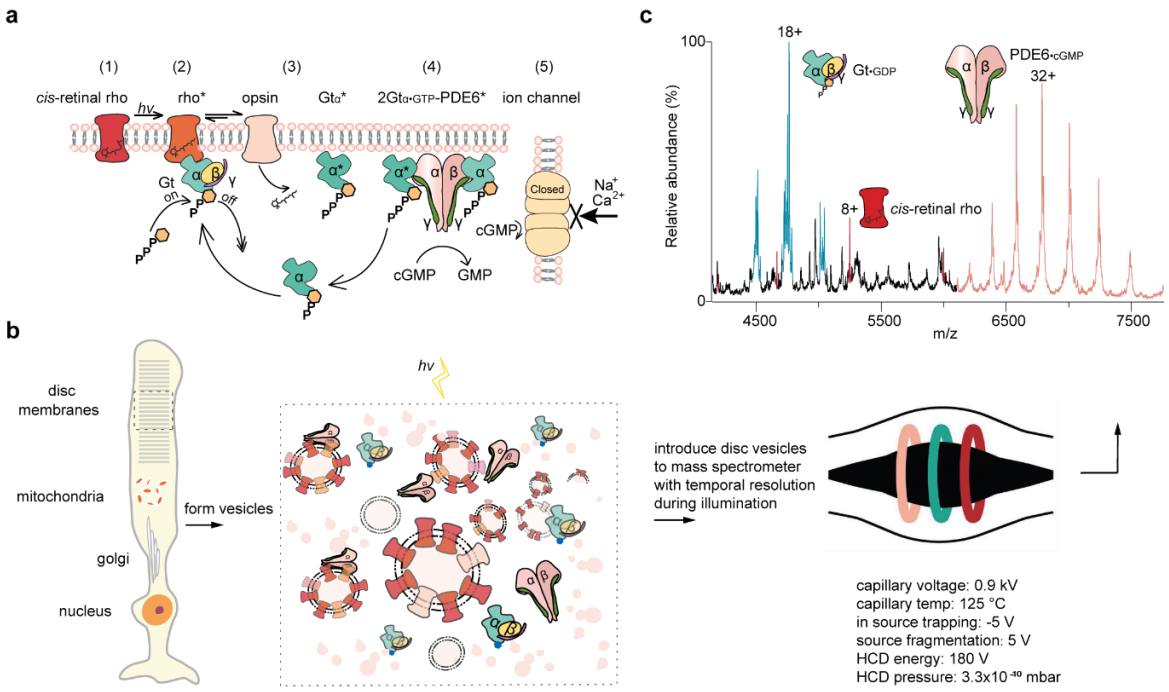


Figure 4.1 Established signalling pathway of rho, experimental conditions for preparation of vesicles for GPCR signalling and representative mass spectrum. (**a**) Following absorption of photons *hv* (1), 11-*cis*-retinal of rho isomerizes to the all-*trans* form (2). The activated states of rho (rho*) engage with transducin (G_t), consisting of $G_t\alpha\beta\gamma$•GDP, and exchange GDP for GTP. G_t dissociates to form $G_t\alpha_{GTP}$ and $G_t\beta\gamma$; loss of retinal from rho* leads to the formation of opsin (3). α-Subunits of G_t interact with the γ-subunits in the PDE6 enzyme, with γ-subunits undergoing a conformational change, relieving inhibition and thereby activating PDE6 to cause hydrolysis of cGMP (4). Depletion of cGMP then closes the ion channel and the 'dark current' is terminated (5). The resulting change in the membrane potential produces the sensation of light. (**b**) Disc membranes of rod cells are homogenized to form a heterogeneous distribution of vesicles that are introduced directly into the mass spectrometer recorded using the parameters stated above. An LED light source is configured to apply timed light intervals before the electrospray ionisation of the vesicles into an Orbitrap mass spectrometer. HCD, higher-energy collisional dissociation. (**c**) Following the addition of a soluble fraction

containing PDE6 and G_t, all proteins along the signalling pathway were detected. The proteins were ejected intact as rho/opsin (red), trimeric G_{tGDP} (cyan) and tetrameric $PDE6_{cGMP}$ (pink). The mass spectrum shown from *m*/*z* 4,000 to *m*/*z* 8,000 represents the raw data, which were repeated three times and shared similar distribution of these three species.

When illuminated, activated all-*trans*-retinal rho signals through Gt, releasing GDP and dissociating to form Gtα•GTP for interaction with PDE6. To monitor real-time signalling through transducin we monitored the decrease in GtGDP:Gt as a function of time, initially for periods of up to 1 min **(Figure 4.2c**), but found little change beyond the first 15 s of light stimulation. The spectra of Gt are complicated by the presence of additional proteoforms (α2 from cone Gt (**Supplementary Figure 4.1a-c**). Control experiments in the absence of light or ROS membranes (**Supplementary Figure 4.1d-f**) confirm that light activation of rho in the membrane is required to decrease population of G_{tGDP} during signaling (**Figure 4.2a**).

We therefore added a soluble fraction containing Gt and PDE6 to our membrane preparations in a 1:2 ratio. Prior to illumination, GDP-bound heterotrimeric G_{tGDP} and cGMP bound heterotetrameric PDE6 ($PDE6_{cGMP}$) were observed as the predominant species (**Figure 4.1c**). To explore the relationship between rho* Gt signalling and $PDE6_{cGMP}$ hydrolysis we investigated three different GTP conditions (i) with endogenous levels of GTP in the soluble fraction (ii) following addition of a molar equivalent of GTP and (iii) in the presence of a molar equivalent of non-hydrolysable GTPγS. After illumination G_{tGDP} levels decreased in all three cases but were replenished in (ii) the supplementary GTP experiment (**Supplementary Figure 4.1g-i**). Linking Gt signalling with $PDE6_{cGMP}$ hydrolysis, we monitored release of the

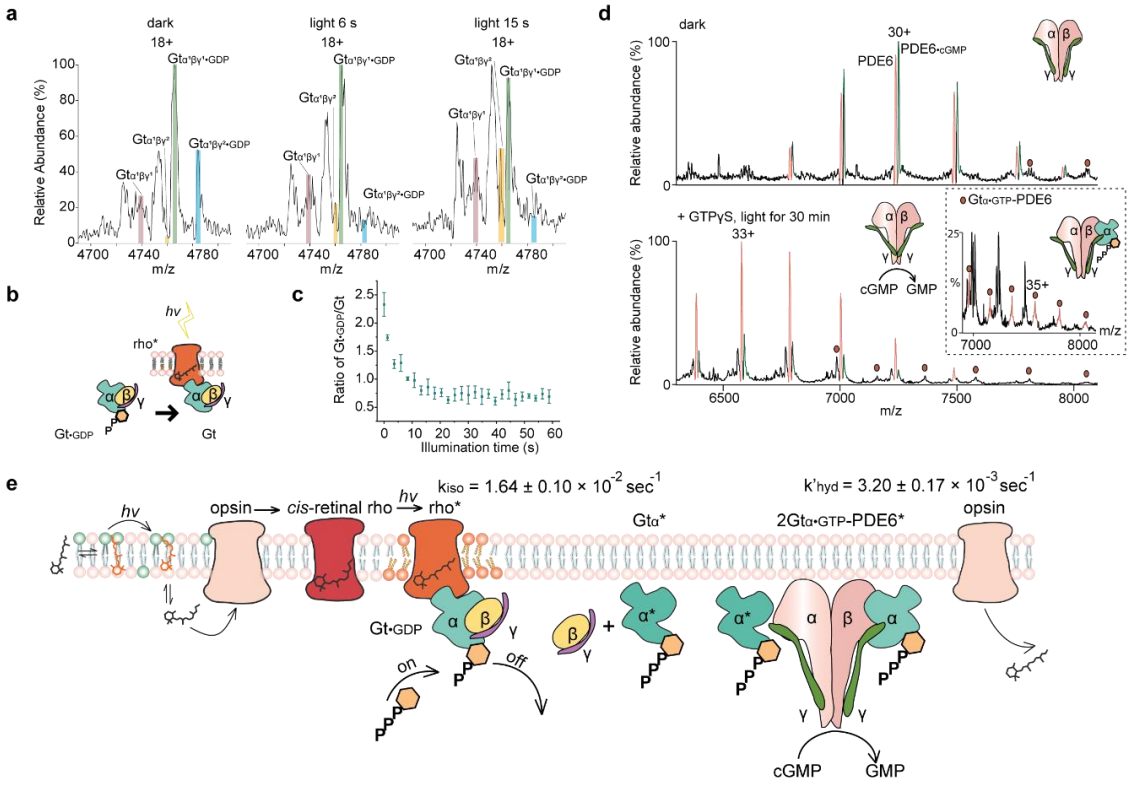


Figure 4.2 Light activation of ROS disc membrane signalling through Gt to release Gtα*GDP to effect hydrolysis of cGMP and its release from PDE6. (**a**) Native mass spectra of Gt_{GDP} with no additional GTP added containing two α isoforms α_1 and α_2 and two γ isoforms γ_1 and γ_2 are catalysed to apo Gt during a 15 s time course of illumination. (**b**), Reaction scheme: rho* catalysed nucleotide exchange of G_t. (**c**) The ratio of ground-state Gt_{GDP} to intermediate apo-Gt decreases rapidly during the first 15 s, post illumination. (**d**) Upper, native mass spectra of PDE6 ejected from ROS disc membrane vesicles under dark conditions reveals that cGMP-binds intact tetrameric PDE6 with ~1:1 ratio. A low population of PDE6:$Gt\alpha^*_{GTP}$ is also observed (red circles). Lower, after exposure to light in the presence of ROS disc membranes, the population of PDE6 bound to cGMP is reduced considerably, consistent with release of GMP following its hydrolysis via activated PDE6. (**e**) Schematic shows the rho* signalling cascade that involves the regeneration via conjugation of all-trans-retinal with PE and its light activated conversion of rho to an intermediate state. Data are presented as mean ± s.e. (n = 3).

takes place with rho* capable of interaction with Gt. Changes in the lipid bilayer are depicted as unsaturated lipids are recruited during rho* signalling (lipids, orange). The $Gt\alpha^{*}_{GTP}$ subunit produced following hydrolysis of Gtα.GDP interacts with PDE6 relieving its inhibition via the gamma subunit and effecting the hydrolysis of cGMP which is then released from PDE6. $Gt\alpha_{GDP}$ is formed for regeneration of Gt_{GDP}. All experiments shown in this figure were repeated at least three times. Data are presented as mean ± s.e. (n = 3).

---

hydrolysed GMP product via the ratio PDE6:$PDE6_{cGMP}$. For dark-adapted membranes, prior to illumination, an PDE6:$PDE6_{cGMP}$ ratio of ~1:1 was observed consistent with full occupancy of one substrate binding site in PDE6 (**Figure 4.2d upper**). Signalling of rho* through Gt prompts further hydrolysis of cGMP with supplementary GTP (the ratio of PDE6 to $PDE6._{cGMP}$ at ~1:0.25) compared to endogenous levels (the ratio of PDE6 to $PDE6._{cGMP}$ at ~1:0.85) (**Supplementary Figure 4.1g, h**). Since $Gt\alpha_{GTP\gamma S}$ is also able to interact with PDE6[326], an intermediate level of hydrolysis of cGMP was observed in the presence of an equimolar aliquot $Gt\alpha_{GTP\gamma S}$ (PDE6:$PDE6_{cGMP}$ of ~1:0.4) (**Figure 4.2d lower and Supplementary Figure 4.1i**). An additional $PDE6.Gt_{\alpha.GTP}$-complex with 1:1 stoichiometry can also be discerned at low intensity in the presence of $Gt\alpha_{GTP\gamma S}$ (**Figure 4.2d inset**). Under these experimental conditions 1:2 complexes of PDE6-$(Gt\alpha_{GTP})_2$ and $PDE6_{(cGMP)2}$ were not observed, consistent with existing mechanistic models. According to one model, although controversial[327], one catalytic subunit of PDE6 binds $Gt\alpha_{GTP}$ with high affinity, but low activity for hydrolyzing cGMP, while the other catalytic subunit of PDE6 binds $Gt\alpha_{GTP}$ with low affinity but is able to hydrolyse cGMP at a high rate[239]. Our data, with only one $Gt\alpha_{GTP}$ subunit binding to PDE6 and with high

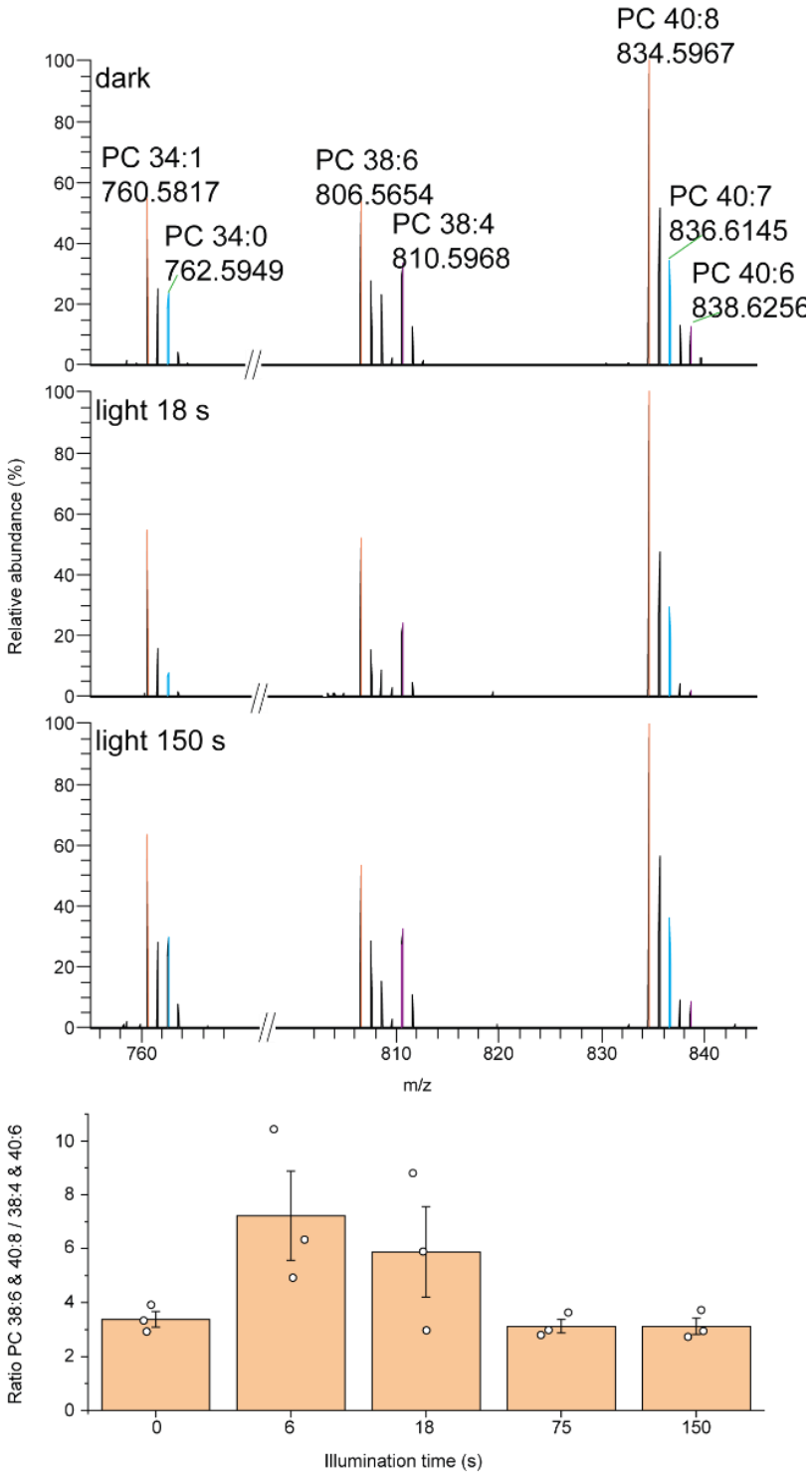


Figure 4.3 Monitoring the lipids that are released with rho while signalling through Gt over the illumination period from dark to light. (**a**) PC lipids released as rho and opsin are ejected from disc membranes during signalling through Gt. Unsaturated lipids are designated as PC

34:1, 38:6 and 40:8 (orange) their more saturated counterparts 34:0, 38:4 and 40:6 (blue/purple). A decrease in the saturated lipids (blue and purple) relative to unsaturated is observed at the early stages of illumination (<18 s). Lipids were detected in the positive ion mode. (**b**) The intensities of PC 38:6 and 40:8 were manually selected to compare with PC 38:4 and PC 40:6 respectively. Plotted is the ratio of unsaturated over saturated lipids as a function of illumination time. After 6 s and 18 s of illumination the extent of unsaturation increases. Data are presented as mean values ± s.e. (n = 3) from three independent experiments. Between 0 s and 6 s, P=0.3146, not significant; two-tailed unpaired Student's t-test. This is due to the low abundance of PC lipids releases from rhodopsin and detected.

---

intensity for a single binding site of cGMP, is in accord with the asymmetry of the activated complex and its ability to hydrolyse one cGMP to GMP per PDE6 tetramer. This reaction in turn is regulated by available GTP, activation of rho* through Gt and causes decreased cGMP concentrations for ion channel closure. Having established a means of capturing the entire signalling cascade we considered further the influence of the membrane through the effect of lipid microenvironments on rho* signalling through Gt. We examined lipids in the spectrum of dark-adapted rho expelled from membranes in the presence of Gt and PDE6 (see **Methods**) and found a predominance of long-chain polyunsaturated phosphatidylcholine lipids (PC 40:6/40:8, 38:6/38:8, 34:1/34:2) in line with observations of distinct regions of retinal rods in nanodisc experiments[264]. After exposure to light (9 -18 s) in the presence of Gt and PDE6, we observed a notable increase in the ratio of unsaturated to saturated PC lipids compared with initial and later time points (**Figure 4.3**). Initial contact with unsaturated lipids is in line with findings of a molecular dynamic simulation study of a GPCR in an inactive or partially active

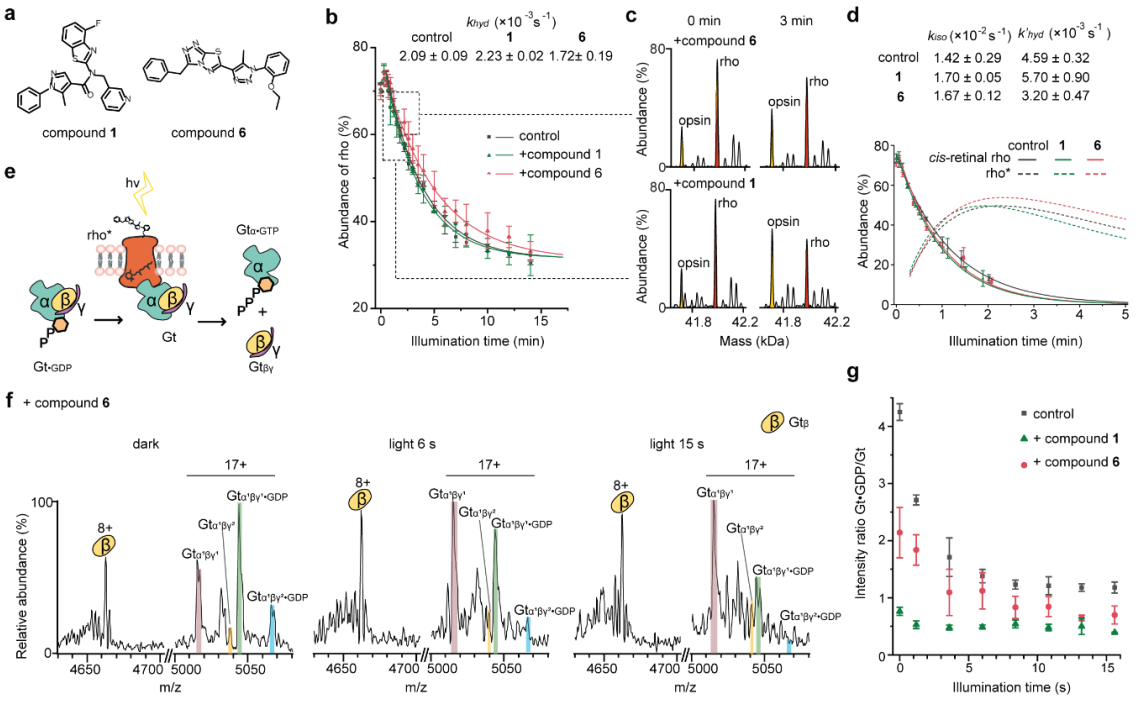


Figure 4.4 The effects of rho targeting compounds on the rho/opsin conversion and hydrolysis of Gt_{GDP}. (**a**) Structures of the two rho targeting compounds **1** and **6** used here. (**b**) Monitoring the conversion of rho to opsin via the change in mass following illumination in the presence of **1** or **6** and in a control. Between control and **1**, control and **6**, **1** and **6**, P=0.2035, 0.1532, 0.0558, not significant respectively; two-tailed unpaired Student's t-test. (**c**) Time points for spectra of rho (red) opsin (yellow) ejected form membranes at 0 and 3 min in the presence of **1** and **6** are shown on a zero-charge scale. (**d**) Plot of the abundance *cis*-retinal rho and rho* as a function of time in the absence (grey) or presence of **6** (pink) or **1** (green). The function for *cis*-retinal rho is derived from fitting data from > 3.6 s (dashed lines) and the function for rho* is derived from the fit of the data from 18 to 300 s (solid lines) and extrapolated (dotted lines). Between control and **1**, control and **6**, **1** and **6**, for k_{iso} P=0.3952, 0.4703, 0.8288, not significant; for k_{hyd} P=0.3098, 0.0709, 0.0695, not significant separately; two-tailed unpaired Student's t-test. (**e**) Reaction Scheme: following activation with light rho* signals through Gt to the nucleotide-free form (Gt) which then dissociates to form βγ and α. (**f**) Mass spectra of

Gt in the presence of **6** added to the native membrane in the absence of light (dark), and after exposure to light for 6 s and 15 s. (**g**) Monitoring signalling through G_t via changes in the ratio of Gt_{GDP} to Gt in the presence or absence of **1** or **6** as a function of illumination time. Three independent experiments are performed and data are presented as mean ± s.e. (n = 3).

---

state receptor[280] and NMR experiments[328]. The increase in contact with unsaturated lipid chains, as opposed to saturated lipids, during light activation provides evidence to support earlier proposals[280] that polyunsaturated chains associate specifically with rho* to enable its conformational change from its ground state and thereby facilitate swift signalling.

Monitoring signalling in a membrane environment provides us with a unique opportunity to probe the effects of rho-targeting molecules, identified through cell-based assays designed to identify compounds that enhance or perturb rho dimerization[324,325] (see **Methods**). We tested nine compounds individually and measured their k_{hyd} against a control (compounds **1** – **9**, **Supplementary Figure 4.2a, b**). The results allowed these ligands to be divided into two groups, capable of either accelerating with a marginal effect or, more commonly, slowing retinal hydrolysis. **1** and **6** represent these two capabilities and were selected for further study (**Figure 4.4a**). Clear differences emerge post illumination (3 min) since rho and opsin predominate in the presence of **1** and **6** respectively (**Figures 4.4b, c**). After 15 min, noticeable differences occur in the rho/opsin conversion which is faster when modulated by **1** and **6,** with **1** being marginally faster than the control. Since we see no evidence for displacement of retinal by these compounds or changes in the conformation of rhodopsin[325], we speculate that **1** and **6** act as allosteric modulators even though the effect is not significant, probably due to the complexity of membrane .

Calculating rates of hydrolysis and isomerization in the presence of hydroxylamine (**Figure 4.4d**) we find that **1** accelerates both hydrolysis and isomerization while **6** has an accelerated isomerization rate but retains retinal by decelerating hydrolysis, thereby potentially maintaining active signalling states for extended periods.

Anticipating that the prolonged activation state of **6**, and the rapid isomerization and hydrolysis of **1** would affect signalling differently through Gt, we established a further assay in which we measure the ability of **1** and **6** to affect phototransduction. As above we monitor their impact by measuring the change in mass of the conversion of Gt_{GDP} to apo Gt (**Figure 4.4e**). Prior to illumination, in the presence of **1** or **6** and a control under dark conditions the ratio of Gt_{GDP} to Gt is 1:0.8, 1:2.1, 1:4.3 respectively. In the presence of the ROS membrane both ligands had considerably effected conversion of Gt_{GDP} to Gt compared to the control (**Figure 4.4f** and **Supplementary Figure 4.3**). Unexpectedly, **1** and **6** induce transduction in the dark, implying that their enhanced isomerization rates enable them to signal through the initial *trans*-retinal rho population prior to illumination. Once illuminated, phototransduction proceeds further (Gt_{GDP}:Gt ratios of 0.4, 0.7 and 1.2 after 15 s for **1**, **6** and a control respectively, **Figure 4.4g**). We conclude that **1** and **6** target rho and amplify signalling through Gt making them ideal starting points for further chemical development. More generally, the fact that we can separate these reactions into isomerization, hydrolysis and influence on a downstream effector, enables us to study rho targeting compounds in molecular detail via their effects on phototransduction, in real-time and within their native membrane environments. Combined with the result from Chapter 3 that rho photocoversion is significantly different in membrane and detergent micelles, evaluation of rhodopsin- and other GPCR-targeting compounds in native membrane is better to be done in native membrane as these compounds screened by conventional assays in vivo

and in detergent cause a milder influence than in detergent micelles. This suggests the importance of choosing a appropriate membrane mimetics for drug discovery.

4.4 Conclusion

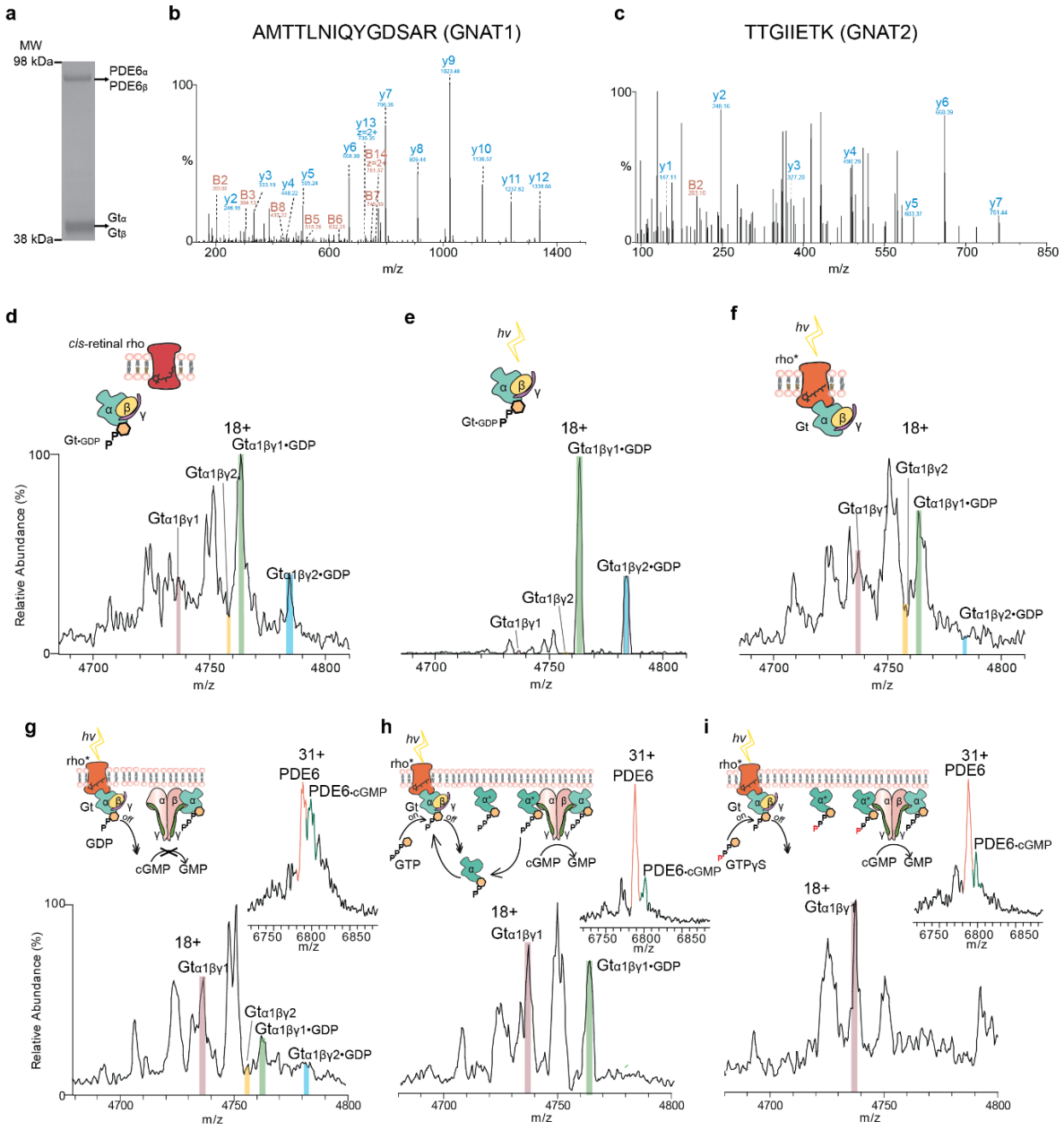
Overall, our results with native mass spectrometry show that we have conserved a signalling cascade for an archetypal class A GPCR across its native membrane. We have shown that all three key players in the cascade can be ejected simultaneously from vesicles formed from native ROS disc membranes. The importance of the membrane also prompted us to consider the changing lipid microenvironment during signalling and documented changes in the association of unsaturated long chain PC lipids in the vicinity of signalling rho.

Considering the effects of rho targeting compounds, we demonstrated their ability to considerably accelerate the rho/opsin conversion or slow down the reaction and hypothesize that the latter would stabilize a signalling competent state. In line with this hypothesis, we demonstrate the ability to modulate the Gt signalling pathway through an increase in the turnover of Gt_{GDP} to Gt and the subsequent dissociation to the Gtβγ complex and $Gt\alpha_{GTP}$ primed for interaction with PDE6. We observed the static intermediate Gtα-PDE6, as well as the effects of endogenous GTP levels on the hydrolysis of cGMP through light activated signalling across the membrane.

By capturing the entire signalling process, we have demonstrated the importance of the lipid micro-environment for signalling, coupling, effector targeting and regeneration. We anticipate that similar approaches will be applicable to other cell surface receptors (e.g. smell receptors) present at high density enabling kinetic studies of molecular events, and identification of endogenous and synthetic receptor ligands that perturb signalling. Therefore, we have

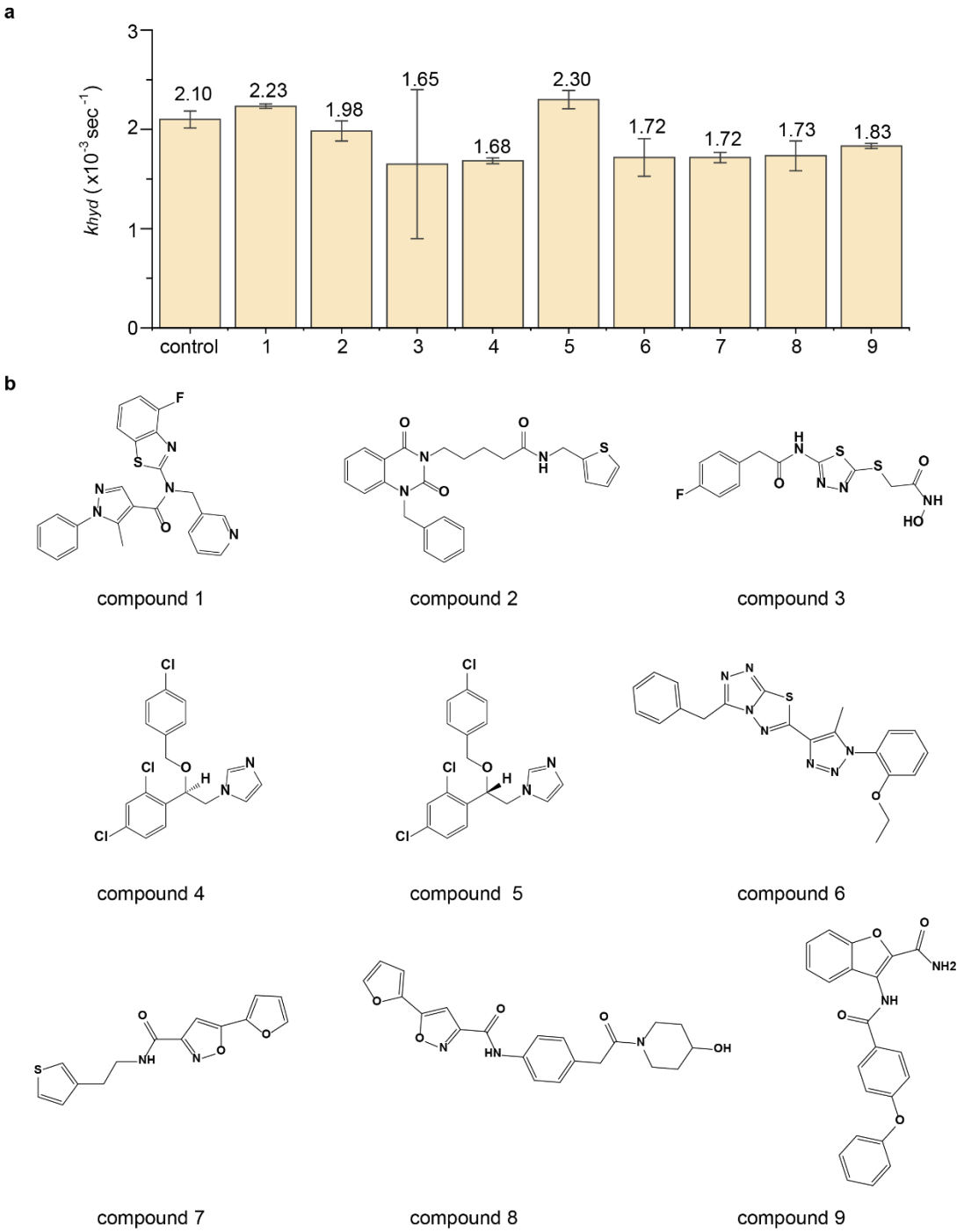
highlighted a new approach to drug discovery in which different stages of a signalling cascade can be targeted across native membranes.

4.5 Supplementary Information

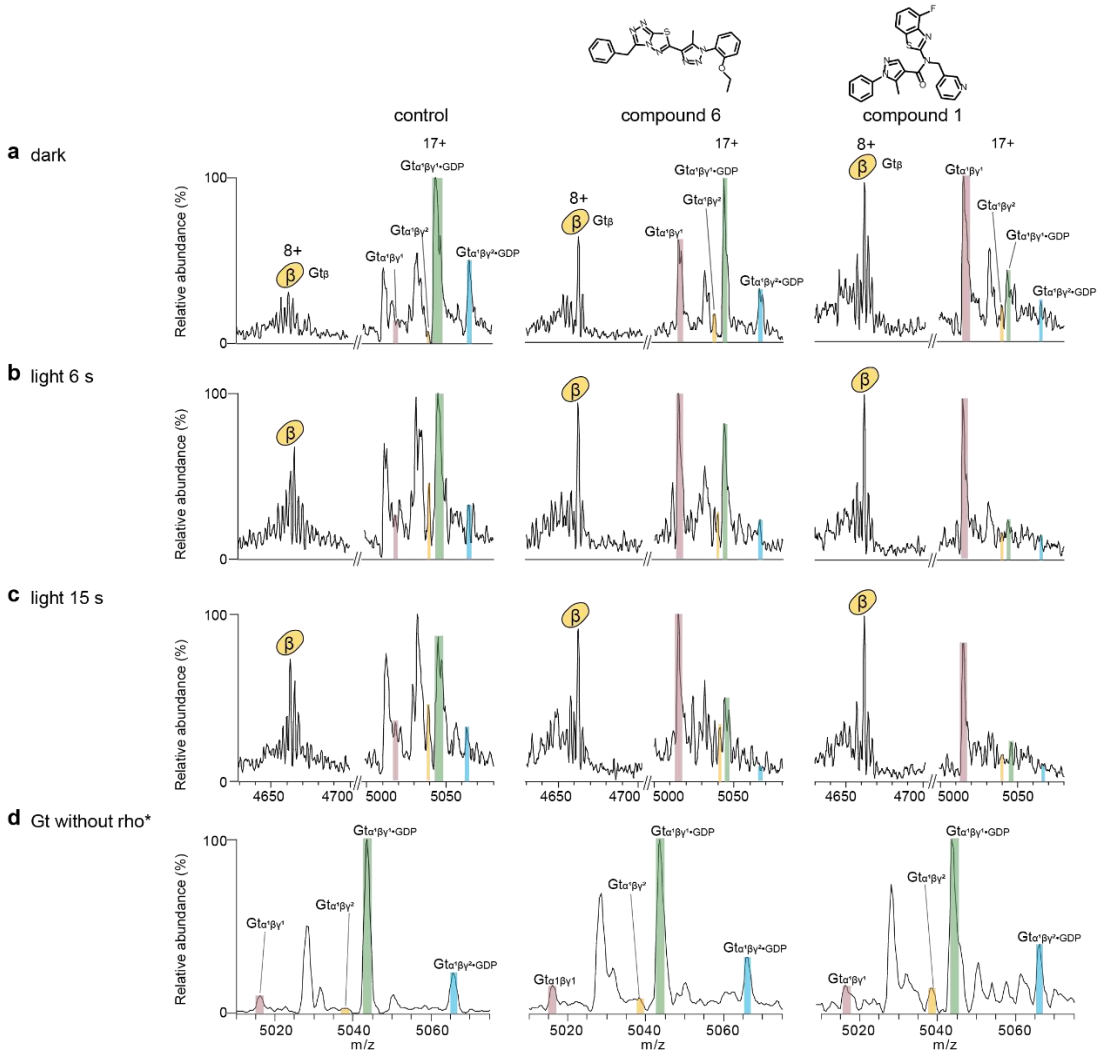


Supplementary Figure 4.1 In-gel digestion proteomics and native mass spectra of the Gt protein complex isolated from dark-adapted rod outer segments under different conditions. (**a**) Separation of the component subunits of the Gt and PDE6 complex on a 2D page gel, pre-stained with protein standards. (**b**) Unique peptides of Gα1 (GNAT1) and (**c**) Gα2 (GNAT2) are identified by proteomics (see **Methods**). (**d**) Gt with rho-containing membrane in the absence of light reveals

the initial extent of endogenous GDP binding to Gt (blue and green for γ^1 and γ^2 isoforms respectively) and without GDP (pink and yellow). (**e**) Illuminated in the absence of membranes with Gt bound to GDP predominant for both isoforms. No reaction is observed in the absence of the membrane. (**f**) with rho in membranes following illumination, GDP bound forms have declined in favour of *apo* states (pink and yellow). (**g**) with endogenous levels of GTP, the PDE6:PDE6cGMP complex remains ~ 1:1 after activation of Gt though rho*. The endogenous quantity of GTP is insufficient to produce sufficient $G\alpha_{.GTP}$ to interact with PDE6 and promote hydrolysis of cGMP. (**h**) in the presence of an additional molar aliquot of GTP, Gt_{GDP} is replenished and further cGMP hydrolysis takes place. **i** in the presence of an additional molar aliquot of GTPγS, further cGMP hydrolysis has occurred but Gt_{GDP} is not replenished. The in-gel proteomics experiment was performed once and the native MS experiments were repeated three times. Data are presented as mean ± s.e. (n = 3).



Supplementary Figure 4.2 The light induced k_{hyd} of rho in the presence or absence of rho regulators. The control was ROS rho vesicles in 200 mM ammonium acetate with 0.1% DMSO at pH 7.0. **1 - 9** were solubilised in DMSO and diluted to 100 µM in ROS rho vesicles in 0.1% DMSO, 200 mM ammonium acetate, pH 7.0. Data are from three replicate experiments and presented as mean ± s.e. (n = 3).



Supplementary Figure 4.3 Gt activation across native membranes during illumination of a control and in the presence of 1 and 6. The ratio of GDP bound Gt to intermediate Gt represents a measure of the activation of transducin. Upon light absorption via rho activated G protein releases GDP in the dark **a**, at 6 s **b,** and after 15 s of the reaction **c**. In the presence of **1** or **6** G protein signalling was accelerated and commenced even under dim red-light conditions when compared to the control experiment. Two different control experiments were performed in the absence of compounds (left hand column) but with the same concentration of DMSO used to dissolve **1** and **6**. **d,** The second control in the presence of the same concentration of DMSO, and **1** and **6**, but in the absence of rho* establish that the agonists do not compete with the nucleotide

(GDP) and that signalling occurs through rho* in the native membrane. The experiment was repeated three times.

Chapter 5 Native Mass Spectrometry Meets Glycomics – Resolving Structural Detail and Occupancy of Glycans on Intact Glycoproteins

Adapted from a publication in Anal. Chem, 2021,

Siyun Chen, Di Wu, Carol V. Robinson and Weston B. Struwe*

Physical and Theoretical Chemistry, Department of Chemistry, University of Oxford, OX1 3QZ, Oxford, United Kingdom

*Corresponding Author

Graphical abstract

Acknowledgements

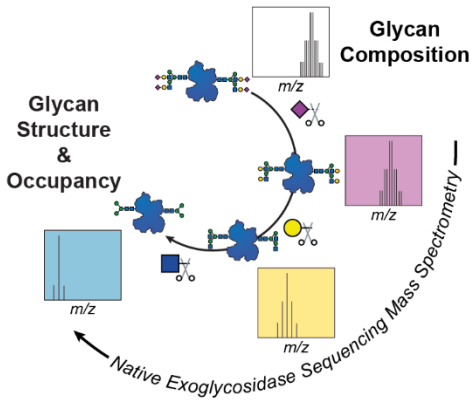
We recognize support from the Medical Research Council (MRC) (program grant MR/N020413/1) and the European Research Council (Advanced Grant ENABLE 695511 and Consolidator Grant PHOTOMASS 819593).

Author contributions

S.C. and W.B.S. designed the study. S.C. optimized and performed the sample preparation and native mass spectrometry experiments. W.B.S performed glycan mass spectrometry experiments. S.C. analyzed and interpreted the native MS data with the help of W.S. and D.W. . All authors contributed to writing the manuscript. W.B.S. and C.V.R. conceived and supervised the project.

Abstract

Glycoproteins are inherently heterogeneous and therefore resolving structures in their entirety remains a major challenge in structural biology. Native mass spectrometry has transformed our ability to study glycoproteins. Despite advances in high-resolution instrumentation, comparatively few studies demonstrate its potential with data for an overall measure of monosaccharide composition for all glycans across glycosylation sites for a given protein.



Clearly, these readouts lack glycan topology information, namely monosaccharide linkage and glycan branching. To address this deficiency, we developed a new approach that joins native mass spectrometry with glycan exoglycosidase sequencing — the combination of which provides remarkable glycoprotein structural detail. We show how N-glycan branching, terminal fucosylation, LacNAc extensions as well as N- and O-glycan occupancy (i.e. total number of glycans) can be directly characterized on intact glycoproteins with minimal sample preparation. Taken together, native exoglycosidase sequencing mass spectrometry (NES-MS) notably improves our ability to characterize protein glycosylation, addressing a significant need in structural biology that will enable new routes to understand glycoprotein function.

5.1 Introduction

Proteins modified with oligosaccharides, or glycans, are widespread among eukaryotes and act to fine-tune protein function and cell signaling. Over 50% of all human/mammalian proteins are predicted to be glycosylated[329] and importantly, changes in glycosylation are observed in nearly all disease states, leading many to identify glycosylation changes as markers for disease[47]. Mass spectrometry (MS) is the leading method to characterize protein glycosylation, typically through analysis of isolated oligosaccharides or glycopeptides generated by enzymatic digestion (e.g. trypsin). These strategies provide two key sets of information. First, glycan analysis provides a detailed picture of topology, branching and linkage of structures present on a glycoprotein. Second, the analysis of glycopeptides informs on the site-specific glycan composition as well as the degree to which a glycan is present at a given protein glycosite. These characteristics are known as macroheterogeneity (i.e. the presence or absence of glycans at a specific glycosite) and microheterogenity (i.e. differences in glycan structure at a specific glycosite). Generally, each analytical approach comes at the cost of the other; for example, when measuring released glycans the site information is obviously lost. Conversely, glycopeptide approaches do not provide comprehensive glycan structural detail. Only recently, and principally through the development of high-resolution instruments, native MS has created new opportunities for studying glycoproteins[114,116,126,330-333].

Native MS is highly versatile for studying protein assemblies, enabling the analysis of stoichiometry, dynamics, ligand binding and post-translational modifications (PTMs) including phosphorylation, acetylation and glycosylation[334]. Its use for glycoproteins is rapidly growing, with published reports demonstrating its value for characterizing protein-protein interactions[126], glycan-specific drug binding[116], lectin recognition[116], glycolipid-protein interactions[159,332,333], O-GlcNAcylation kinetics[335], biotherapeutic stability[336], similarity[337] and glycoengineering[338]. Intact glycoprotein analysis is complicated by the diversity of glycan types (e.g. N-glycosylation, O-glycosylation, O-GlcNAcylation) as well as microheterogeneity and/or macroheterogeneity within each type. Consequently, the principal output from many native MS glycoprotein analyses is limited to monosaccharide composition across the entire protein and a mass distribution of glycoproteoforms that are present.

The main challenge with analyzing glycans on intact glycoproteins is the combined diversity across glycosites and overlap in composition (i.e. the presence of structural isomers) between different glycan structures or types. For example, a single *N*-acetylhexosamine residue (HexNAc, 203 Da) can be present as an O-GlcNAc, a GalNAc on an O-glycan or a bisecting GlcNAc on an N-glycan. In principle, native MS can differentiate classes of glycans but this typically involves removal of one type, such as enzymatic cleavage of N-glycans via PNGaseF, to reveal the presence of others[330]. Such steps are informative, and have enabled the identification of N-glycosylation on Asn at an Asn-Ala-Ala sequence, not the recognized Asn-X-Ser/Thr sequon[339]. Of course, the removal of N-glycans in these examples prevents their

structural interrogation in an intact approach. Therefore, most native MS studies of glycoproteins are complemented by peptide-based methods (i.e. glycoproteomics to identify glycan diversity at a specific site) to validate assignments, confirm occupancy and site of attachment[330,337]. Such orthogonal strategies are indeed powerful; however, they require extensive sample workflows and various mass spectrometry setups. Ideally, native glycoprotein analyses will provide information on both glycan structure and glycan occupancy in a straightforward method.

Conceivably, the route towards uncovering microheterogeneity, and even macroheterogeneity, by native MS will exploit existing glyomics tools, namely enzymatic digestion prior to MS profiling[340]. Exoglycosidases cleave individual monosaccharides on a given glycan that are located distal from the reducing end residue. Exoglycosidase sequencing of released glycans is routinely used in MS and HPLC analyses but the application of glycan digesting enzymes in a native MS approach has not been explored beyond the removal of terminal sialic acids[126], β1-4 galactose[114] or the use of endoglycosidases for removal of whole N- or O-glycans by PNGaseF or O-glycanase, respectively[330,337].

Exoglycosidases are linkage and residue specific, so to uncover glycan structures it is constructive to combine enzymes sequentially or in thoughtful combinations based on well-known principles of glycoconjugate biosynthesis. In this way, glycan topology and linkage information are revealed by relating changes in glycoprotein mass with known enzyme

specificities (**Figure 5.1a**). Here, we demonstrate how native exoglycosidase sequencing mass spectrometry **(**NES-MS) provides significant glycan structural detail as well as quantitative occupancy information, namely the total number of glycans, on intact glycoproteins in a simple technique. Specifically, we show how NES-MS can resolve N-glycan terminal fucosylation, LacNAc extensions, branching as well as identification of O-glycan core types on several glycoproteins.

5.2 Methods

Materials

Human apo-transferrin (product code: T2036), bovine fetuin (product code: F3004), human α1-acid glycoprotein (product code: G9885) and NuPAGE™ LDS sample buffer (4X), ammonium acetate 7.5M solution were from Sigma-Aldrich. Eprex® (epoetin alfa) erythropoietin (EPO) was from Janssen Pharmaceutical. α2-3, 6, 8, 9 neuraminidase A (product code: P0722S), β1-4 galactosidase S (product code: P0745S), β-N-acetylglucosaminidase S (product code: P0744S), O-glycosidase (product code: P0733S), α1-3,4 fucosidase (product code: P0769S) and α1-2, 3, 4, 6 fucosidase (product code: P0748S) were purchased from New England Biolabs. PNGaseF was expressed and purified in-house.

Native exoglycosidase digestion

Intact glycoproteins were desialylated using α2-3, 6, 8, 9 neuraminidase A (ABS). For each digestion, ~18 μg of glycoprotein was digested with 80 U ABS in 50 mM sodium acetate (Fluka Analytical, Steinheim, Germany) pH 5.5, in a final volume of 24 μL. The samples were incubated for 24 h at 37 °C while shaking. Following sialic acid removal, each was then digested using 32 U β1-4 galactosidase S (BGS) with or without 16 U of β-N-acetylglucosaminidase S (GUH) or with 16 U of α1-3,4 fucosidase (AMF) or O-glycosidase.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Exoglycosidase activity and glycan cleavage of human transferrin, bovine fetuin, EPO and α1-acid glycoprotein were confirmed by SDS-PAGE prior to native MS analysis (**Supplementary Figure 5.1**). For each reaction, 9 μL of NuPAGE™ LDS sample buffer was added to 18 μg of each glycoprotein sample. Samples were boiled at 95 °C for 5 min, run on a NuPAGE™ 4-12%, Bis-Tris gel and stained with Coomassie blue.

Native mass spectrometry

Samples were buffer exchanged and sequentially washed in 1 M ammonium acetate, pH 6.9 until adequately desalted using 10 or 30 kDa molecular weight cut-off centrifugal spin filters (GE Healthcare). Samples were diluted to 10 μM and measured using a modified Q-Exactive mass spectrometer (Thermo Scientific). Several μL sample was loaded into gold-coated borosilicate glass capillaries and infused into the instrument via a nanoelectrospray (nESI) ion source. The capillary voltage was +1-1.3 kV, source was 100°C, S-lens RF was 200, maximum

injection time was 100 ms. The in-source trapping, source fragmentation and HCD setting were off. Data was acquired (m/z 1,000-15,000, 5 microscans 5 average) and processed with XCalibur v4.1 (Thermo Scientific). Spectra were deconvoluted with UniDec[308] and annotated manually and with in-house software.

Data interpretation

Average masses were used for PTM calculations, including hexose/mannose/galactose (Hex/Man/Gal, 162.14 Da), N-acetylhexosamine/N-acetylglucosamine/N-acetylgalactosamine (HexNAc/GlcNAc/GalNAc, 203.20 Da), deoxyhexose/fucose (dHex/Fuc, 146.14Da), N-acetylneuraminic acid (NeuAc, 291.26 Da), and phosphorylation (Phos, 79.98 Da). The backbone sequence of Tf, fetuin, EPO and AGP and their reported PTM sites are in **Supplementary Figure 5.2**.

High performance liquid chromatography - mass spectrometry

N-linked glycans were enzymatically cleaved in-solution from 20 µg AGP with PNGaseF. Released N-glycans were fluorescently labelled with 2-aminobenzamide (2-AB) as previously described[341]. Labelled N-glycans were analyzed by hydrophilic liquid interaction chromatography ultrahigh performance liquid chromatography (HILIC-UHPLC) using a Waters Acquity-Xevo G2 instrument (with a 2.1 x 150mm, 1.7 µm particle size BEH glycan column)[341]. Solvent A was 50 mM ammonium formate adjusted to pH 4.4, and solvent B was

ACN. A linear gradient was used, which was 35 to 47% A for 48 min, 47 to 100% A for 1 min, 4 min 100% solvent A, 35 to 100% solvent A for 1 min, and 35% solvent A for 6 min. Samples were diluted in 65% ACN before analysis. The instrument was set to negative mode and MS settings were as follows: capillary voltage, 0.8-1.0 kV; sample cone, 100 V; extraction cone, 25 V; cone gas, 40 l/h; source temperature, 120 °C. N-glycans were measured as $[M-2H]^{2-}$ ions. Data was acquired and processed with MassLynx v4.1 (Waters, Manchester, UK). Glycan data was annotated manually.

5.3 Results and Discussion

Native exoglycosidase sequencing (NES) mass spectrometry (MS) of human transferrin N-glycans

To demonstrate the feasibility and benefits of a NES-MS approach, we first examined a relatively simple glycoprotein, albeit one with notable diagnostic value in disease detection, human transferrin (Tf). Tf is a serum glycoprotein with two N-glycans that regulates iron concentration in the body. Tf is the principal biomarker for human congenital disorders of glycosylation (CDG), a severe disease that is classified as type-I or type-II depending on the cellular location of glycosylation deficiency[342]. CDG subtypes present with different structural characteristics on glycoproteins. A type-I CDG is indicated by Tf with loss one or two N-glycans (i.e. under-glycosylation) while type-II CDG is revealed by a decrease in terminal sialic acid monosaccharides on both N-glycans[343]. The most common diagnostic method for CDG is

isoelectric focusing[344], which is not only time consuming but suffers from irreproducibility. Proteomics is emerging as an alternative[345], but with it comes added sample preparation steps, longer sample run times and involved data analysis.

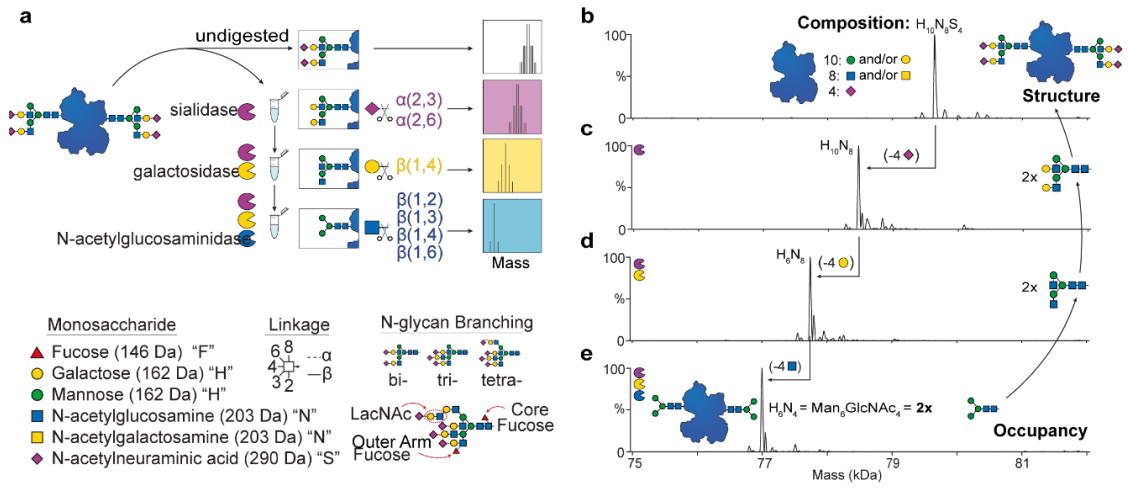


Figure 5.1 (**a**) Schematic of native exoglycosidase sequencing mass spectrometry (NES-MS). (**b**) Native MS spectrum of human transferrin with monosaccharide compositions labelled. Spectra following removal of (**c**) N-acetylneuraminic acid (NeuAc), (**d**) NeuAc and galactose (Gal), (**e**) NeuAc, Gal and N-acetylglucosamine (GlcNAc) residues reveal N-glycosylation occupancy and N-glycan structures on transferrin. **Bottom**: legend for glycan structure nomenclature used in this study.

A native MS approach would address these shortcomings and the simplicity of the data makes its use for CDG screening plausible. The deconvoluted zero-charge state mass spectrum of

undigested Tf contained a single, major peak of 79568 Da. Comparison of this mass to the protein backbone mass (75157 Da), revealed the remaining 4411 Da could be assigned to N-glycans (**Figure 5.1b**). From this information we can calculate what glycans may be theoretically present on Tf. We know that Tf has two potential N-glycosylation sites from the presence of the N-glycan sequon, Asp-X-Thr/Ser (X ≠ Pro), at two positions in amino acid sequence[346,347]. First, after we incorporate the mass of a key conserved feature of N-glycans, the trimannosyl core that consists of 2 N-acetylglucosamine (GlcNAc, blue squares) and three mannose (Man, green circles) and a mass of 892.7 Da, the remaining glycan mass of the two N-glycans is 2626 Da (4411-(2*892.7) Da). Next, single GlcNAc residues (203 Da) are likely linked to the Man residues on the trimannosyl core, which dictates the degree of glycan branching as shown in the **Figure 5.1** legend. Third, GlcNAc residues are extended with galactose (Gal, yellow circles), which are isomers of Man (162 Da). If each Tf N-glycan has two antennary N-glycans (i.e. a bi-antennary structure) with an additional Gal residue per N-glycan antennae, then we would observe a mass shift of 365*4 from 2626 (365 indicating a Gal-GlcNAc), leaving 1166 Da (=2626-1460 Da). Finally, Gal residues can be terminated with N-acetylneuraminic acid (NeuAc) residues, commonly referred to as sialic acid, each having a mass of 291 when conjugated to a glycan. The 1166 Da specified the presence of 4 NeuAc residues on Tf. Therefore, Tf has a potential N-glycan composition of 10 hexose (Hex, H) residues (Gal and/or Man), 8 N-acetylhexosamine (HexNAc, N), here most likely GlcNAcs, but theoretically N-acetylgalactosamine (GalNAc) is possibly elsewhere on the protein, and

finally 4 NeuAc residues (S). However, this type of calculation, which is seemingly how glycoprotein native MS spectra are assigned, makes several assumptions and is in no way absolute in terms of glycan structural assignment.

NES-MS is essentially the experimental equivalent of this theoretical calculation but has the added benefit of distinguishing structural features that would not appear otherwise (discussed further below). Tf is a straightforward example of NES-MS and demonstrates that exoglycosidases can access and trim terminal monosaccharide residues on protein-bound glycans. This observation is important since steric hindrance can preclude glycan-cleaving enzymes, in particular endoglycosidases (i.e. those between the peptide backbone and glycan, for example PNGaseF) on a folded substrate glycoprotein. We did not see any such problems as sequential NES reactions with Tf were complete. Sialidase (herein referred to as ABS) digestion resulted in an 1166 Da mass shift, corresponding to cleavage of four sialic acid resides (291 Da each) (**Figure 5.1c**). Further digestion of de-sialylated Tf (78402 Da, composition = $H_{10}N_8$) with β1-4galactosidase (referred to as BGS) yielded a 77753 Da peak (composition = H_6N_8), arising from the loss of four galactose residues (**Figure 5.1d**). The final digestion with *N*-acetylglucosaminidase S (*Streptococcus pneumoniae* hexosaminidase, referred to as GUH herein) yields Tf with a mass of 76940 Da (a composition of H_6N_4), 1783 Da of which is the remaining two trimannosyl cores ($76940_{\text{truncated glycoprotein}}$ - $75157_{\text{peptide backbone}}$) (**Figure 5.1e**). At this stage, further N-glycan digestion, such as with a mannosidase, does not provide any additional information. The final $Hex_6HexNAc_4$ composition is therefore $Man_6GlcNAc_4$

revealing the presence of two N-glycans as expected. Using this information, we can work backwards and confirm that undigested Tf is 100% occupied with two bi-antennary, di-sialylated structures, effectively converting composition information in **Figure 5.1b**, top spectrum to a fully characterised structure.

Distinguishing N- and O-glycan structures and occupancy on bovine fetuin

Protein glycosylation is often more complex than that found on Tf both in terms of the number, type and structure of glycans. The presence of other types of glycosylation needs to be addressed to establish the utility of NES-MS and how well it performs, for example in differentiating N- from O-glycans. O-glycans are prevalent on proteins[348] and occur on serine/threonine residues but do not require an amino acid consensus sequence like N-glycans. Therefore, we cannot predict the presence or potential number of O-glycans from protein sequences alone. Understandably, most native MS analyses of glycoproteins that include N- and O-glycans only report the overall monosaccharide composition for both as stated above. This is because it is not possible to discern the presence of some O-glycans from modifications on N-glycans as these oligosaccharides are isomeric. A good example is the occurrence of a core 1 O-glycan structure (Galβ1-3GalNAc-Ser/Thr) versus a Galβ1-4GlcNAc- disaccharide on an N-glycan, which contributes to branching as discussed above. However, the products of exoglycosidase digestions should distinguish the two modifications. To explore this, we examined bovine fetuin, a glycoprotein with N- and O-glycosylation as well as phosphorylation.

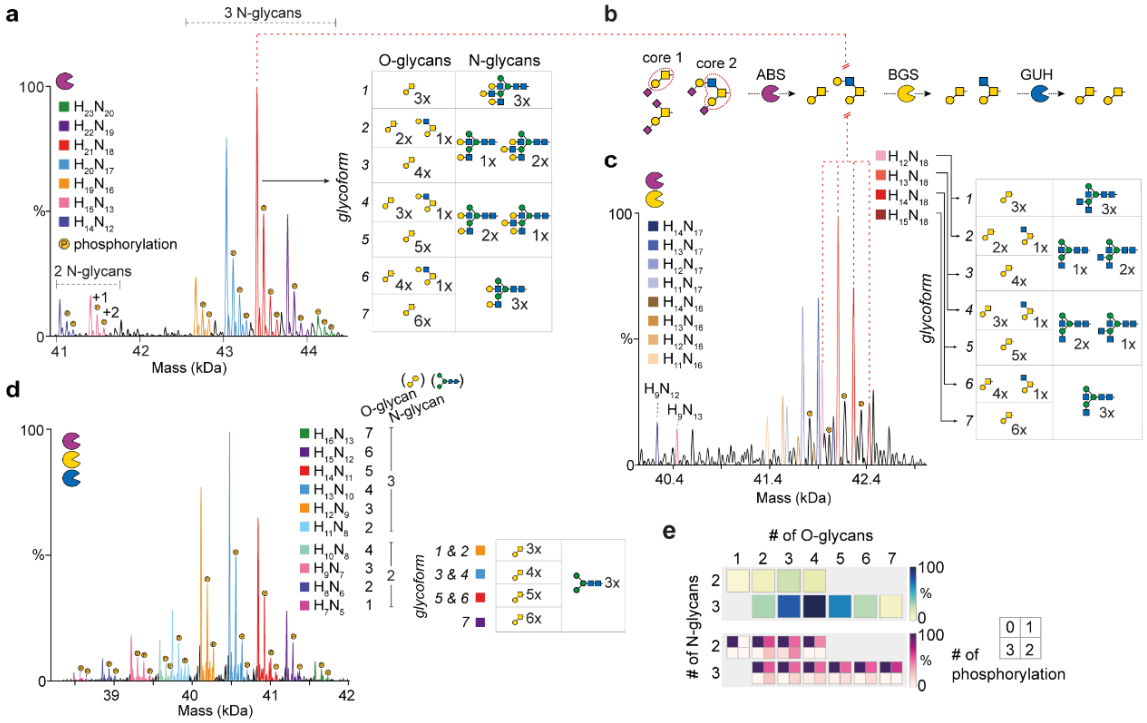


Figure 5.2 NES-MS of bovine fetuin. (**a**) De-sialylated fetuin with seven potential N- and O-glycan glycoforms of the major peak (red). The monosaccharide compositions of all ions are also shown. (**b**) Schematic of the O-glycan exoglycosidase digestion. Spectra of fetuin treated with (**c**) ABS and BGS, and (**d**) ABS BGS and GUH. (**e**) Quantitative heat map of N-glycans and O-glycans (top) with phosphorylation (bottom). H=hexose (Gal/Man), N=HexNAc (GlcNAc/GalNAc).

Fetuin is a good test case as the number of O-glycans has been variably reported, from four to eight, and is based on bottom-up glycopeptide analyses[349-351]. Here, the native MS spectrum of undigested fetuin had two distributions centring around 44325 Da and 47190 Da (**Supplementary Figure 5.3**). This difference corresponded to the approximate mass a single

N-glycan, suggesting fetuin has two principal species, one containing two N-glycans and one with three N-glycans. The complexity of the intact fetuin spectrum arises from the combination and overlap of O-glycosylation, N-glycan branching, sialylation and phosphorylation. After removing the sialic acids with ABS, the heterogeneity was significantly reduced (**Figure 5.2a**) and the two species were clearly visible as well as the presence of one to three phosphate groups per resolved monosaccharide compositional peaks ($H_{14}N_{12}$ to $H_{23}N_{20}$). The fetuin form with two N-glycans had peaks corresponding to two monosaccharide compositions, $H_{14}N_{12}$ and $H_{15}N_{13}$. The higher mass distribution had peaks corresponding to five monosaccharide compositions ranging from $H_{19}N_{16}$ to $H_{23}N_{20}$.

Considering the most abundant peak, $H_{21}N_{18}$, we calculated seven possible glycoforms, differing in the degree of N-glycan branching and the number of O-glycans (**Figure 5.2a, inset**). Again, the presence of a core 1 O-glycan (shown in **Figure 5.2b**) has the same mass shift between a bi-antennary and tri-antennary N-glycan. Thus, *glycoform 1* has three tri-antennary N-glycans and three core 1 O-glycans, whereas *glycoform 7* has three bi-antennary N-glycans and six core 1 O-glycans. The presence of core 2 O-glycans cannot be excluded as they fit with the compositions (i.e. four core 1 O-glycans (H_4N_4) could also be assigned to two core 1 (H_2N_2) plus one galactose extended core 2 (H_2N_2) structures as indicated for *glycoform 2* and *3*. In principle, these species will be revealed following down-stream NES digestion.

After GUH digestion, all β1-4 linked Gal residues on N-glycans and select Gal residues on core 2 O-glycans were removed (shown in **Figure 5.2b**), but the Gal on core 1 O-glycans are not cleaved as they are β1-3 linked and not a substrate for BGS. The result from this digestion was that glycoforms present as a single $H_{21}N_{18}$ peak in **Figure 5.2a** progressed into four different populations ($H_{12}N_{18}$, $H_{13}N_{18}$, $H_{14}N_{18}$, and $H_{15}N_{18}$, shown in red dotted lines, **Figure 5.2c**). We observed shifts from de-sialylated fetuin consistent with galactose removal (yellow circle) from core 2 O-glycans, confirming their presence. Loss of a Gal from *glycoform 2* in **Figure 5.2a** resulted in the same mass as *glycofrom 1* as well as *glycoform 4* with *3, glycoform 6* with *5*, thus initially confirming the seven glycoforms. Overall, the spectrum following NeuAc and Gal removal was less resolved and due to overlapping signals across numerous fetuin glycoforms with an interval difference corresponding to Hex (162 Da) and phosphate (80 Da). The assignment of phosphorylation on fetuin is in line with proteomics data[352].

The peaks resolution was notably improved following GlcNAc removal with GUH (**Figure 5.2d**). Here, all core 2 O-glycans were converted to core 1 structures (i.e. trimming of β1-6 GlcNAc shown in **Figure 5.2b**) and N-glycans were present as trimannosyl cores, $Man_3GlcNAc_2$ as shown for Tf in **Figure 5.1e**. The seven glycoforms present for $H_{21}N_{18}$ in **Figure 5.2a** were confirmed as all mass shifts were consistent with the exoglycosidases used. For example, *glycoforms 1* and *2*, which had different masses due to variation in O-glycosylation structure, became identical in glycan composition after ABS, BGS and GUH digestion, this verifying the presence of three O-glycans and three N-glycans ($H_{12}N_9$, orange

peaks). This was also true for *glycoforms 3* and *4* ($H_{13}N_{10}$), *glycoforms 5* and *6* ($H_{14}N_{11}$) and *glycoform 7* ($H_{15}N_{12}$). Importantly, we could clearly see the distribution of fetuin with two N-glycans plus between one and four O-glycans (H_7N_5, H_8N_6, H_9N_7 and $H_{10}N_8$) as well as fetuin with three N-glycans plus between two and seven O-glycans ($H_{11}N_8$, $H_{12}N_9$, $H_{13}N_{10}$, $H_{14}N_{11}$, $H_{15}N_{12}$ and $H_{16}N_{13}$). We also observed phosphorylation mass shifts from these data and can directly quantify the exact occupancy levels of N-glycans, O-glycans and phosphate, all from the single spectrum (**Figures 5.2d-e**). Therefore, the fetuin evaluated here had between two and three N-glycans (14% and 86 %), one to seven O-glycans (1%, 21%, 25%, 28%, 16%, 7.0%, 2%) and zero to three phosphorylations at 59%, 30%, 10% and 1%, respectively.

Differentiating glycan branching and LacNAcs extended glycoforms on erythropoietin

Next, we sought to characterise a glycoprotein with more complex N-glycosylation, specifically heightened N-glycan branching, fucosylation and N-acetyllactosamine (LacNAc) extensions (Galβ1-4GlcNAc). We chose a recombinant therapeutic form of the human erythropoietin (EPO), a glycoprotein hormone used to treat anaemia that can be caused by chronic kidney disease or chemotherapy[353]. Glycosylation on EPO, namely the degree of sialylation, is important as it affects its half-life[354-356], making it a critical quality attribute in drug development. Most therapeutic proteins are glycosylated and therefore the ability to fully characterize glycosylation on biologics is important as demonstrated by increasing reporting guidelines from regulatory agencies[357].

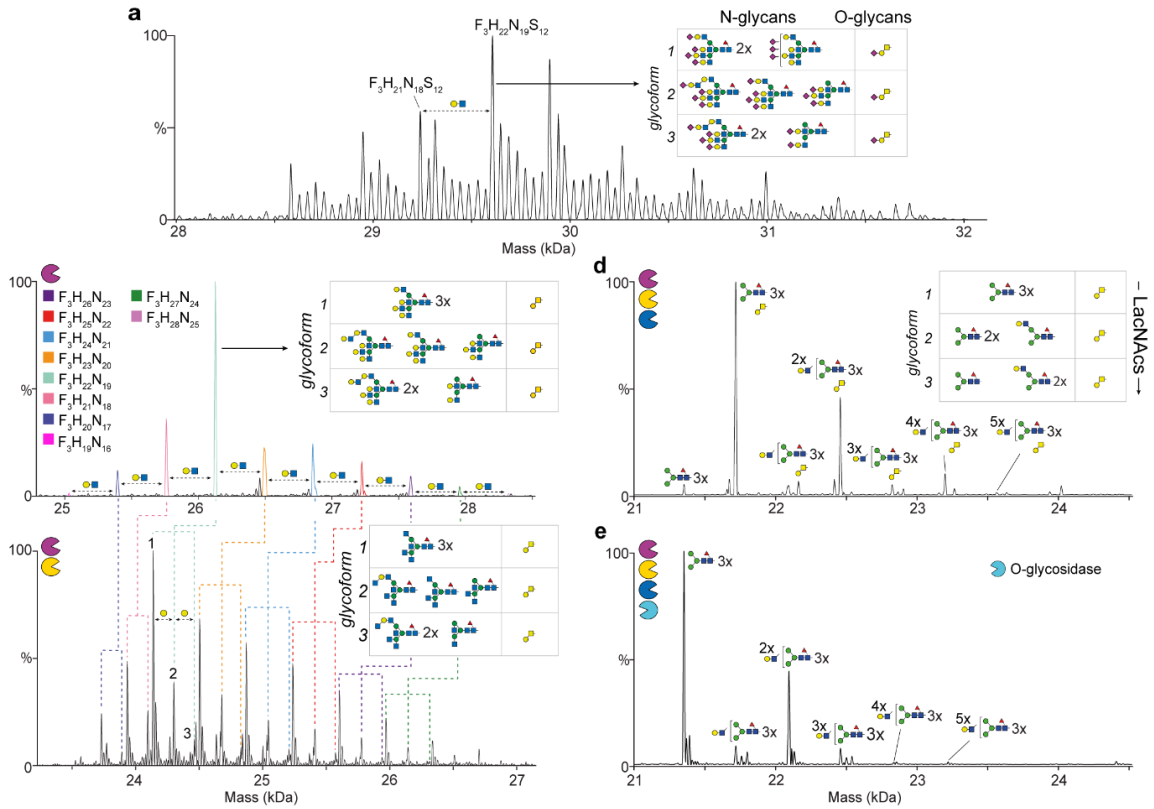


Figure 5.3 NES-MS of erythropoietin (EPO). (**a**) Undigested EPO with the three potential glycoforms of the major peak (shown from a-d). Native MS spectra following NeuAc removal (**b**), NeuAc and Gal removal (**c**), NeuAc Gal and GlcNAc removal (**d**) and NeuAc, Gal, GlcNAc and core 1 O-glycan removal via O-glycosidase (**e**)

Considering its importance in healthcare, EPO has been measured by native MS[337,338], as well as via glycomics/glycoproteomics[358-360]. As such, we know EPO has a single O-glycan and three N-glycosylation sites comprised of complex-type bi-antennary to tetra-antennary structures with different degrees of LacNAc extensions and terminal sialylation. Our native

MS spectra of non-digested and desialylated EPO are similar to published data (**Figure 5.3a-b**) and reveal monosaccharide compositions from $F_3H_{20}N_{17}$ to $F_3H_{27}N_{24}$. The mass difference between peaks is 365 Da, corresponding to a Hex-HexNAc and presumably these are differences in glycan branching or LacNAcs. Differentiating LacNAcs from increased N-glycan branching (tri- vs. tetra-antennary structure) is not straightforward by glycomics and involves generating and detecting diagnostic ions, such as cross-ring A-type or D-type fragments in negative MS/MS approaches[361]. Here, we focus on NES-MS for identifying the N-glycan antennae with LacNAc units.

Intact EPO had numerous peaks corresponding to glycoforms that differed in the number of Hex, HexNAc and Neu5Ac residues, ranging from $F_3H_{20}N_{17}S_{11}$ to $F_3H_{26}N_{23}S_{13}$ (**Figure 5.3a**). However, within each peak and a given monosaccharide composition there contained several potential glycoprotein structures. For example, the predominant peak had a composition of $F_3H_{22}N_{19}S_{12}$, and based on the presence of three N-glycosylation sites, we calculated three putative glycoforms. *Glycoform 1* has one sialylated core 1 O-glycan plus three tetra-antennary N-glycans. *Glycoform 2* has one sialylated core 1 O-glycan plus one tri-antennary, one tetra-antennary and LacNAc-extended tetra-antennary N-glycans. *Glycoform 3* has one sialylated core 1 O-glycan plus one bi-antennary and two LacNAc-extended tetra-antennary N-glycans (**Figure 5.3a, inset)**. After sialic acid removal, the spectrum was notably simplified as NeuAc heterogeneity was removed, and ten glycan compositions were evident (**Figure 5.3b** and **Supplementary Table 5.1**). Still, it is not possible to confirm any glycan structural features

within each peak at this stage. The three major glycoforms from **Figure 5.3a** are isomeric, possessing the same monosaccharide composition, however, after β1-4 galactosidase digestion, the product of each glycoform differs due to the presence of LacNAc extensions. This is because a single LacNAc (Galβ1-4GlcNAc) effectively shields the inner-most Gal present at the penultimate Galβ1-4GlcNAc (i.e. the upper-branch of the tetra-antennary N-glycan as shown in **Figure 5.3c, inset**). Following galactose removal, the presence and abundance of each glycoform was evident; *glycoform 1* was the major species, followed by *glycoform 2* and *3*. The pattern we observed of a single peak splitting into three peaks after galactose removal (**Figure 5.3c)** was observed for each de-sialylated species (colored dotted lines). In short, branched glycans lose a Gal, while a LacNAc extension preserves a Gal and is evident by a +162 Da mass shift per LacNAc.

If the structural assignments in **Figure 5.3c** were correct, then upon GlcNAc removal the inner-most Galβ1-4GlcNAc disaccharide would remain attached to the trimannosyl N-glycan core, which we detected (**Figure 5.3d).** As a result of GUH digestion several EPO glycoform peaks, including *glycoforms 1-3* from **Figure 5.3b-c**, merge with other glycoforms based on the number of LacNAc extensions. As the peak intensities are a quantitative measure of LacNAcs present on N-glycans, the major EPO glycoforms had one core 1 O-glycan and three N-glycans that were not LacNAc extended (**Supplementary Table 5.1**). The second major ion was EPO glycoforms with two LacNAcs followed by four LacNAc extensions.

It is possible that the assignments of LacNAc extensions on N-glycans in **Figure 5.3d** were in fact due to the presence of more than one core 1 O-glycan, although this occurrence has not been reported previously. To exclude this possibility, we digested EPO with ABS, BGS, GUH plus O-glycosidase, an endoglycosidase that specifically trims core 1 O-glycans (**Figure 5.3e**). The series of peaks observed in **Figure 5.3d** shifted 365 Da as expected for the loss of the single core 1 O-glycan, however the peak intensity of the EPO glycoform with four LaNAcs was reduced. Interestingly the odd number of LacNAcs species were less abundant than even numbered species from both **Figure 5.3d-e**. We postulate that $LacNAc_2$ occur linearly on a single N-glycan branch (i.e. poly LacNAcs) as opposed to two separate branches. Overall, we measured EPO with between one and five LacNAcs, which has not been quantified to this extent previously either by native MS or bottom-up approaches. These data, and the ability of NES-MS to readily quantify LacNAc extensions, will be valuable in the glycosylation characterization of EPO and other biologics.

Identifying outer-arm fucosylation on α1-acid glycoprotein

One structural feature of complex-type N-glycans we have not addressed is fucosylation, which can occur at the reducing end GlcNAc (as seen for EPO) or at GlcNAc residues on the antennae (**Figure 5.1, legend**). To examine the proficiency of NES-MS for characterising outer-arm fucosylation, we used α1-acid glycoprotein (AGP), a plasma protein, belonging to the group of acute-phase proteins. AGP has five N-linked glycans, which are di-, tri-, and tetra-antennary

structures ± fucose (Fuc)[362]. N-glycan branching can differ during inflammation[363,364] and pregnancy[365] and the degree of N-glycan branching of AGP is biologically important as it influences immunomodulatory functions[366-368]. Theoretically, NES-MS can provide a high-resolution structural overview of AGP glycosylation, namely branching and quantification of outer-arm fucosylation, which has not yet been achieved on an intact level.

We have previously examined intact AGP and its binding characteristics to the drug warfarin and identified fucose dependent interactions, specifically decreased drug binding affinity to fucosylated AGP compared to non-fucosylated forms[126]. We did not however confirm AGP glycan structures on intact forms and therefore relied on compositional data and inferences from glycan biosynthesis to make structural assignments, a process that is time-consuming and ambiguous. Using NES-MS, we aim to build upon our previous findings and fully assign AGP glycoforms, the results of which will be broadly applicable for characterising glycoproteins with outer-arm or core fucosylation.

Following sialic acid removal, we found that all five N-glycosylation sites on AGP were occupied and N-glycans ranged from bi-antennary to tetra-antennary structures with at least two LacNAc extensions on AGP (**Figure 5.4a**). Each major peak (shown with cartoon glycan structures) had additional mass shifts corresponding the presence of one to three fucose residues (red dotted line from non-fucosylated peaks), consistent with our previous published results[126]. Interestingly, following GUH and BGS digestion, several glycoforms retained Gal

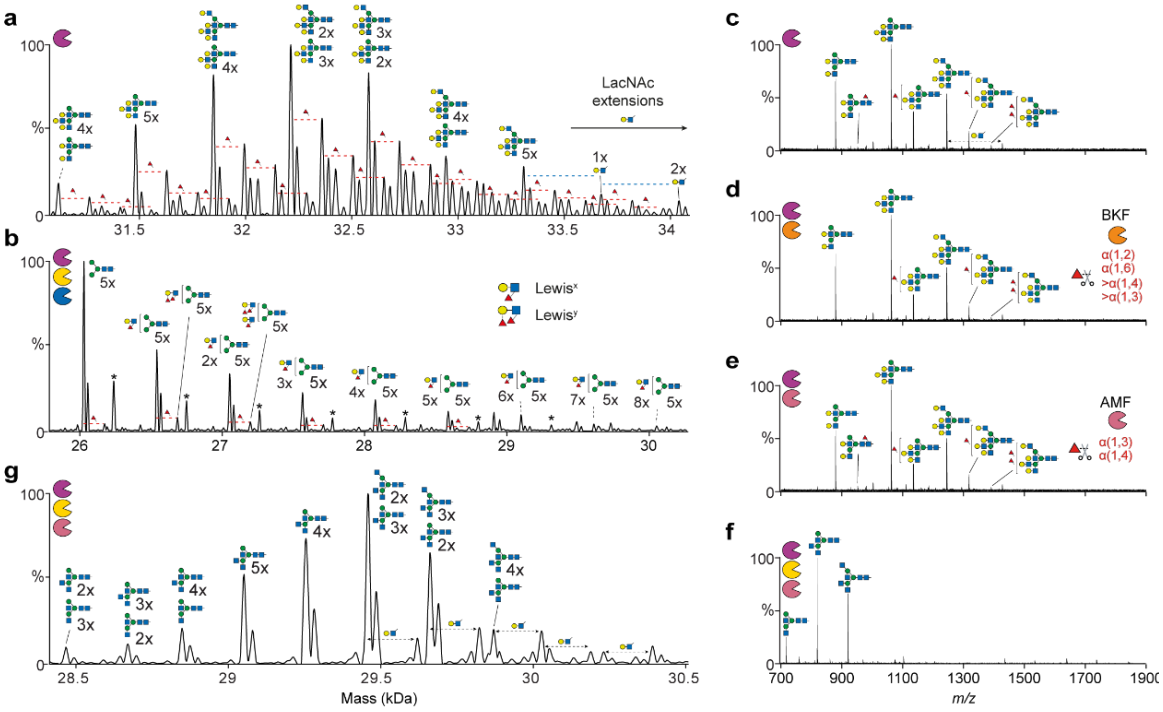


Figure 5.4 Native MS of de-sialylated AGP (**a**) and after sialic acid, galactose and N-acetylglucosamine removal (**b**). The peaks labelled with asterisks are assigned as a cysteinylated A variant. Glycomics analysis of AGP N-glycans treated with sialidase (**c**) with bovine kidney fucosidase (BKF) (**d**), with almond meal fucosidase (AMF) (**e**) and with AMF plus β1-4 galactosidase (BGS) (**f**). Native MS of AGP digested with ABS, BGS and AMF (**g**). Fucose mass shifts (146 Da) are shown with red dotted lines and LacNAc extensions (365 Da) are indicated with black dotted lines.

---

and GlcNAc residues but only when fucose was present (from one to eight fucose residues in total), suggesting terminal fucose prevents the action of β1-4 galactosidase (**Figure 5.4b**).

These mass shifts corresponded to the presence of a Galβ1-4(Fucα1-3)GlcNAc trisaccharide, consistent with a Lewis^{x} epitope. At the lower mass range, we also observed a composition that could only be assigned as AGP containing a (Fucα1-2)Galβ1-4(Fucα1-3)GlcNAc tetrasacchairde (i.e. Lewis^{y} epitope) at 26686 and 27197 Da in Figure 4b. Upon closer examination, we also detected a minor peak (26174 Da) corresponding to AGP with core-fucosylation as outer-arm fucose residues are generally linked to GlcNAc, which was absent based on this composition (i.e. five trimannosyl cores, one with fucose).

To confirm the assignments of Lewis^{y} and core-fucose glycans, we performed an LC-MS analysis of purified N-glycans. The de-sialylated N-glycan spectrum revealed several important features, first the presence of two fucose residues on a tetra-antennary N-glycan supporting the likelihood of a Lewis^{y} epitope (**Figure 5.4c**). A peak corresponding to a bi-antennary N-glycan with one fucose was also present in low amounts and could account for the core-fucosylated glycoform in Figure 4b. A double exoglycosidase digestion with sialidase and bovine kidney fucosidase (BKF), which is proficient at removing α1-6 linked core fucose, effectively removed this minor peak (**Figure 5.4d**). Performing a similar digestion, but replacing BKF with almond meal fucosidase (AMF), which does not target α1-6 linked core Fuc but does cleaves α1-3/4 linked Fuc more efficiently than BKF, resulted in the presence of the core fucose peak (**Figure 5.4e**) supporting the assignment in **Figure 5.4b**. Importantly, the remaining fucose residues were not cleaved, maintaining the view and observation in **Figure 5.4b** that outer-arm fucose residues prevent β1-4 galactosidase, even on free N-glycans. Steric hindrance prevents this

digestion and is caused by a strong nonbonded interaction between Fuc and Gal as each ring presents its most hydrophobic face to the other creating nonpolar van der Waals contacts[369]. This steric hindrance was overcome if β1-4 galactosidase and α1-3/4 fucosidase exoglycosidases were added together; as all N-glycans were digested to their GlcNAc residues on bi-, tri- or tetra-antennary N-glycans (**Figure 5.4f**). This digestion also worked for intact AGP where all five N-glycans were trimmed to the GlcNAc residue, revealing the level of branching for each peak (**Figure 5.4g**).

The glycomics analysis also underscores issues with limited de-glycosylation steps in native MS analyses when inferring glycan structural information, for instance, the use of desialylated spectra to assign the level of branching. Here glycomics showed that the relative abundance of di-, tri-, and tetra-antennary N-glycans was 26%, 50% and 24% on AGP, respectively. These abundances can be reasonably accounted for after Gal and Fuc removal in **Figure 5.4g**, but the combinatorial possibilities of branching are greater than what is shown (as illustrated in **Supplementary Figure 5.4**). For example, the five tri-antennary N-glycan peak (29051 Da) could also be one tetra-, three tri- and one bi-antennary or two tetra-, one tri- and two bi-antennary structures. Differentiating branching in a NES-MS approach when multiple N-glycans are present will require N-acetylglucosaminidases with high specificity to β1-2, β1-3, β1-4 or β1-6 linkages. Nonetheless, our analysis uncovered both outer-arm fucose, including minor amounts of Lewis^{y} epitopes as wells as core fucose, which have not been reported for AGP to the best of our knowledge.

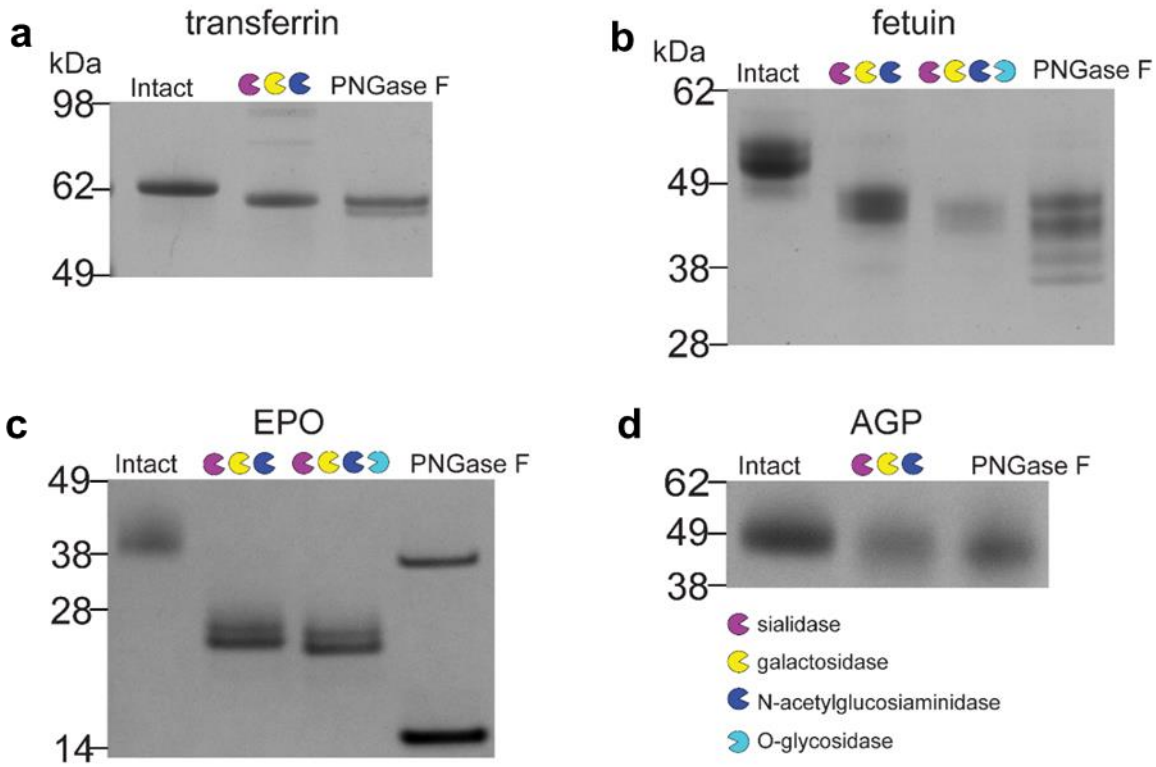
5.4 Conclusion

NES-MS is a remarkably simple, yet exceedingly powerful approach for studying glycosylated proteins – divulging glycan structure and occupancy information simultaneously. We believe this new method will not only streamline structural, or even functional, studies but also help uncover unique protein glycoforms that would go undetected by glycomics, glycoproteomics or conventional native MS analyses. NES-MS complements existing MS methods, namely glycan and glycopeptide analysis, however, the method cannot link an individual glycan structure to a specific glycosylation site when there are more than one glycosylation sites on a protein.

It is important to note the value of a native approach opposed to a denaturing measurement (i.e. top-down MS). The advantages of native MS are two-fold. First, by measuring charge-reduced ions the mass resolution is improved. Second, it allows for the measurement of non-covalent complexes, including glycan-specific interactions with small molecules[126]. The model systems presented here were selected with the aim of establishing the method, but consequently we identified glycoforms or structural features that were previously unknown or contested in the literature, such as the relationship of fetuin N- and O-glycosylation, amount of LacNAc extensions on EPO and the presence of core-fucose and Lewis^y epitopes on AGP. These glycoproteins have been widely studied by numerous methods and these findings underscore the value of NES-MS. Given the use of exoglyosidase enzymes for the analysis of non-

mammalian glycans, including plant and insect structures[44,45], we view NES-MS as having widespread applications in the field. As a final point, characterising some glycan structural features, such as branching, is logically reliant on available enzymes and their specificity. By and large this is not a limitation, rather an opportunity for further exploration and development – testing the range of exoglycosidases, their numerous combinations and the countless glycoprotein structures to be discovered by NES-MS.

5.5 Supplementary Information



Supplementary Figure 5.1 SDS-PAGE of glycoproteins used in this study with and without glycan enzymatic digestions. (**a**) human transferrin, (**b**) bovine fetuin, (**c**) epoetin alfa (EPO) and (**d**) α1-acid glycoprotein (AGP).

a

human Transferrin
average mass: 75157Da (19 disulfide bonds)
VPDKTVRWCAVSEHEATKCQSFRDHMKSVIPSDGPSVACVKKASYLDCIRAIAANEAD
AVTLDAGLVYDAYLAPNNLKPVVAEFYGSKEDPQTFYYAVAVVKKDSGFQMNQLRGKK
SCHTGLGRSAGWNIPIGLLYCDLPEPRKPLEKAVANFFSGSCAPCADGTDFPQLCQLC
PGCGCSTLNQYFGYSGAFKCLKDGAGDVAFVKHSTIFENLANKADRDQYELLCLDNTR
KPVDEYKDCHLAQVPSHTVVARSMGGKEDLIWELLNQAQEHFGKDKSKEFQLFSSPH
GKDLLFKDSAHGFLKVPPRMDAKMYLGYEYVTAIRNLREGTCPEAPTDECKPVKWCA
LSHHERLKCDEWSVNSVGKIECVSAETTEDCIAKIMNGEADAMSLDGGFVYIAGKCGL
VPVLAENYNKSDNCEDTPEAGYFAIAVVKKSASDLTWDNLKGKKSCHTAVGRTAGWN
IPMGLLYNKINHCRFDEFFSEGCAPGSKKDSSLCKLCMGSGLNLCEPNNKEGYYGYT
GAFRCLVEKGDVAFVKHQTVPQNTGGKNPDPWAKNLNEKDYELLCLDGTRKPVEEY
ANCHLARAPNHAVVTRKDKEACVHKILRQQQHLFGSNVTDCSGNFCLFRSETKDLLF
RDDTVCLAKLHDRNTYEKYLGEEYVKAVGNLRKCSTSSLLEACTFRRP

b

bovine fetuin A
average mass: 36353 Da
IPLDPVAGYKEPACDDPDTEQAALAAVDYINKHLPRGYKHTLNQIDSVKVWPRRPTGEVY
DIEIDTLETTCHVLDPTPLANCSVRQQTQHAVEGDCDIHVLKQDGQFSVLFTKCDSSPDS
AEDVRKLCPDCPLLAPLNDSRVVHAVEVALATFNAESNGSYLQLVEISRAQFVPLPVSVS
VEFAVAATDCIAKEVVDPTKCNLLAEKQYGFCKGSVIQKALGGEDVRVTCTLFQTQPVIP
QPQPDGAEAEAPSAVPDAAGPTPSAAGPPVASVVVGPSVVAVPLPLHRAHYDLRHTFS
GVASVESSSGEAFHVGKTPIVGQPSIPGGPVRLCPGRIRYFKI

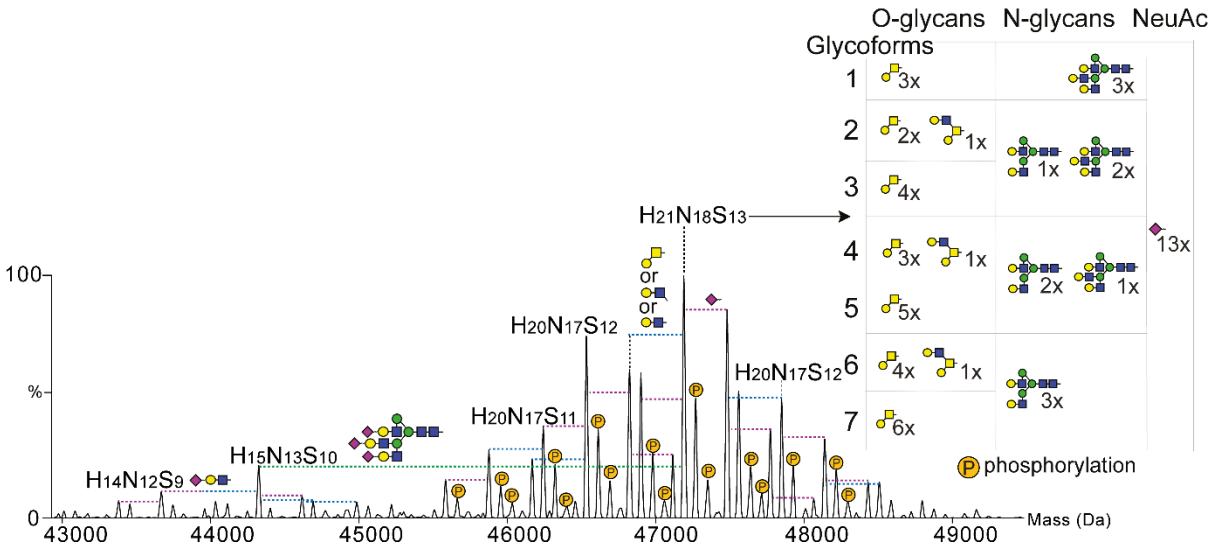
c

epoetin alfa (EPO)
average mass: 18236 Da (2 disulfide bonds)
APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQ
AVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKE
AISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD

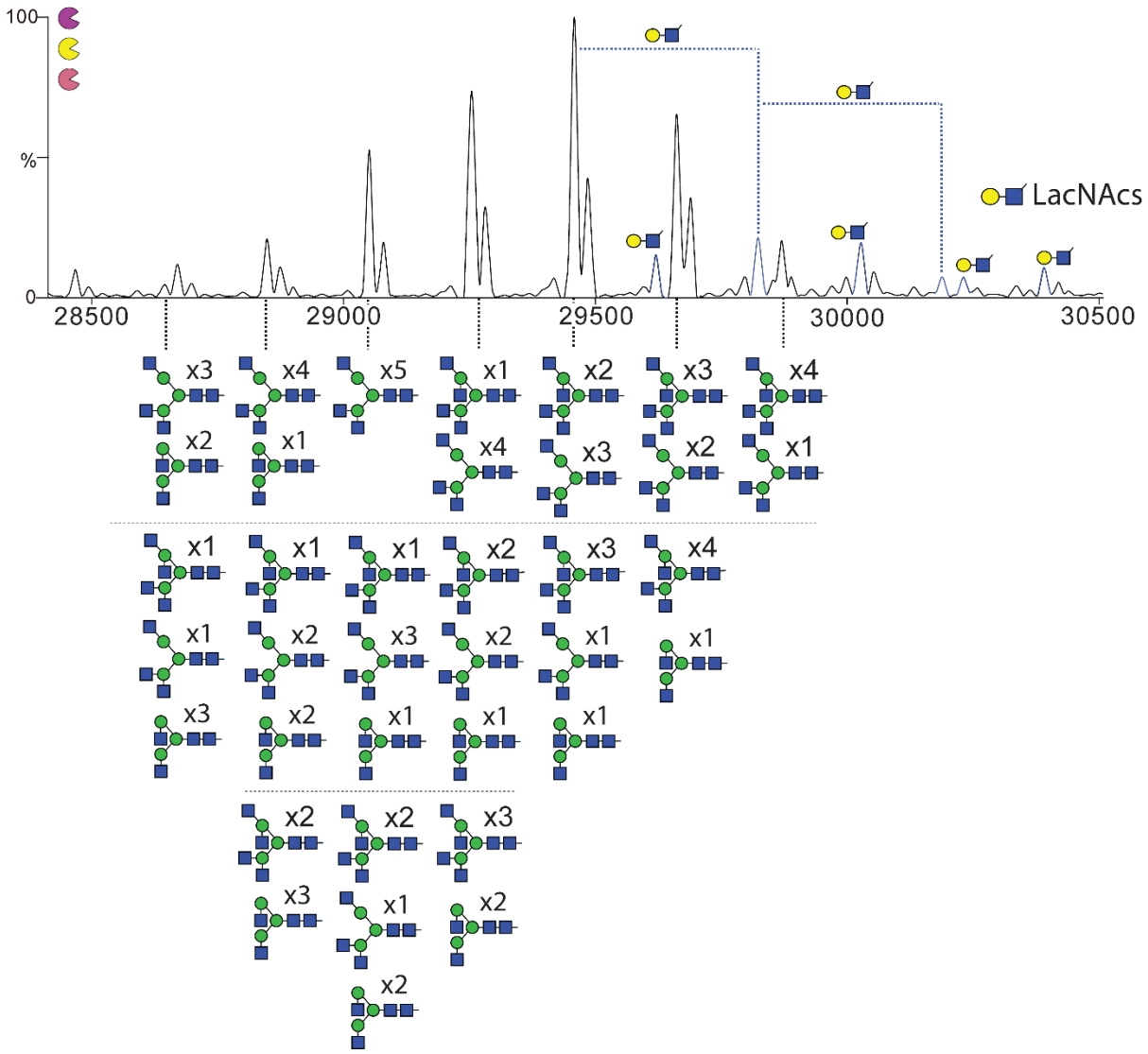
d

human α1-acid glycoprotein
average mass: 21560 Da
QIPLCANLVPVPITNATLDQITGKWFYIASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIFLR
EYQTRQDQCIYNTTYLNVQRENGTISRYVGGQEHFAHLLILRDTKTYMLAFDVNDEKNWG
LSVYADKPETTKEQLGEFYEALDCLRIPKSDVVYTDWKKDKCEPLEKQHEKERKQEEGES

Supplementary Figure 5.2 Glycoprotein amino acid sequences and average masses. (**a**) human transferrin (uniprot: P02787); (**b**) bovine fetuin (uniprot: P12763); (**c**) epoetin alfa and (**d**) α1-acid glycoprotein (uniprot: P02763). N-glycosylation (purple) and O-glycosylation sites (blue) are coloured.



Supplementary Figure 5.3 Native MS spectrum of intact fetuin. The major glycoforms of the most abundant peak are labelled. Interval difference of N-acetylneuraminic acid, 291 Da (violet dashed line), hexose plus N-acetylhexosamine, 365 Da (blue dashed line), phosphate group (orange cycle), the N-glycan with a H6N5S3 composition (green dashed line) are indicated.



Supplementary Figure 5.4 Native MS spectrum of AGP digested by sialidase, galactosidase and fucosidase. All possible antennary structures and compositions based on the accurate mass of the digested glycoprotein are shown. The LacNAc extension of the main peak is indicated by the blue dotted line.

ABS				ABS, BGS				Structure details		
Mass (Da)	Hex	HexNAc	Fuc	Mass (Da)	Hex	HexNAc	Fuc	Branches	LacNAcs	O-glycan
25006	19	16	3	23546	10	16	3	9	0	1
				23705	11	16	3	8	1	1
25369	20	17	3	23748	10	17	3	10	0	1
				23909	11	17	3	9	1	1
25735	21	18	3	23952	10	18	3	11	0	1
				24113	11	18	3	10	1	1
26102	22	19	3	24154	10	19	3	12	0	1
				24317	11	19	3	11	1	1
				24478	12	19	3	10	2	1
26470	23	20	3	24520	11	20	3	12	1	1
				24682	12	20	3	11	2	1
				24843	13	20	3	10	3	1
26832	24	21	3	24885	12	21	3	12	2	1
				25047	13	21	3	11	3	1
				25208	14	21	3	10	4	1
27119	25	22	3	25250	13	22	3	12	3	1
				25412	14	22	3	11	4	1
				25573	15	22	3	10	5	1
27563	26	23	3	25616	14	23	3	12	4	1
				25778	15	23	3	11	5	1
				25938	16	23	3	10	6	1
27930	27	24	3	25982	15	24	3	12	5	1
				26143	16	24	3	11	6	1
				26305	17	24	3	10	7	1
28298	28	25	3	26346	16	25	3	12	6	1

Supplementary Table 5.1 Native MS peak assignments of de-sialylated EPO and de-sialylated and de-galactosylated EPO.

Chapter 6 Probing Endogenous Ligands of Potassium Ion Channel Beta Subunit Potential for Sleep Regulation

6.1 Introduction

Sleep regulation by voltage-gated potassium channel

The voltage-gated potassium channel family (Kv) is a group of transmembrane potassium ion channels that broadly regulate membrane potential and excitability in a wide range of physiological contexts across the nervous, cardiovascular, endocrine, and immune systems[370,371].

In the studies with Drosophila, K^+ channel mutant flies have reduced sleep[372,373], and similar findings have been demonstrated in mice with corresponding mutants in Kvα subunits[374,375], highlighting the importance of Kv ion channals in sleep regulation. Recently, Kvβ subunit couples the oxidative byproducts of mitochondrial electron transport to regulate the activity of the sleep-promoting neurons, dorsal fan-shaped body (dFB) in the brain[376,377], by modulating Kv channel formed by α subunits[378,379] (**Figure 6.1**). Hence, the aim of this chapter is to investigate possible endogenous ligands of K^+ channel β subunits which regulates its function and influences sleep by means of native MS and other MS-based technologies.

Voltage-dependent potassium ion channel α and β subunits

Voltage-gated potassium channel family α and β subunits bind to each other at the plasma membrane to form an ion channel complex[380].

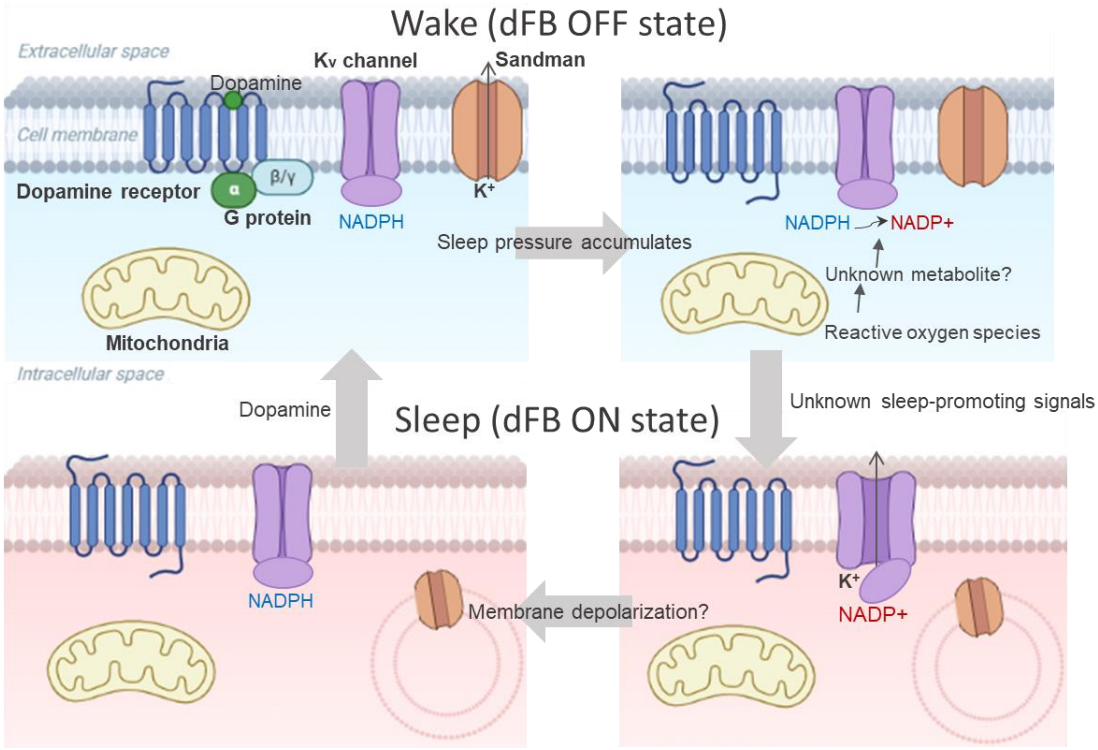


Figure 6.1 Model of sleep regulation in the dFB neuron. (Top) To initiate waking, dopamine binds to dopamine receptor, Dop1R2 (GPCR) on the dFB, which leads to downregulation of Kv channel through T1 domain, currents and leads to insertion of Sandman, a 2-pore domain potassium channel (K2P) channel, in the plasma membrane. Potassium leak currents through Sandman then predominate and lead to electrical silence within the dFB while awake. When sleep pressure accumulates throughout the day, mitochondrial reactive oxygen species production accumulates and leads to oxidation of NADPH to NADP+, thereby allowing the oxidation state of Kvβ's cofactors to act as biochemical memory and couple sleep pressure to electrical excitability. However, Kv channel currents cannot help drive to sleep since Sandman leak currents still prevail. (Bottom) When currently unknown sleep promoting signals arrive, they lead to the internalization of Sandman, thereby allowing dFBs to excite and delayed Shaker

currents to prolong spiking. Near the end of the night, NADPH cofactors replace NADP+, and the cycle repeats itself when awakening dopaminergic signals arrive. Schematic is adapted from Lyons & Rihel and Xu et al..

Kvα channels are composed of tetrameric, transmembrane assemblies that are highly selective for potassium ions, K^+ [381-384]. Voltage-dependent potassium channels at the plasma membrane regulates K^+ current from the cytosol to the exterior. These are six transmembrane-spanning K^+ channels with a voltage sensor[385]. Triggered by changes in membrane voltage, the ion channel undergoes a conformational change and opens the channel to allow ion current into the cell and adjust the membrane pressure.

Kvβ possesses a variety of functions i) promoting expression of the pore-forming alpha subunits at the cell membrane and improving pore stability[386-388]; ii) modulating potassium channel states of open or closure, so called gating[389-393]; iii) being as aldo-keto reductases and catalyzing the peroxidised products during reactive oxidative stress[374,394].

Kvβ exerts multiple chaperone-like effects on Kvα, i.e. assisting potassium channel folding and promote expression, including mediation of the biosynthetic maturation[387] and mediation of surface expression of voltage-gated K^+ channel complexed by enhancing the glycosylation of Kvα[388]. However, latter studies showed in contrast that assembly with β subunits did not increase the rate or extent of Shaker protein maturation[388]. Overall, it is difficult to exclude the possibility that one is quite likely to affect the other's abundance since β and α subunits co-express and co-assemble in endoplasmic reticulum (ER) as a tight complex in the cell membrane. Although some

influence on levels of expression and stability is evident, the chaperone hypothesis does not seem conclusive.

The gating function of Kvβ, i.e. modulating potassium channel open or closure is based on the observation that certain β subunit family members reduce ion conduction enabling inactivation. But different variants contribute to the gate inactivation distinguishably in different ways. For Kvβ1, this is commonly referred to as ball-and-chain or A-type inactivation and is best characterized in the K^+ channel from Drosophila, where the inactivation gate is formed by the α subunit's own N-terminal 20 or so amino acids located on the cytoplasmic side of the membrane[395]. This ability to mediate inactivation, however, is not universal for other β subunits. The β2 subunit, for example, is almost identical to β1 throughout most of its amino acid sequence except having a shorter N terminus and lacks the inactivation gate of β1[396,397]. Due to the missing inactivation gate, the β2 subunit has a reduced ability to inactivate the open K^+ channels.

Besides Kvβ is one of aldo-keto reductases and shares significant structural similarity to other aldo-keto reductases. Kvβ TIM barrel structure houses C-terminal active site that catalyzes the reduction of substrates with its cofactor, NADPH[383,398,399], where *NADPH donates one hydrogen and two associated electrons* and is oxidized to $NADP^+$.

As an aldo-ketoredutase, Kvβ can respond to intracellular oxidative stress by catalyzing a wide range of oxidative substrates[393,400,401], for example 4-oxonenal, 9,10-phenanthrenequinone, prostaglandin J2, 4-nitrobenzaldehyde, 4-nitroacetophenone and 4-oxo-trans-2-nonenal.

Unlike other soluble aldo-keto reductases found in the cytoplasm[380,383,402], Kvβ subunits bind to Kvα subunits at the plasma membrane[381]. The catalytic activity of Kvβ is able to regulate Kv

channel state. Kvβ subunits interact to Kvα T1 domains through a docking platform formed by the N-terminal domains of tetrameric Kvα[381,384].When reducing aldehyde-containing molecules is reduced and the Kvβ-bound NADPH is oxidized, Kv channel function is modulated [393,397,401].

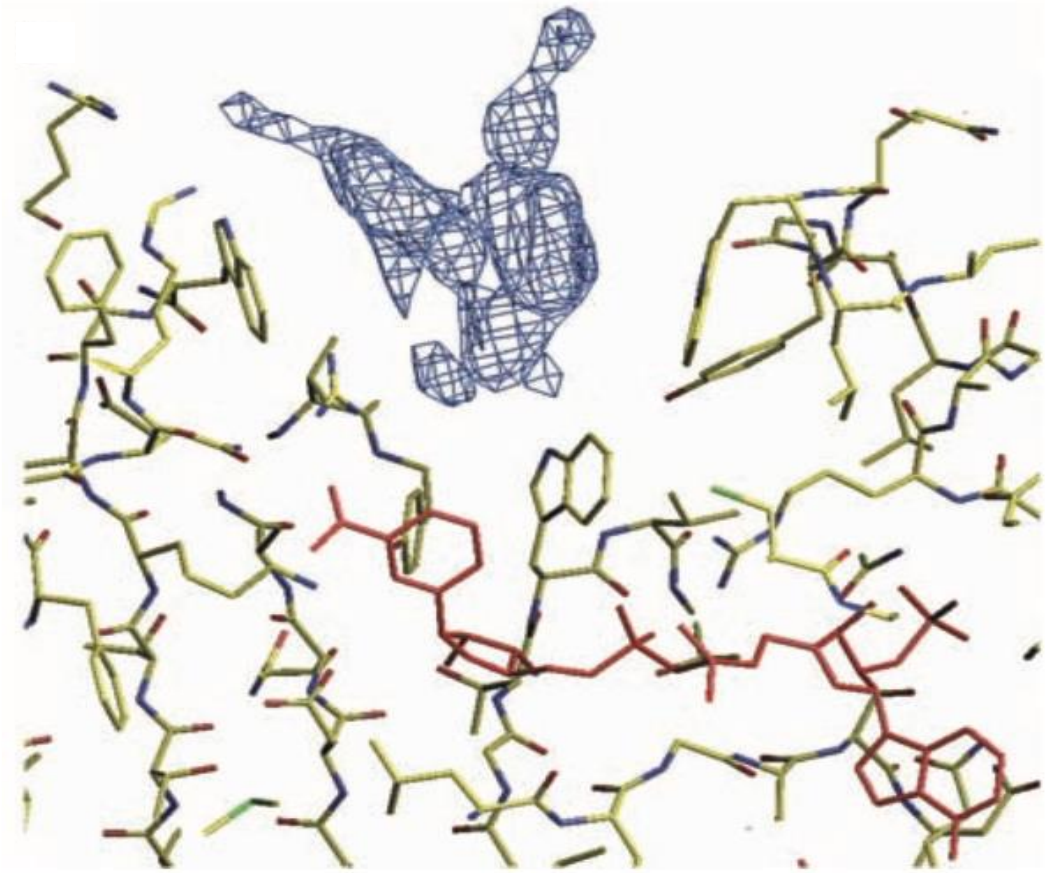


Figure 6.2 An unresolved electron density overlying the Kvβ2 catalysis site. In the structure of the Kv1.2-β2 complex, an unresolved electron density was found in the reduction pocket of Kvβ2 next to the reductase cofactor NADP+ or NADPH. This could represent endogenous substrates or inhibitor in all possible forms like lipids, peptides or small molecules. Figure is adapted from Long et al..

Rationale of linking Kvβ2 redox state to sleep regulation

A recent study shows that an NADPH-dependent ion channel β subunit couples the oxidative byproducts of mitochondrial electron transport to regulate the activity of dFB neurons[378,379]. Strong evidence has accumulated proposing the dFB neurons as a major 'on-off' switch for sleep regulation[403]. The intracellular redox state and mitochondrial metabolism are linked to the activity of dFB neurons via NADPH-dependentβ. Increased reactive oxygen species production from the mitochondria led to oxidation of Kvβ's NADPH cofactor and a change in Kv activity[378,404,405]. Indeed, it has been shown that mitochondrial reactive oxygen species production was elevated in the dFB neurons following sleep deprivation[378]. When sleep need accumulates, dFB neurons stay in OFF state. Low metabolic level in the dFB might result in a buildup of ATP that leads to the leakage of electrons from the mitochondrial transport chain. Reaction between electrons and oxygen forms negative oxygen ions and thereby produce reactive oxygen species[404,406,407]. Since reactive oxygen species are unstable and short-lived[408] and since the mitochondria is spatially separated from the plasma membrane where Kvβ acts, secondary signaling molecules to convey the redox relay are needed[408].

One possibility concerns lipid peroxidation, whereby oxidized polyunsaturated fatty acids generate lipid hydroperoxides and downstream radicals to perpetuate redox signalling[409], which is critical for various physiological processes including neurodegenerative disorders[410]. These lipid hydroperoxides and downstream radicals turn into more stable lipid peroxidation products with aldehydes or ketones. Then, these potential molecules with aldehyde or ketone can diffuse to the plasma membrane and bind to Kvβ[383,407]. These lipid peroxidation byproducts containing aldehyde or ketone can be reduced by NADPH-dependent Kvβ[393,401]. The changed redox state of Kvβ, i.e.

the ratio of NADP+ : NADPH, mediates the gating of Kv ion channel to regulate the dFB neuron's activity and hence sleep. Thus these lipid peroxidation intermediates or products could be potential key modulators through Kv ion channel for sleep regulation.

Even though a crystallography structure of Kv1.2-β2 complex revealed an unresolved electron density overlying the hydrophobic active site of Kvβ2 next to its cofactor NADPH/NADP+, its identity was not clear (**Figure 6.2**). The unresolved electron density at the catalytic site suggests that a heterogeneous mix of ligands bind to Kvβ subunits which could be peptides, phosopholipids or other small molecules[381].

However, it still remains unknown what endogenous substrates or inhibitors bind to Kvβ subunits. Therefore, the aim of experiments in this chapter is to look for endogenous ligands of Kvβ and seeks to identify the redox state of K^+ channels under the influence of the ligands by native MS and other MS-based technologies. This investigation of endogenous ligands in Kvβ by mass spectrometry may help reveal the type of molecule or provide further information of the potential ligands. These ligands may have a chance for a drug prototype for sleep regulation.

6.2 Methods

Cell culture

Frozen HEK293 cells were stored in liquid nitrogen in cryotubes containing a fetal bovine serum solution with 10% dimethylsulfoxide (DMSO). Human embryonic kidney 293 cells (HEK293) cells were grown in cell culture flasks in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum and penicillin-streptomycin. Cells were passaged via

trypsinization every 48 hr. Cells were pelleted by centrifuging at 1000 rpm for 5 min and resuspended in fresh media. These experiments were optimized together and carried out by Patrick Liu in Miesenböck group.

Plasmids and mutagenesis

Plasmids were amplified in and purified from One Shot TOP10 chemically competent E. coli according to the manufacturer's specifications. Mutagenesis experiments were performed using Q5 Site-Directed Mutagenesis Kit (New England Biolands). All constructs were validated via Sanger sequencing. Standard LB agar and broth with antibiotic were used to culture bacteria. These experiments These experiments were optimized together and carried out with Patrick Liu in Miesenböck group.

Transfection and stable cell line generation

HEK-293 cells were transfected using Lipofectamine 3000. To create a stable cell line, cells were transfected with a plasmid containing an Internal Ribosome Entry Sites (IRES) sequence with puromycin resistance. After 48-hr post-transfection, cells were passaged into fresh media containing puromycin (1 ug/ml) to select for puromycin resistant cells. Selected cells were continually passaged in media containing puromycin. For the expression of Kvβ2, the gene for the full-length mouse Kv1.2 (UniProt P62482) with a C-terminal Flag tag and the gene for the conserved mouse β2-core (residues 36-367, untagged) were ligated into multiple cloning sites of the vector (Invitrogen Life Technologies). This vector, containing both the Kv1.2 and the β2-core subunits, was used to transform human embryonic kidney 293 cells (HEK293 cells). For the co-expression, the gene for the full-length mouse Kv1.2 (UniProt P62482) with a C-terminal Flag tag

and the gene for the conserved mouse β2-core (residues 36-367, untagged) were ligated into multiple cloning sites of the vector (Invitrogen Life Technologies). This vector, containing both the Kv1.2 and the β2-core subunits, was used to transform human embryonic kidney 293 cells (HEK293 cells). These experiments were optimized together and carried out with Patrick Liu in Miesenböck group.

Expression in HEK293 cells.

Transformants were selected on YPDS plates (Invitrogen) containing 800-2000 µg/ml Zeocin (Invitrogen). Resistant colonies were grown in liquid culture and tested for expression by anti-HIS tag Western blotting. Transformed HEK 293 cells were grown in 4-litre batches in a Bioflo 3000 fermentor (New Brunswick Scientific). Cultures were grown at 37 °C, pH 5.0 and 40% dissolved O_2. Cells were grown using glycerol as a carbon source by fed-batch method to a density of 250 g/liter before induction at this point with methanol (0.07 ml/min for the first hour, then 0.17 ml/min for 24 hr total). Pelleted cells were immediately frozen in liquid N2 and stored at –80 °C.

Expression in bacteria and purification of His-tagged constructs

The sequence for Kvβ isoform RE was subcloned into a pNIC28-Bsa4 vector with an N-terminal His tag. BL-21(DE3) bacteria containing the plasmids with Hyperkinetic wild-type and various mutants were grown in terrific broth at 37 °C and induced with 1 mM isopropyl b-D-thiogalactopyranoside at 18 °C for 16 hr. Cells were harvested via centrifugation at 5,000 RPM for 20 min at 4 °C. Throughout the purification process, a 500 mM NaCl, 20 mM Tris, pH 7.9 buffer was used. Cells were lysed in buffer supplemented with 5 mM imidazole with a sonication probe using 80% amplitude with 5 s on and 5 s off for 12 min and then centrifuged at 14,000 RPM

for 1 hr. The supernatant was then incubated with Ni resin (Ni Sepharose 6 Fast Flow) on a shaking incubator for 1 hr at 4 °C before being loaded on a column pre-equilibrated with buffer and washed 3 times with a 5 mM imidazole buffer. The resin was then three times with a 20 mM imidazole buffer to wash non-specifically bound proteins before the protein was eluted with a 300 mM imidazole buffer and loaded on a size exclusion column pre-equilibrated with reaction buffer. Fractions containing the protein of interest were then concentrated. This sample preparation was prepared with Patrick Liu in Miesenböck group.

Purification of mouse Kvβ2

Cells were lysed with sonication probe (Vibra-Cell VCX-500 Watt, Sonics), at amplitude 60%, with a circle of 10 s on with the 20 s off for 8 min on ice. The lysis buffer contained 150 mM KCl, 50 mM TRIS-HCl, protease inhibitor, pH 7.5. The lysate was then centrifuged at 20,000 xg, 20 min, 17°C). The supernatant was added to 0.5 ml Pierce™ Anti-DYKDDDDK Magnetic Agarose (Thermo Fisher) previously equilibrated with 150 mM KCl, 50 mM TRIS-HCl, pH 7.5. After incubation for 45 min, beads were washed three times with 20 vol buffer (150 mM KCl, 50 mM TRIS-HCl, protease inhibitor, pH 7.5). 0.05mg/ml FLAG peptides in 150 mM KCl, 50 mM TRIS-HCl, pH 7.5 was added to eluted protein prior to concentration. The eluted protein was concentrated to minimal volume at 30 µl with 30k cut-off Amicon® microfilter (Merck Millipore).

Purification of mouse Kv1.2-β2 complex

Cells were lysed in 1% n-dodecyl-β-D-maltoside (DDM) buffer contained 150 mM KCl, 50 mM TRIS-HCl, protease inhibitor with sonication probe, 60% amplitude, 10 s on with the 20 s off for 20 min on ice. The lysate was then centrifuged at 20,000g, 20 min, 17°C). The supernatant was

added to 0.5 ml Pierce™ Anti-DYKDDDDK Magnetic Agarose (Thermo Fisher) previously equilibrated with 0.5% DDM, 150mM KCl, 50 mM TRIS-HCl, pH 7.5. After incubation for 45 min, beads were washed with 20 vol buffer (0.5% once and then 0.2% DDM twice, 150 mM KCl, 50 mM TRIS-HCl, protease inhibitor, pH 7.5). 0.1 mg/ml FLAG peptides in 150 mM KCl, 50 mM TRIS-HCl, pH 7.5 was added to eluted protein prior to concentration. The eluted protein is concentrated with Amicon microfilter 100k cut-off to 30 µl.

Western blot and coomassie staining

Protein samples were boiled at 95 °C for 5 min in Laemmli buffer and bmercaptoethanol and were run in pre-cast 4-12% Bis-Tris protein gels. InstantBlue staining was used for 1-hour Coomassie visualization. For Western blots, following protein transfer to a PVDF membrane, membranes were blocked in 5% milk in TBST (0.1% Tween-20 in trisbuffered saline) for 1 hr at room temperature and incubated with primary antibodies diluted in 5% milk in TBST overnight at 4 °C. Membranes were then washed three times for 5 min each with TBST, incubated with HRP conjugated secondary antibodies diluted in 5% milk in TBST for 1 hr, washed three times for 5 min each with TBST at room temperature, and imaged using enhanced chemiluminescence.

Native mass spectrometry of mouse Kvβ2

The purified Kvβ2 was buffer-exchanged with BioSpin (BioRad) into 200 mM ammonium acetate (AmAc) prior to native mass spectrometry measurement according the BioRad biospin guideline. The Kvβ2 sample was measured at the following settings: capillary voltage: 1.1 kV; in source trapping: off; source fragmentation: 0 V; HCD energy 0-100 V; HCD pressure 3.3e-10 mbar. Data

was acquired (m/z 2,000-15,000) and processed with XCalibur 2.2 (Thermo Fisher Scientific). Spectra were annotated manually and with in-house software.

Native mass spectrometry of mouse Kv1.2-β2 complex

The complex was buffer-exchanged with BioSpin into 200 mM AmAc containing 0.0174% DDM and then washed three times with 200 mM AmAc with 0.06% Octyl Glucose Neopentyl Glycol (OGNG) in 100 kDa-cutoff Amicon® centrifugal filter (Merck Millipore) prior to MS measurement. The Kv1.2-β2 complex was measured at the following settings: capillary voltage: 1.1 kV; in source trapping: -75 V; source fragmentation: 50 V; HCD energy 200 V; HCD pressure 4.42e-10 mbar. Data was acquired (m/z 2,000-15,000) and processed with XCalibur 2.2 (Thermo Fisher Scientific). Spectra were annotated manually and with in-house software.

Identification of Kvβ2 cofactor with mass spectrometry

Purified proteins were buffer-exchanged with 200 ammonium and were boiled at 95°C for 30 min. The proteins were ultracentrifuged at 20,000 g for 20 min. The supernatant was harvested for detection on Orbitrap Eclipse™ Tribrid™ mass spectrometer (Thermo Fisher Scientific) at standard pressure at positive ion mode. Data was acquired (m/z 500-1, 200) and processed with XCalibur 2.2 (Thermo Fisher Scientific). Spectra were annotated manually.

POVPC reduction and binding to Kvβ2

To test Kvβ2 activity with POVPC, purified Kvβ2 was buffer exchanged into 200 mM ammonium acetate and a 17 mM POVPC stock solution in water was made. Further dilutions were done as necessary to achieve the proper concentrations. For reduction studies, 40 uM of Kvβ2 was

incubated with 40 uM of POVPC at a 1:1 ratio. The mixture was incubated at 37 °C for 3 hr. The samples were then boiled at 95 °C for 30 min and centrifuged at 20,000 g for 20 min to release the substrates. The supernatant was harvested for MS analysis on on Orbitrap Eclipse™ Tribrid™ mass spectrometer (Thermo Fisher Scientific) at standard pressure at positive ion mode. Data was acquired (m/z 500-1, 200) and processed with XCalibur 2.2 (Thermo Fisher Scientific). Spectra were compared with POVPC standard and annotated manually. For POVPC binding experiments with naitve MS, 40 uM of Kvβ2 was incubated with either 40 uM or 2.8 mM POVPC at a 1:1 ratio at either 4 or 37 °C for 5 min. The incubation time for binding experiments was shorter than reduction experiments, which was attempt to capture the binding that precedes catalysis on native MS.

Immunofluorescence imaging

HEK-293 cells were grown on glass microscope coverslips. They were then fixed in 4% paraformaldehyde in (Phosphate Buffered Saline) PBS for 15 min, washed three times with PBS for 5 min per wash, permeabilized with 0.1% Triton X-100 in PBS for 15 min, washed with 0.06% Tween-20 in PBS for 5 min, washed twice with PBS for 5 min, and blocked for 1 hor at room temperature in 10% normal goat serum in PBS. Primary antibodies were diluted in 10% normal goat serum in PBS and incubated with the cells overnight at 4°C. Cells were then washed three times with PBS for 5 min, incubated with fluorescent secondary antibodies diluted in 10% normal goat serum in PBS for 1 hr away from light, stained with Hoescht, and washed three times with PBS for 5 min at room temperature before being mounted on a microscope slide with VECTASHIELD. Coverslips were then imaged on a Leica confocal laser scanning microscope and images were processed using ImageJ 2.2.

6.3 Results and Discussion

To investigate possible ligand bindings to Kvβ2, we first expressed the protein with a FLAG tag at C-terminus in HEK293 cells. A mammalian cell possesses the complete metabolite pathway of mitochondria and this attempts to mimic a similar cellular environment as the ex-vivo and in vivo studies showing reactive oxygen species generating from mitochondria chain tend to influence he redox state of mitochondria. Hence the relevant peroxidative by-products generated from mitochondrial electron transport may interact with Kvβ2. Kvβ2 was then purified with immunoprecipitation using magnetic beads and eluted with FLAG peptides (see **Methods**) for native mass spectrometry and other complementary MS studies. In this way, the protein is preserved in a native state so that the endogenous cofactors or ligands should still stay bound[72,411].

Then Kvβ2 was measured with native MS after buffer exchange into 200 mM ammonium acetate. A series of board peaks at 336 kDa is obtained on the spectrum without activation (**Figure 6.3a**). Compared to the calculated mass of a tetramer bringing 4 cofacters, 166 kDa, it is slightly larger than the double of the tetramer, which could be an octamer structure. The peaks ranging from 332-340 kDa appear broad and heterogeneous. This implies that probably a variety of substrates are bound to Kvβ2. In order to further look into the composition of this 366k species, the activation energy in the high collision dissociation cell is gradually elevated to track the change of the complex. At 75 V, another series of peaks generated from the original series are 10 kDa smaller than the original peaks (**Figure 6.3b**). With a supplemental increase to 100 V (**Figure 6.3c**), yielded predominant peaks at 334 kDa which could be resolved, suggesting the presence of an octamer with eight cofactors. Along with the main species, the 327 kDa species is at lower

abundance than the apo octamer of Kvβ2. It is unexpected to find Kvβ2 in an octameric state with native mass spectrometry, because Kvβ2 was recognized as tetramer with X-ray crystallography. But an interface between two tetramers was also observed and captured in crystal latiice and it was

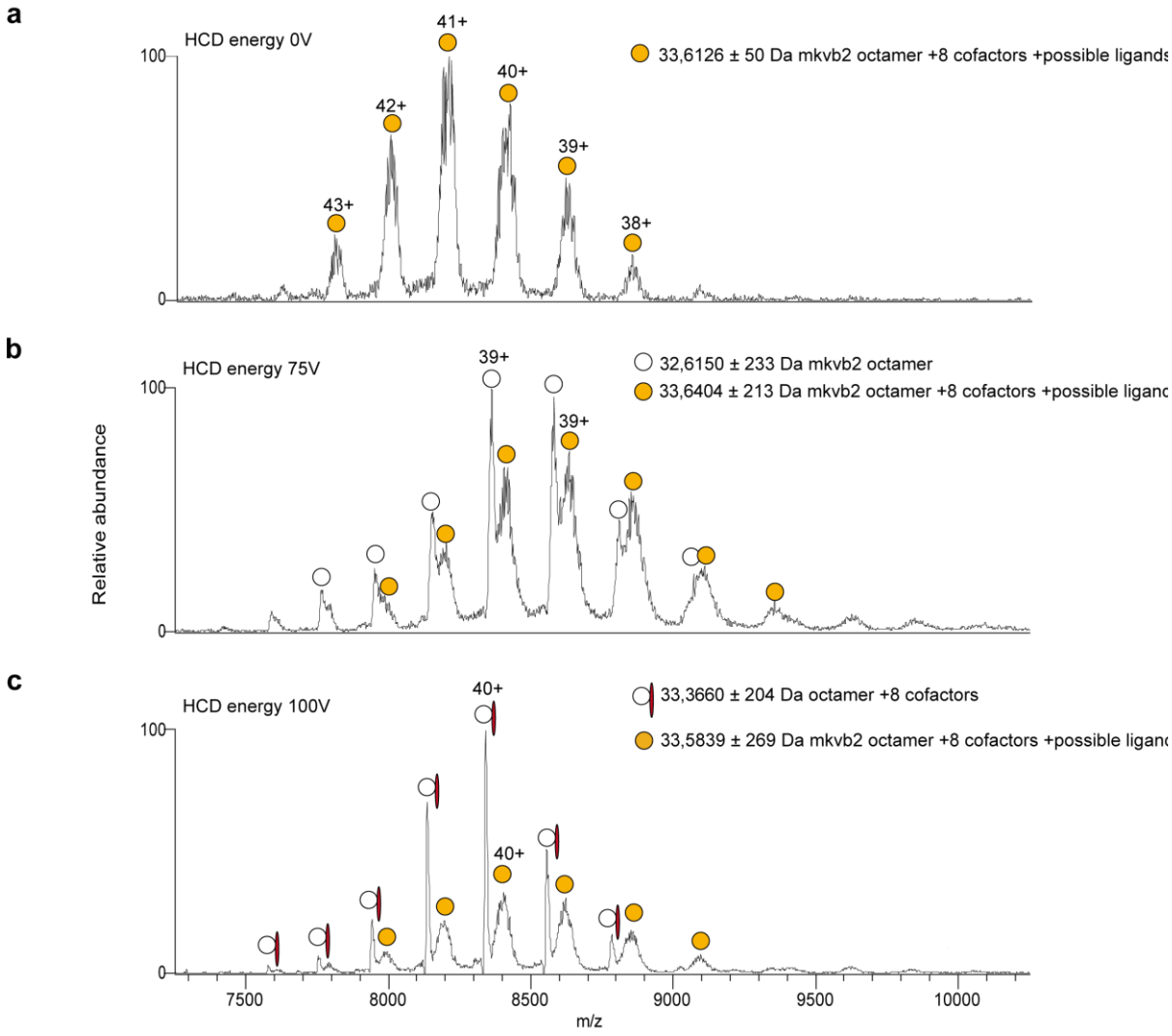


Figure 6.3 Structural characterization of mouse Kvβ2 with native MS. Mouse Kvβ2 octamer bound with endogenous ligands, apo Kvβ2 partially released, Kvβ2 octamer bound with cofactors activated at 0 **(a)**, 75 **(b)**, 150 V **(c)** in high collision dissociation cell at pressure: 3.3 e-10 mbar.

yet difficult to prove if it is a native interface rather an artefact in X-ray crystallography[383] (**Figure 6.4**). On the other hand, based on the structure of Kv1.2-β2 complex[381], α subunits form tetramer in a similar way like β subunits. The C-terminal of β subunits interacts with T1 domain of alpha subunits. With the assumption that β subunit octamer is still functional and can bind to alpha subunits, two β subunit tetramers are most likely and come together and bind at its N-terminal surface. This inference matches the finding on the crystallography grid (**Figure 6.4**), in which a homo-octamer is composed of two ring-like tetramers (PDB:1QRQ) mentioned above. Two tetramers are in contact via their N-terminal side. Two tetramers do not overlap with each other completely but are rotated by about 15°. This is possibly attributed to preserving accessibility for

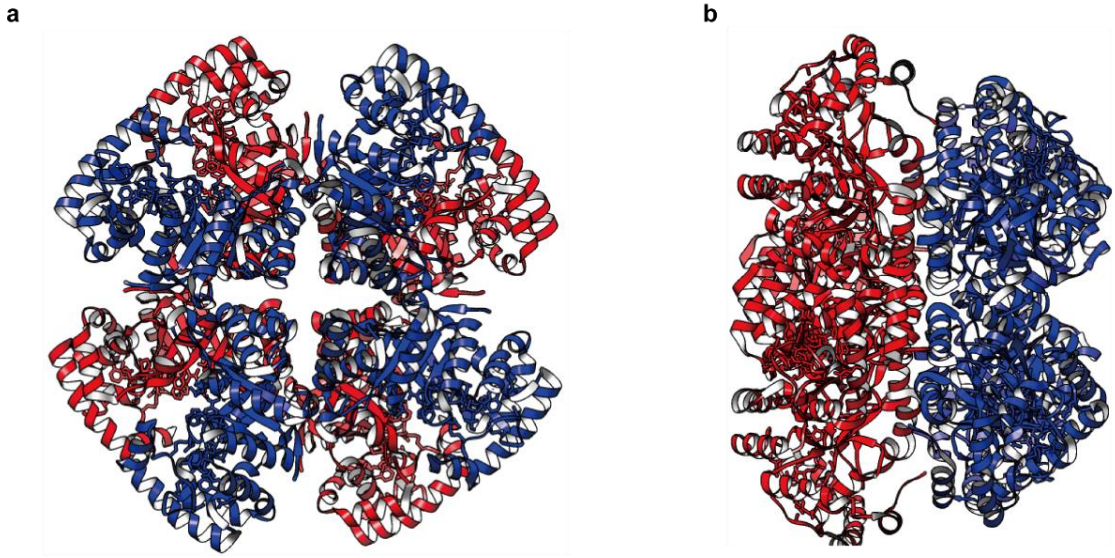


Figure 6.4 Structural characterization of Kvβ2 by crystallography. Mouse Kvβ2 octamer was found in the crystallography (PDB). The C-terminal surfaces of mouse Kvβ2 remain accessible and two N-terminal surfaces of tetramer interact with each other to form a homo octamer. Image from the RCSB PDB (rcsb.org) of PFDB Code: 1QRQ.

NADPH going into or NADP+ leaving redox pocket so that Kvβ2, a NADP(H)-dependent oxidoreductase, stays functional. Hence all monomers are potentially to retain catalytical activities as an aldo-keto reductase.

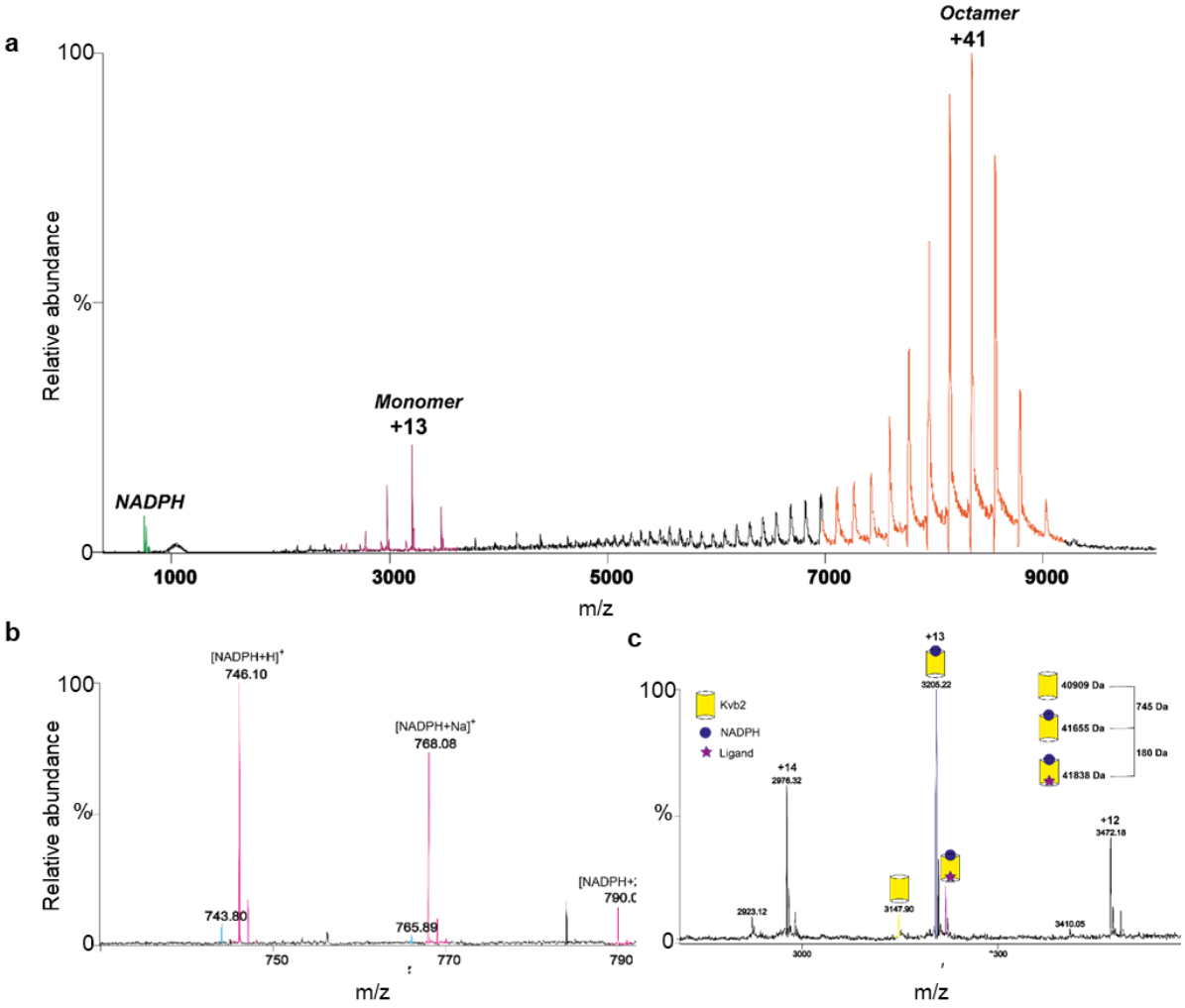


Figure 6.5 The confirmation of homo-octamer of Kvβ2 and its bound cofactor. The native mass spectrum of mouse Kvβ2. The octamer Kvβ2 is activated in mass spectrometry and dissociated into the monomer and cofactor, NADPH (a). The monomeric mouse Kvβ2 (b) and NADPH (c) were released from octamer. Figure data is acquired with Dr Dinesh Chinthapalli.

The octamer was successfully activated in the HCD cell of the MS with using adequate amount of urea so that monomeric Kvβ2 dissociates from the octamer, which further confirms Kvβ2 as a homo-octamer (**Figure 6.5a**). At the same time, NADPHs, the cofactors of Kvβ2 are released from Kvβ2 (**Figure 6.5a, c**). NADPH is the reductive agent in the reaction[401,412]. Kvβ2 recruits and binds NADPH. In response to the metabolic and redox state of the cell, Kvβ2 reduces aldehydes and ketones into alcohol by transferring two electrons from NADPH. This leads to NADP+, and oxidized products. Interestingly, there is a +180 Da adduct peak next to the monomeric apo Kvβ2 (**Figure 6.5c**). Based on the amino acid sequence and post-translation modifications of Kvβ2[413], this possibly represents a non-covalent ligand bound to monomeric Kvβ2. For NADPH bound Kvβ2, the ratio of ligand-bound form to free form is about 0.2. Additionally, 10% apo Kvβ2 and the ligand-bound Kvβ2 are in lower abundance but also exist next to the main peaks of NADPH-bound Kvβ2. This suggests that the ligand of 180 Da is bound independently regardless of whether

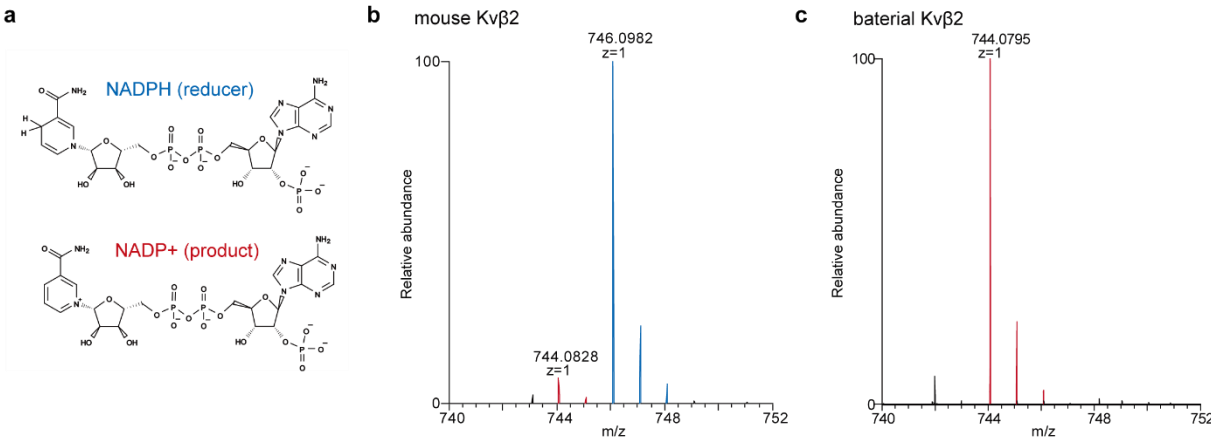


Figure 6.6 Cofactors released from mouse and bacterial Kvβ2. NADPH and NADP+ are a redox couple for mouse Kvβ2 **(a)**. The reactant NADPH is predominantly co-purified with mouse Kvβ2 **(b)** but oxidized NADP+ is released from bacterial Kvβ2 **(c)**.

Kvβ2 is bound to NADPH. Further investigation was carried out for identification of the redox state of cofactors. By boiling and denaturing the protein, the cofactors were released from Kvβ2. By comparing mouse and bacterial Kvβ2 (**Figure 6.6**), we could determine that 90% of cofactors of mouse Kvβ2 is NADPH but bacterial cofactor is an oxidised product, NADP+. This difference in the redox state of cofactor between mouse and bacterial Kvβ2 is unexpected. The discrepancy in the ratio of NADPH to NADP+ from mouse and bacterial species indicates the different reductase activity of Kvβ2. Since NADPH is the reducer and NADP+ is the oxidised product of the reductase, mouse Kvβ2 is less active and reactive compared to bacterial Kvβ2. The identity of a cofactor in rat Kvβ2 expressed in the Sf9 cell system is NADP+[383], similar to the findings in bacterial Kvβ2. This implies that the eukaroyotic biological system might provide specific metabolites to influence the reductase activity of Kvβ2. It will be interesting to find out the exact reason underlying the inability of mouse Kvβ2 to oxidise NADPH.

Following previous studies on the inhibition of Kvβ2[414], the inhibitors used ranged from 110 to 360 Da. They inhibit Kvβ2 from Aldo-keto reductase activity depending on their concentration but never give rise to complete inhibition. However, the inhibition hasn't been thoroughly studied. It is difficult to conclude whether the inhibitors compete with substrates by binding to the active site of Kvβ2 or induce an allosteric effect by binding to a site that is not the active site. Among these inhibitors, cortisone has been shown to have a significant effect on channel activity site compared to its limited inhibition on aldehyde turnover by Kvβ2.

This drives us to further probe if this ligand also exists in the Kv1.2-β2 complex and if so its role in regulating channel activity. The Kv1.2-β2 was encoded and expressed in HEK293 cells. Untagged Kvβ2 is co-purified with Kv1.2 affinity by FLAG immunoprecipitation

(**Supplementary Figure 6.1**). It is also shown that Kvβ2 is bound to the Kv1.2 channel on the membrane in the confocal image (**Supplementary Figure 6.1**). This was in agreement with recent reports indicating that Kvβ2 enhances the localization and trafficking of Kv1.2.

The whole Kv1.2-β2 complex or Kv1.2 tetramer was not able to be detected due to the limitation of ionisation of large membrane assemblies in mass spectrometry. However, Kvβ2 octamer disassociated from the Kv1.2-β2 complex are bound with 3-8 cofactors (**Figure 6.7**). This probably resulted from the activation in HCD cell and cofactors are falling off from Kvβ2 octamer which

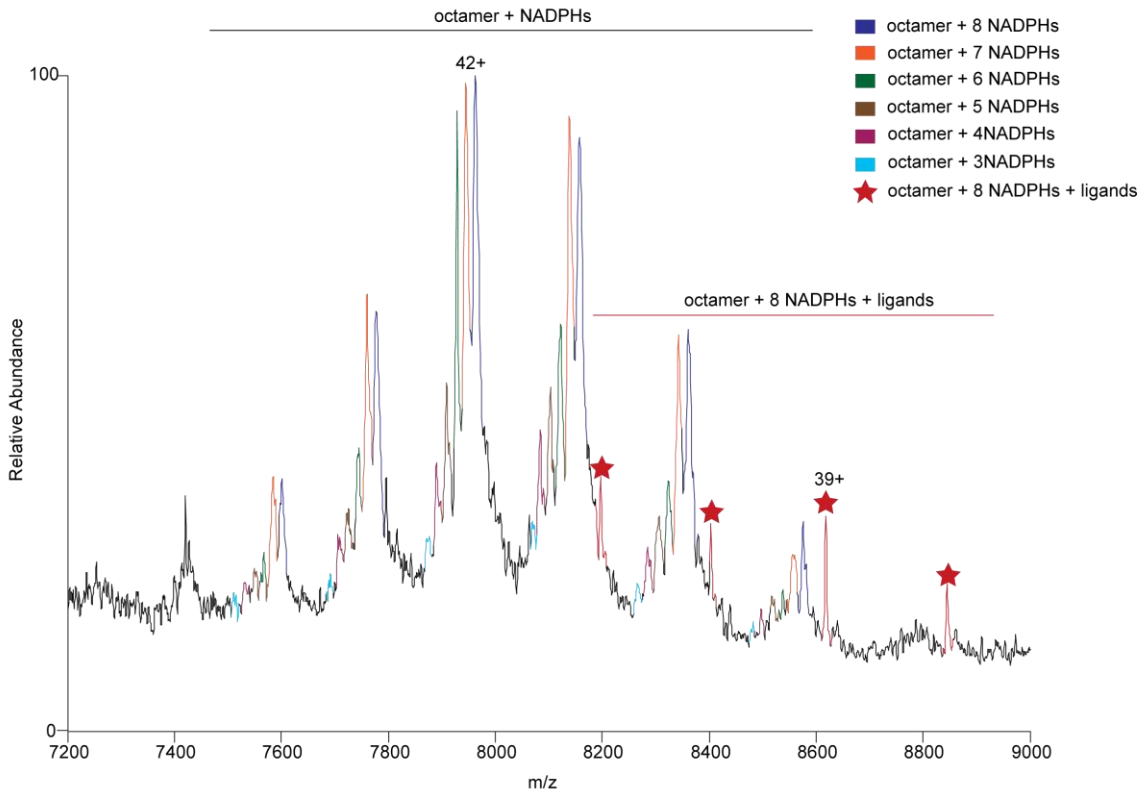


Figure 6.7 Structural characterization of Kvβ2 co-purified with Kv1.2 by native MS. Mouse Kvβ2 octamer bound with NADPH ranging from 3 to 8, left to right accordingly and Kvβ2 bound with 8 cofactors and ligands in a homogenous manner.

originally binding at 1:1. Interestingly, surprising peaks shows a ligand bound to the Kvβ2 fully with NADPH docking. The ligands are estimated to be 1.5 kDa, which is close to the sum of eight 180-Da ligands. This result further supports the inference of the small-molecule inhibitor to Kvβ2.

Previous studies about a reduction on a wide range of aldehydes and ketones[393,401] concluded substrate specificity. All these substrates are substrates of AKR. To examine if these substrates can bind to Kvβ2, a binding study with PGJ2 and POVPC incubated with β2 is performed. PGJ2 and POVPC were selected, because they are large enough to be resolved and visible in native MS while having a high substrate specificity among various aldehydes and ketones. Unfortunately, native mass spectrometry results suggest that Kvβ2 does not show detectable binding activity to PGJ2 or POVPC after incubating with a 50-fold excess of these (**Supplementary Figure 6.2**). One potential explanation is that NADPH-bound Kvβ2 rapidly reduces the substrate and releases the reduced product transiently. But this possibility is not against the observation on catalytic activity of Kvβ2 examined on POVPC. Under the same temperature, POVPC is not successfully reduced by Kvβ2 (**Supplementary Figure 6.2**). The alternative explanation is that the endogenous inhibitory ligand binds to a site and affects the binding of substrates so that the catalytic activity is blocked.

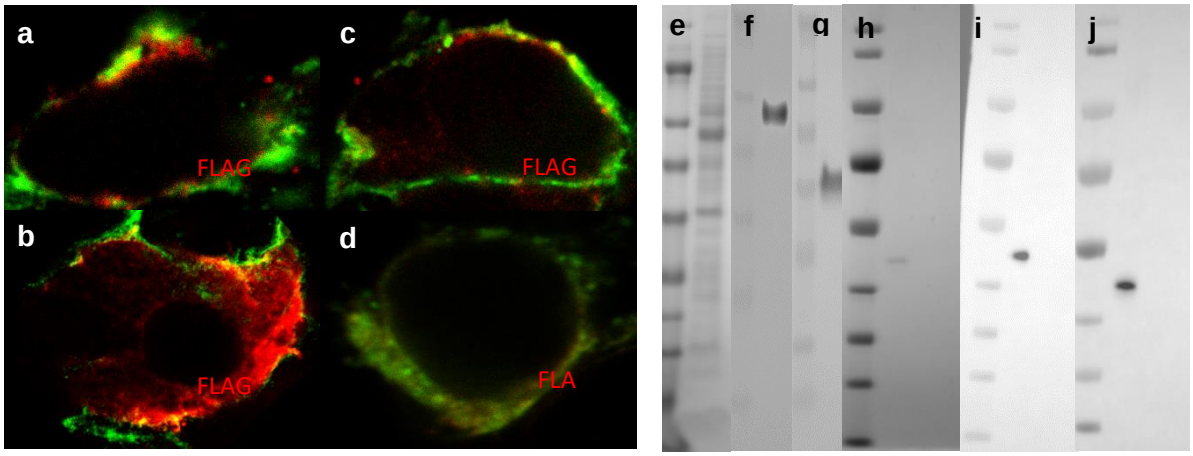
6.4 Conclusion

The investigation of Kvβ2 by native MS shows that the stoichiometry of Kvβ2 is a homo-octamer. This is supported by the previous crystallography data that two Kvβ2 tetamers have an interface at their N-terminus. We also found that mouse Kvβ2 expressed in HEK 293 cell line is mostly bound with NADPH, present at a high redox state. Meanwhile, purified Kvβ2 is found to be bound

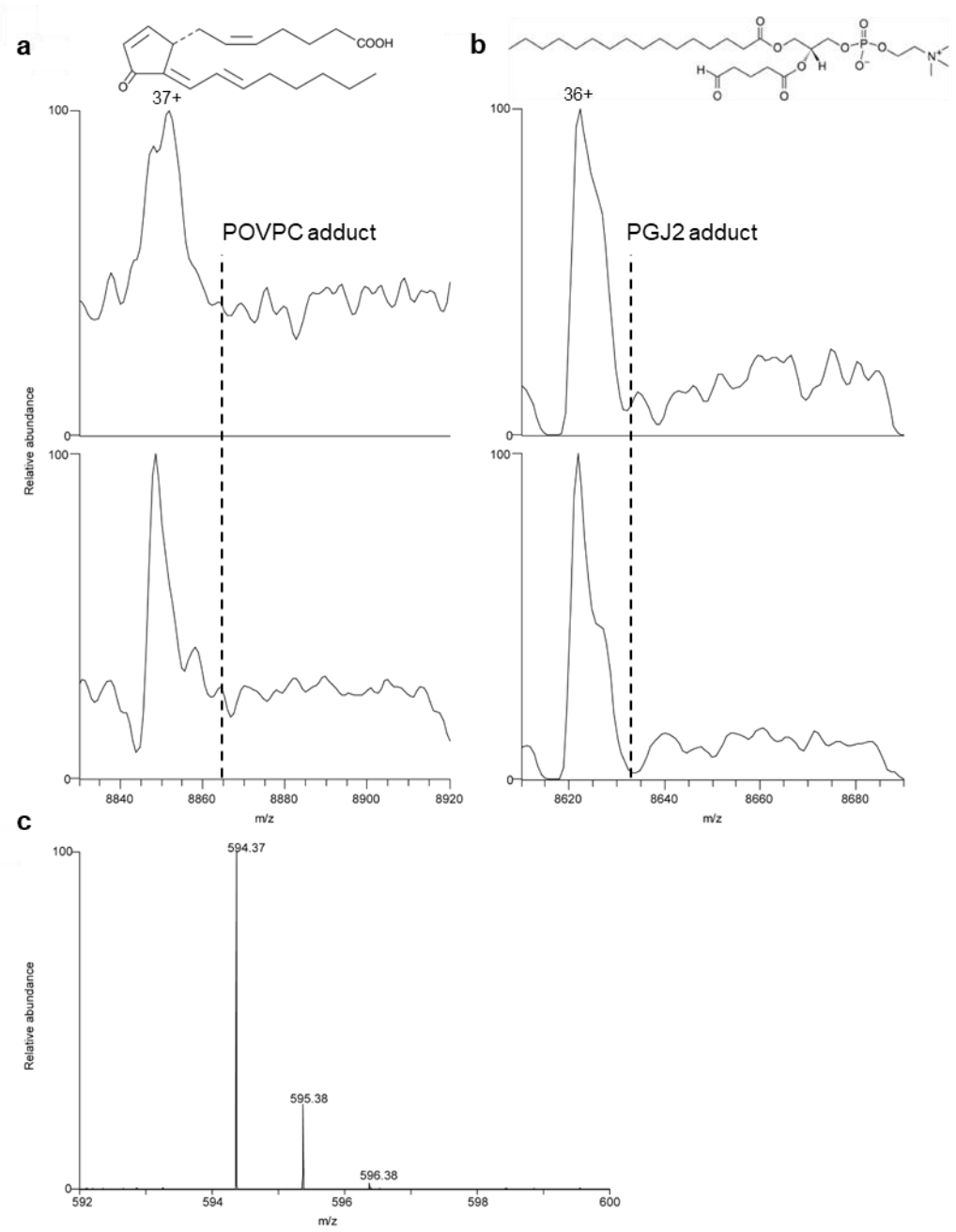
with a large group heterogenous ligands. Supported by the inability of substrate binding and reduction on a known substrate of Kvβ2, POVPC, the discovery of a 180-Da ligand bound to Kvβ2 suggests the possibility of an endogenous inhibitor for the catalytic activity of Kvβ2.

However, the understanding of this small molecule is lacking so it is challenging to apply gas/liquid chromatography–mass spectrometry in a suitable condition without knowing the chemical property of the molecule (data not shown). Second, nativeomics was difficult to perform because of octameric Kvβ2 exceeding the m/z range of Orbitrap Eclipse™ Tribrid™ mass spectrometer. What is more, the analysis of small-molecule metabolites also remains challenging as the analysis of MS^n spectra from gas/liquid chromatography–mass spectrometry or nativeomics are still limited by manual analysis and not all the fragment peaks are analysed. In the future, new advances in instrument, spectrum analysis software or experimental design are needed to identify the chemical structure of the inhibitor.

6.5 Supplementary Information



Supplementary Figure 6.1. Representative confocal images of individual HEK 293T cells (**a**) transfected with Kv1.2-FLAG and stained for FLAG (red) and N-cadherin (green); (**b**) transfected with Kvβ2-FLAG and stained for FLAG (red) and N-cadherin (green); (**c**) transfected with Kv1.2-FLAG and Kvβ2 and stained for FLAG (red) and N-cadherin (green); (**d**) transfected with Kv1.2-FLAG and Kvβ2 and stained for FLAG (red) and Kvβ2 (green). To confirm that Kvβ2 binds to the ion channel at the plasma membrane, FLAG immunoprecipitation using the detergent beta-DDM was done on cells transfected with Kv1.2-FLAG and Kvβ2. The eluted protein sample was (**e**) run on a Coomassie gel and western blot stained for (**f**) FLAG and (**g**) Kvβ2. Purified Kvβ2-FLAG was then (**h**) run on a Coomassie gel, and stained for (**i**) FLAG and (**j**) Kvβ2 on a Western blot. Confocal experiments were carried out by Patrick Liu.



Supplementary Figure 6.2. Examination of substrate binding by native MS. (**a**) Mouse Kvβ2 is incubated with POVPC, bottom and compared to control, top. (**b**) Similarly, mouse Kvβ2 is incubated with PGJ2, bottom, compared to the control, top. (**c**) After 2 hour-incubation of POVPC

with Kvβ2 and measured at positive ion mode at low m/z, POVPC is not reduced by NADPH-bound Kvβ2.

Chapter 7 Concluding Remarks and Outlook

In the past 30 years, an increasing number of questions in biology have been explored with various emerging interdisciplinary technologies. In this dissertation, insights have been provided with native MS and other complementary omics-based technologies. By bringing together these approaches it is possible to investigate the structural biology of glycoproteins, the regeneration and activation of rhodopsin and its signal amplification in its native membrane microenvironment (first for GPCRs) as well as investigating endogenous ligands of the potassium ion channel beta subunit affecting its redox state for sleep regulation.

This dissertation first introduces the recent advances and applications using native MS (Chapter 1) and the studies in rhodopsin, particularly focusing rhodopsin's relationship to the disc membranes and its function (Chapter 2). It provides a basis into the investigation of rhodopsin in native membrane, which is a focus of this thesis.

In more detail, we use native MS to enable the study of rhodopsin in disc membranes in Chapter 3. A new native MS platform is established to explore rhodopsin in disc membranes. We successfully detected rhodopsin ejecting from its disc membrane for the first time. This demonstrates the potential for using native MS to characterize GPCRs without solubilising them first into detergent micelles meanwhile preserving as many interacting molecules as possible. Such a method opens up opportunities to obtain new information that is ignored by other existing methods involving detergents. Combined with an experimental setup enabling real-time monitoring, we provide a time-resolved observation of photoactivation and regeneration of

rhodopsin. Opsin formation in membranes, monitored with native mass spectrometry, supplements measurement of metarhodopsin II decay by fluorescence studies. As well, this experimental method enables observation of rapid rhodopsin regeneration via native lipids from the membrane. Again, this methodology enables detection of rhodopsin molecules from native membrane despite the complexity of lipid composition and validates that the activation energy in mass spectrometer is sufficient to eject GPCRs.

Considering the high concentration of rhodopsin in membranes, up to about 95%, the next challenge is whether this approach is applicable for other membrane receptors with lower relative abundance. To achieve this, there is some transferrable experience based on this study that suggests this might be possible in the following scenarios: (i) a purified membrane with few soluble proteins – the membrane needs to be purified with isotonic and hypotonic buffer washes; (ii) after desalting of membranes – our hypothesis (which needs further investigation) is that this method minimizes membrane fraction loss thus elevating the abundance of vesicles in the sample while facilitating their ionization and desalting of the protein in the mass spectrometer. Overall, we assume that this led to a signal intensity sufficient to perform our experiments; (iii) overexpressed membrane proteins, or dominantly existing receptors e.g. olfactory receptors are considered compatible with this method. In general, this method takes a lead in investigating the puzzle of rhodopsin — understanding of the relationship and interaction of rhodopsin and its chromophore as well as the mechanism of rhodopsin regeneration and photoactivation. This would also potentially help develop drugs for vision diseases based on rhodopsin mutations such as retinitis pigmentosa.

On the basis of the real-time monitoring of rhodopsin photoconversion, we expanded the investigation into rhodopsin and its signalling partners in Chapter 4. Here we use native mass

spectrometry as a tool for biological mechanisms looking at rhodopsin and interacting partners, transducin and PDE6. The activation of G protein transducin is characterised by the molecular-shift due to GDP dissociating from activated transducin. The transducin activation process is observed is captured in seconds. Meanwhile, more unsaturated lipids adhere to rhodopsin during signalling.

The activated transducin α subunit further activates PDE6 asymmetrically where a complex of $G_t\alpha$ and PDE6 at 1:1 molar ratio is observed. This supports the hypothesis that PDE6 α and β subunits interacting with $G_t\alpha$ have a different binding affinity. After activation of PDE6, cGMP is released as it is hydrolyzed by activated PDE6.

By monitoring rhodopsin downstream signalling by native mass spectrometry, we are able to observe the phototransduction process and relevant lipid interaction, catalysis of cGMP by PDE6 and activated PDE6 stoichiometry. We anticipate that similar approaches can be applied to other low-expressed cell surface receptors to observe the dynamics of molecular events in a complex environment complementing other assays or methods. In summary, we present a new approach to drug discovery in which different stages of a signalling cascade can be targeted across native membranes. This potentially enables the identification of endogenous and synthetic receptor ligands that perturb signalling.

In addition to molecular events of signalling, glycoprotein structures are investigated in Chapter 5 by employing native MS. A method called native exoglycosidase sequencing-mass spectrometry (NES-MS) is established to better address the comprehensive structure of heavily glycosylated proteins and the topology of glycans. The topology and heterogeneity of glycans is highly

correlated to the spatial structure of proteins and their ligand binding propeorties. Our method is designed to divulge glycan structure and occupancy information simultaneously at the native intact glycoprotein level. NES-MS, as a consequence, identify glycoforms or structural features that were previously unknown or contested in the literature. These features include the relationship of fetuin N- and O-glycosylation, amount of LacNAc extensions on EPO and the presence of core-fucose and Lewis^{y} epitopes on AGP and branching. It is important to note the value of a native approach opposed to a denaturing measurement largely focusing on the glycan itself. This allows for the measurement of non-covalent complexes, for example, glycan-specific interactions with small molecules, glycan-specified interactions between proteins such as with lectin, or protein aggregation mediated by glycan structures. It is a further opportunity for further exploration and development – testing the range of exoglycosidases, their numerous combinations and the countless glycoprotein structures to be discovered by NES-MS.

In addition to rhodopsin membranes and glycoproteins, we extend the application of native MS to potassium ion channels and attempt to discover its endogenous ligands that regulate circadian rhythyms and enable drug candidate development. As shown in Chapter 6 in the example of the voltage-dependent potassium channel beta subunit, which we observed binding to NADPH, and an unknown small chemical. In Chapter 6, several applications of native MS are employed for the study of the ion channel structure, redox state of aldo-keto reductases and discovery of endogenous ligands attached to Kvβ2. We attempted to apply native MS to probe these endogenous ligands, potentially as well as intermediate or transient metabolites. This study hopefully expands application of native mass spectrometry to metabolomics but is distinct from untargeted metabolomics or target metabolomics by quantitatively measuring hypothetical metabolites. We

attempted to identify endogenous ligands from a target protein by native MS, also named nativeomics, will enable discovery and identification of small ligands, even intermediate metabolites, that are pathologically or physiologically significant. Continual instrumental improvement is required to achieve this and ongoing enhancement in data analysis tool are required. A delicate and robust ion transmission design is required to ensure that low-mass ligands transfer efficiently to the detector after dissociation from the proteins for reduced loss of ligands. Also needed is a broad dynamic range of detection and efficient MS^n activation of small endogenous ligands. The difficulty in analysing MS^n spectra is estimated to be an analogous to that of untargeted metabolomics where only 10% of molecular features can be annotated[415]. To enhance the usage and analysis of chemical fragments generated from MS^n, a reference data-driven workflow is essential and urgently required e.g. standard electron ionization mass spectra as an analytical reference for small chemicals. These current challenges, as well as opportunities, will eventually elevate the power of native mass spectrometry (or nativeomics) so that physiologically significant ligands could be discovered and this potentially advance biology.

Collectively, this dissertation contributes to expanding the application of native mass spectrometry in receptor activation and signalling in the native membrane, without using detergents, along with monitoring and evaluation of active compounds, structural details of universal glycoproteins, and probing physiologically important endogenous ligands for drug prototypes. Native mass spectrometry is addressing more and more challenges in biology to probe signalling across native membrane and endogenous ligands in native environments.

We anticipate that native MS not only acts as a complementary tool for structure biology, but also potentially goes above protein structures for new discoveries in protein function, integrated with

other MS-based techniques, such as native electrospray ion-beam deposition[416,417], limited proteolysis-coupled mass spectrometry[418], hydrogen−deuterium exchange[419], hydroxy radical footprinting[420,421], fast photochemical oxidation of proteins[422] as well as coupling microfluidics with MS for ultrahigh-throughput analysis of complex mixtures[423]. Exciting times lie ahead for further development of native mass spectrometry applied to wild-type receptors in native environments, potentially including intact glycosylation, as explored in this thesis, for therapeutic discovery and intervention.

References

1 Brent, M. R. How does eukaryotic gene prediction work? *Nature Biotechnology* **25**, 883-885, (2007).

2 Carninci, P. *et al.* The Transcriptional Landscape of the Mammalian Genome. **309**, 1559-1563, (2005).

3 Venter, J. C. *et al.* The Sequence of the Human Genome. *Science* **291**, 1304-1351, (2001).

4 Strauss, O. The Retinal Pigment Epithelium in Visual Function. *Physiological Reviews* **85**, 845-881, (2005).

5 Cournia, Z. *et al.* Membrane Protein Structure, Function, and Dynamics: a Perspective from Experiments and Theory. *The Journal of Membrane Biology* **248**, 611-640, (2015).

6 Erdogmus, S. *et al.* Helix 8 is the essential structural motif of mechanosensitive GPCRs. *Nature Communications* **10**, 5784, (2019).

7 van Rhee, A. M. & Jacobson, K. A. Molecular architecture of G protein-coupled receptors. *Drug Development Research* **37**, 1-38, (1996).

8 Sanchez-Reyes, O. B. *et al.* G Protein-Coupled Receptors Contain Two Conserved Packing Clusters. *Biophysical Journal* **112**, 2315-2326, (2017).

9 Yeagle, P. L. & Albert, A. D. G-protein coupled receptor structure. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1768**, 808-824, (2007).

10 Welte, W., Nestel, U., Wacker, T. & Diederichs, K. Structure and function of the porin channel. *Kidney International* **48**, 930-940, (1995).

11 Bainbridge, G., Gokce, I. & Lakey, J. H. Voltage gating is a fundamental feature of porin and toxin β-barrel membrane channels. *FEBS Letters* **431**, 305-308, (1998).

12 Fairman, J. W., Noinaj, N. & Buchanan, S. K. The structural biology of β-barrel membrane proteins: a summary of recent reports. *Current Opinion in Structural Biology* **21**, 523-531, (2011).

13 Lillemeier Björn, F., Pfeiffer Janet, R., Surviladze, Z., Wilson Bridget, S. & Davis Mark, M. Plasma membrane-associated proteins are clustered into islands attached to the cytoskeleton. *Proceedings of the National Academy of Sciences* **103**, 18992-18997, (2006).

14 Brunsveld, L., Waldmann, H. & Huster, D. Membrane binding of lipidated Ras peptides and proteins — The structural point of view. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1788**, 273-288, (2009).

15 von Heijne, G. Membrane-protein topology. *Nature Reviews Molecular Cell Biology* **7**, 909-918, (2006).

16 Vögler, O., Barceló, J. M., Ribas, C. & Escribá, P. V. Membrane interactions of G proteins and other related proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778**, 1640-1652, (2008).

17 Zhang, Z. *et al.* Domain Organization and Conformational Plasticity of the G Protein Effector, PDE6. *Journal of Biological Chemistry* **290**, 12833-12843, (2015).

18 Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. The role of hydrophobic interactions in positioning of peripheral proteins in membranes. *BMC Structural Biology* **7**, 44, (2007).

19 Wee, C. L., Balali-Mood, K., Gavaghan, D. & Sansom, M. S. P. The Interaction of Phospholipase A2 with a Phospholipid Bilayer: Coarse-Grained Molecular Dynamics Simulations. *Biophysical Journal* **95**, 1649-1657, (2008).

20 Sriram, K. & Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Molecular Pharmacology* **93**, 251, (2018).
21 Superti-Furga G Fau - Lackner, D. *et al.* The RESOLUTE consortium: unlocking SLC transporters for drug discovery. *Nature* **19**, 2, (2020).
22 Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery* **16**, 19-34, (2017).
23 Palczewski, K. G Protein–Coupled Receptor Rhodopsin. *Annual Review of Biochemistry* **75**, 743-767, (2006).
24 Spehr, M. & Munger, S. D. Olfactory receptors: G protein-coupled receptors and beyond. *Journal of Neurochemistry* **109**, 1570-1583, (2009).
25 Gravina, S. A., Yep Gl Fau - Khan, M. & Khan, M. Human biology of taste. *Ann Saudi Med.* **33**, 217-222, (2013).
26 Horn, F. *et al.* GPCRDB information system for G protein-coupled receptors. *Nucleic Acids Research* **31**, 294-297, (2003).
27 Davies, M. N. *et al.* On the hierarchical classification of G protein-coupled receptors. *Bioinformatics* **23**, 3113-3118, (2007).
28 Samama, P., Pei, G., Costa, T., Cotecchia, S. & Lefkowitz, R. J. Negative antagonists promote an inactive conformation of the beta 2-adrenergic receptor. *Molecular Pharmacology* **45**, 390, (1994).
29 Chidiac, P., Hebert, T. E., Valiquette, M., Dennis, M. & Bouvier, M. Inverse agonist activity of beta-adrenergic antagonists. *Molecular Pharmacology* **45**, 490, (1994).
30 Gainetdinov, R. R., Premont, R. T., Bohn, L. M., Lefkowitz, R. J. & Caron, M. G. Desensitization of G Protein-Coupled Receptors and Neuronal Functions. *Annual Review of Neuroscience* **27**, 107-144, (2004).
31 Hanyaloglu, A. C. & Zastrow, M. v. Regulation of GPCRs by Endocytic Membrane Trafficking and Its Potential Implications. *Annual Review of Pharmacology and Toxicology* **48**, 537-568, (2008).
32 Rosenbaum, D. M., Rasmussen, S. G. F. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356-363, (2009).
33 Vasiliou, V., Vasiliou, K. & Nebert, D. W. Human ATP-binding cassette (ABC) transporter family. *Human Genomics* **3**, 281, (2009).
34 Calisto, F., Sousa, F. M., Sena, F. V., Refojo, P. N. & Pereira, M. M. Mechanisms of Energy Transduction by Charge Translocating Membrane Proteins. *Chemical Reviews* **121**, 1804-1844, (2021).
35 Armstrong, C. M. & Hille, B. Voltage-Gated Ion Channels and Electrical Excitability. *Neuron* **20**, 371-380, (1998).
36 Pizzagalli, M. D., Bensimon, A. & Superti-Furga, G. A guide to plasma membrane solute carrier proteins. *The FEBS Journal* **288**, 2784-2835, (2021).
37 Ayka, A., ehirli, A. & zer. The Role of the SLC Transporters Protein in the Neurodegenerative Disorders. *Clin Psychopharmacol Neurosci* **18**, 174-187, (2020).
38 Zhang, Y., Zhang, Y., Sun, K., Meng, Z. & Chen, L. The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *Journal of Molecular Cell Biology* **11**, 1-13, (2019).
39 Yan, A. & Lennarz, W. J. J. J. o. B. C. Unraveling the Mechanism of Protein N-Glycosylation*. **280**, 3121 - 3124, (2005).

40 Bieberich, E. in *Glycobiology of the Nervous System* (eds Robert K. Yu & Cara-Lynne Schengrund) 47-70 (Springer New York, 2014).
41 Wolfert, M. A. & Boons, G.-J. Adaptive immune activation: glycosylation does matter. *Nature Chemical Biology* **9**, 776-784, (2013).
42 Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annual Review of Biochemistry* **77**, 521-555, (2008).
43 Bojarová, P. & Křen, V. Glycosidases: a key to tailored carbohydrates. *Trends in Biotechnology* **27**, 199-209, (2009).
44 Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nature Reviews Molecular Cell Biology* **13**, 448-462, (2012).
45 Stanley, P. Golgi glycosylation. *Cold Spring Harb Perspect Biol.* **3**, a005199, (2011).
46 Ohtsubo, K. & Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **126**, 855-867, (2006).
47 Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nature Reviews Nephrology* **15**, 346-366, (2019).
48 Parodi, A. J. Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *Biochem J* **348 Pt 1**, 1-13, (2000).
49 Hoffman, S. & Edelman, G. M. Kinetics of homophilic binding by embryonic and adult forms of the neural cell adhesion molecule. *Proceedings of the National Academy of Sciences* **80**, 5762-5766, (1983).
50 Lee Sena, J. *et al.* Mannose Receptor-Mediated Regulation of Serum Glycoprotein Homeostasis. *Science* **295**, 1898-1901, (2002).
51 Belenkaya, T. Y. *et al.* Drosophila Dpp Morphogen Movement Is Independent of Dynamin-Mediated Endocytosis but Regulated by the Glypican Members of Heparan Sulfate Proteoglycans. *Cell* **119**, 231-244, (2004).
52 Ornitz, D. M. *et al.* Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Molecular and Cellular Biology* **12**, 240-247, (1992).
53 Tedder, T. F., Tuscano, J. M., Sato, S. & Kehrl, J. H. J. A. r. o. i. CD22, a B lymphocyte-specific adhesion molecule that regulates antigen receptor signaling. **15**, 481-504, (1997).
54 Partridge Emily, A. *et al.* Regulation of Cytokine Receptors by Golgi N-Glycan Processing and Endocytosis. *Science* **306**, 120-124, (2004).
55 Suzuki, N. Glycan diversity in the course of vertebrate evolution. *Glycobiology* **29**, 625-644, (2019).
56 Amano, M. *et al.* Threshold in Stage-specific Embryonic Glycotypes Uncovered by a Full Portrait of Dynamic N-Glycan Expression during Cell Differentiation *. *Molecular & Cellular Proteomics* **9**, 523-537, (2010).
57 Chen, Y. *et al.* Role of malectin in Glc2Man9GlcNAc2-dependent quality control of α1-antitrypsin. *Molecular Biology of the Cell* **22**, 3559-3570, (2011).
58 Dall'Olio, F., Malagolini, N., Trinchera, M. & Chiricolo, M. Mechanisms of cancer-associated glycosylation changes. **17**, 670-699, (2012).
59 Toegel, S. *et al.* Phenotype-related differential α-2,6- or α-2,3-sialylation of glycoprotein N-glycans in human chondrocytes. *Osteoarthritis and Cartilage* **18**, 240-248, (2010).

60 Seeberger, P. H. & Werz, D. B. Automated synthesis of oligosaccharides as a basis for drug discovery. *Nature Reviews Drug Discovery* **4**, 751-763, (2005).
61 Schweizer, F. & Hindsgaul, O. Combinatorial synthesis of carbohydrates. *Current Opinion in Chemical Biology* **3**, 291-298, (1999).
62 Chen, S., Wu, D., Robinson, C. V. & Struwe, W. B. Native Mass Spectrometry Meets Glycomics: Resolving Structural Detail and Occupancy of Glycans on Intact Glycoproteins. *Analytical Chemistry* **93**, 10435-10443, (2021).
63 Awad, H., Khamis, M. M. & El-Aneed, A. Mass Spectrometry, Review of the Basics: Ionization. *Applied Spectroscopy Reviews* **50**, 158-175, (2015).
64 Barber, M., Bordoli, R. S., Elliott, G. J., Sedgwick, R. D. & Tyler, A. N. Fast atom bombardment mass spectrometry. *Analytical Chemistry* **54**, 645-657, (1982).
65 Fenn John, B., Mann, M., Meng Chin, K., Wong Shek, F. & Whitehouse Craig, M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **246**, 64-71, (1989).
66 Banerjee, S. & Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry* **2012**, 282574, (2012).
67 Spengler, B. *et al.* The detection of large molecules in matrix-assisted UV-laser desorption. *Rapid Communications in Mass Spectrometry* **4**, 301-305, (1990).
68 Mädler, S., Erba, E. B. & Zenobi, R. in *Applications of MALDI-TOF Spectroscopy* (eds Zongwei Cai & Shuying Liu) 1-36 (Springer Berlin Heidelberg, 2013).
69 Mager, F., Sokolova, L., Lintzel, J., Brutschy, B. & Nussberger, S. LILBID-mass spectrometry of the mitochondrial preprotein translocase TOM. *Journal of Physics: Condensed Matter* **22**, 454132, (2010).
70 Peetz, O. *et al.* LILBID and nESI: Different Native Mass Spectrometry Techniques as Tools in Structural Biology. *Journal of The American Society for Mass Spectrometry* **30**, 181-191, (2019).
71 Ambrose, S. *et al.* Native Desorption Electrospray Ionization Liberates Soluble and Membrane Protein Complexes from Surfaces. *Angewandte Chemie International Edition* **56**, 14463-14468, (2017).
72 Chen, G. *et al.* Advances in MS Based Strategies for Probing Ligand-Target Interactions: Focus on Soft Ionization Mass Spectrometric Techniques. *Frontiers in Chemistry* **7** (2019).
73 Cech, N. B. & Enke, C. G. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrometry Reviews* **20**, 362-387, (2001).
74 Rohner, T. C., Lion, N. & Girault, H. H. Electrochemical and theoretical aspects of electrospray ionisation. *Physical Chemistry Chemical Physics* **6**, 3056-3068, (2004).
75 Rayleigh, L. J. P. M. S. XX. On the equilibrium of liquid conducting masses charged with electricity. **14**, 184-186.
76 Tang, K. & Smith, R. D. Physical/chemical separations in the break-up of highly charged droplets from electrosprays. *Journal of the American Society for Mass Spectrometry* **12**, 343-347, (2001).
77 Nakorchevsky, A. & Yates, J. R. in *Comprehensive Biophysics* (ed Edward H. Egelman) 341-375 (Elsevier, 2012).

78 Wilm, M. S. & Mann, M. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes* **136**, 167-180, (1994).
79 Wilm, M. & Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry* **68**, 1-8, (1996).
80 Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics* **64**, 2287-2294, (1976).
81 Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Analytical Chemistry* **85**, 2-9, (2013).
82 Iavarone, A. T. & Williams, E. R. Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *Journal of the American Chemical Society* **125**, 2319-2327, (2003).
83 Mehmood, S., Allison, T. M. & Robinson, C. V. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annual Review of Physical Chemistry* **66**, 453-474, (2015).
84 Fernandez de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica Chimica Acta* **406**, 93-104, (2000).
85 Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *International Journal of Mass Spectrometry* **219**, 39-62, (2002).
86 Konermann, L., Rodriguez, A. D. & Liu, J. On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. *Analytical Chemistry* **84**, 6798-6804, (2012).
87 Zubarev, R. A. & Makarov, A. Orbitrap Mass Spectrometry. *Analytical Chemistry* **85**, 5288-5296, (2013).
88 Yost, R. A. & Enke, C. G. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Analytical Chemistry* **51**, 1251-1264, (1979).
89 Dawson, P. H. Quadrupole mass analyzers: Performance, design and some recent applications. *Mass Spectrometry Reviews* **5**, 1-37, (1986).
90 Yu, F. *et al.* Mass measurement accuracy of the Orbitrap in intact proteome analysis. *Rapid Communications in Mass Spectrometry* **30**, 1391-1397, (2016).
91 Adams, F. & Barbante, C. in *Comprehensive Analytical Chemistry* Vol. 69 (eds Freddy Adams & Carlo Barbante) 67-124 (Elsevier, 2015).
92 de Hoffmann, E. & Stroobant, V. *Mass Spectrometry: Principles and Applications*. (Wiley, 2013).
93 Henchman, M. & Steel, C. Understanding the Quadrupole Mass Filter through Computer Simulation. *Journal of Chemical Education* **75**, 1049, (1998).
94 Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chemical Reviews* **107**, 3544-3567, (2007).
95 Marshall, A. G. Fourier transform ion cyclotron resonance mass spectrometry. *Accounts of Chemical Research* **18**, 316-322, (1985).
96 Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry* **72**, 1156-1162, (2000).

97 Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry* **40**, 430-443, (2005).
98 Boone, C. & Adamec, J. in *Proteomic Profiling and Analytical Chemistry (Second Edition)* (eds P. Ciborowski & J. Silberring) 175-191 (Elsevier, 2016).
99 Ben-Nissan, G. *et al.* Triple-Stage Mass Spectrometry Unravels the Heterogeneity of an Endogenous Protein Complex. *Analytical Chemistry* **89**, 4708-4715, (2017).
100 Belov, M. E. *et al.* From Protein Complexes to Subunit Backbone Fragments: A Multi-stage Approach to Native Mass Spectrometry. *Analytical Chemistry* **85**, 11163-11173, (2013).
101 van de Waterbeemd, M. *et al.* High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles. *Nature Methods* **14**, 283-286, (2017).
102 Liko, I., Allison, T. M., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry guided structural biology. *Current Opinion in Structural Biology* **40**, 136-144, (2016).
103 Robitaille, A. M., Huguet, R., Bomgarden, R. D., Canterbury, J. D. & Lopez-Ferrer, D.
104 Morris, M., Thibault, P. & Boyd, R. K. Characterization of a high-pressure quadrupole collision cell for low-energy collision-indneed dissociation. *Journal of the American Society for Mass Spectrometry* **5**, 1042-1063, (1994).
105 Shinholt, D. L., Anthony, S. N., Alexander, A. W., Draper, B. E. & Jarrold, M. F. A frequency and amplitude scanned quadrupole mass filter for the analysis of high m/z ions. *Review of Scientific Instruments* **85**, 113109, (2014).
106 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* **113**, 8534-8535, (1991).
107 Ganem, B., Li, Y. T. & Henion, J. D. Detection of noncovalent receptor-ligand complexes by mass spectrometry. *Journal of the American Chemical Society* **113**, 6294-6296, (1991).
108 Van Berkel, W. J. H., Van Den Heuvel, R. H. H., Versluis, C. & Heck, A. J. R. Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry. *Protein Science* **9**, 435-439, (2000).
109 Rostom Adam, A. *et al.* Detection and selective dissociation of intact ribosomes in a mass spectrometer. *Proceedings of the National Academy of Sciences* **97**, 5185-5190, (2000).
110 Tito, M. A., Tars, K., Valegard, K., Hajdu, J. & Robinson, C. V. Electrospray Time-of-Flight Mass Spectrometry of the Intact MS2 Virus Capsid. *Journal of the American Chemical Society* **122**, 3550-3551, (2000).
111 Barrera Nelson, P., Di Bartolo, N., Booth Paula, J. & Robinson Carol, V. Micelles Protect Membrane Complexes from Solution to Vacuum. *Science* **321**, 243-246, (2008).
112 Norris, C. E. *et al.* Native mass spectrometry reveals the simultaneous binding of lipids and zinc to rhodopsin. *International Journal of Mass Spectrometry* **460**, 116477, (2021).
113 Rose, R. J., Damoc, E., Denisov, E., Makarov, A. & Heck, A. J. R. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nature Methods* **9**, 1084-1086, (2012).
114 Yen, H.-Y. *et al.* Correlating Glycoforms of DC-SIGN with Stability Using a Combination of Enzymatic Digestion and Ion Mobility Mass Spectrometry. *Angewandte Chemie International Edition* **59**, 15560-15564, (2020).
115 Yen, H.-Y. *et al.* PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, 423-427, (2018).

116 Wu, D., Li, J., Struwe, W. B. & Robinson, Carol V. Probing N-glycoprotein microheterogeneity by lectin affinity purification-mass spectrometry analysis. *Chemical Science* **10**, 5146-5155, (2019).
117 Chorev, D. S. *et al.* Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science* **362**, 829, (2018).
118 Bolla, J. R. *et al.* Direct observation of the influence of cardiolipin and antibiotics on lipid II binding to MurJ. *Nature Chemistry* **10**, 363-371, (2018).
119 Abbas Yazan, M., Wu, D., Bueler Stephanie, A., Robinson Carol, V. & Rubinstein John, L. Structure of V-ATPase from the mammalian brain. *Science* **367**, 1240-1246, (2020).
120 Chen, S. *et al.* Capturing a rhodopsin receptor signalling cascade across a native membrane. *Nature* **604**, 384-390, (2022).
121 Fiorentino, F. *et al.* Dynamics of an LPS translocon induced by substrate and an antimicrobial peptide. *Nature Chemical Biology* **17**, 187-195, (2021).
122 Agasid, M. T., Sørensen, L., Urner, L. H., Yan, J. & Robinson, C. V. The Effects of Sodium Ions on Ligand Binding and Conformational States of G Protein-Coupled Receptors—Insights from Mass Spectrometry. *Journal of the American Chemical Society* **143**, 4085-4089, (2021).
123 Oluwole, A. O. *et al.* Peptidoglycan biosynthesis is driven by lipid transfer along enzyme-substrate affinity gradients. *Nature Communications* **13**, 2278, (2022).
124 Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nature Methods* **13**, 333-336, (2016).
125 Tamara, S., den Boer, M. A. & Heck, A. J. R. High-Resolution Native Mass Spectrometry. *Chemical Reviews* **122**, 7269-7326, (2022).
126 Wu, D., Struwe, W. B., Harvey, D. J., Ferguson, M. A. J. & Robinson, C. V. N-glycan microheterogeneity regulates interactions of plasma proteins. *Proceedings of the National Academy of Sciences* **115**, 8763, (2018).
127 Gault, J. *et al.* Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. *Nature Methods* **17**, 505-508, (2020).
128 Chorev, D. S. The use of sonicated lipid vesicles for mass spectrometry of membrane protein complexes. *Nat. Protoc.* **15** (2020).
129 Rozen, S. *et al.* CSNAP Is a Stoichiometric Subunit of the COP9 Signalosome. *Cell Reports* **13**, 585-598, (2015).
130 Jackson, V. A. *et al.* Super-complexes of adhesion GPCRs and neural guidance receptors. *Nature Communications* **7**, 11184, (2016).
131 Olerinyova, A. *et al.* Mass Photometry of Membrane Proteins. *Chem* **7**, 224-236, (2021).
132 Aguilar-Gurrieri, C. *et al.* Structural evidence for Nap1-dependent H2A–H2B deposition and nucleosome assembly. *The EMBO Journal* **35**, 1465-1482, (2016).
133 Pukala, T. & Robinson, C. V. Introduction: Mass Spectrometry Applications in Structural Biology. *Chemical Reviews* **122**, 7267-7268, (2022).
134 Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **541**, 421-424, (2017).
135 Gupta, K. *et al.* Identifying key membrane protein lipid interactions using mass spectrometry. *Nature Protocols* **13**, 1106-1120, (2018).

136 Mehmood, S. *et al.* Mass spectrometry captures off-target drug binding and provides mechanistic insights into the human metalloprotease ZMPSTE24. *Nature Chemistry* **8**, 1152-1158, (2016).
137 Bechara, C. & Robinson, C. V. Different Modes of Lipid Binding to Membrane Proteins Probed by Mass Spectrometry. *Journal of the American Chemical Society* **137**, 5240-5247, (2015).
138 Powl, A. M., East, J. M. & Lee, A. G. Importance of Direct Interactions with Lipids for the Function of the Mechanosensitive Channel MscL. *Biochemistry* **47**, 12175-12184, (2008).
139 Pyle, E. *et al.* Structural Lipids Enable the Formation of Functional Oligomers of the Eukaryotic Purine Symporter UapA. *Cell Chemical Biology* **25**, 840-848.e844, (2018).
140 Yeung, Y.-G., Nieves, E., Angeletti, R. H. & Stanley, E. R. Removal of detergents from protein digests for mass spectrometry analysis. *Analytical Biochemistry* **382**, 135-137, (2008).
141 Lee, A. G. Lipid–protein interactions in biological membranes: a structural perspective. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1612**, 1-40, (2003).
142 Hendrickson, W. A. Atomic-level analysis of membrane-protein structure. *Nature Structural & Molecular Biology* **23**, 464-467, (2016).
143 Urner, L. H., Maier, Y. B., Haag, R. & Pagel, K. Exploring the Potential of Dendritic Oligoglycerol Detergents for Protein Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **30**, 174-180, (2019).
144 Urner, L. H. *et al.* Modular detergents tailor the purification and structural analysis of membrane proteins including G-protein coupled receptors. *Nature Communications* **11**, 564, (2020).
145 Urner, L. H. *et al.* A new azobenzene-based design strategy for detergents in membrane protein research. *Chemical Science* **11**, 3538-3546, (2020).
146 Yen, H.-Y. *et al.* Ligand binding to a G protein–coupled receptor captured in a mass spectrometer. *Science Advances* **3**, e1701016, (2017).
147 Urner, L. H. in *Heterologous Expression of Membrane Proteins: Methods and Protocols* (ed Isabelle Mus-Veteau) 359-374 (Springer US, 2022).
148 Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters* **2**, 853-856, (2002).
149 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* **126**, 3477-3487, (2004).
150 Bayburt, T. H. & Sligar, S. G. Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Science* **12**, 2476-2481, (2003).
151 Marty, M. T., Hoi, K. K., Gault, J. & Robinson, C. V. Probing the Lipid Annular Belt by Gas-Phase Dissociation of Membrane Proteins in Nanodiscs. *Angewandte Chemie International Edition* **55**, 550-554, (2016).
152 Marty, M. T., Hoi, K. K. & Robinson, C. V. Interfacing Membrane Mimetics with Mass Spectrometry. *Accounts of Chemical Research* **49**, 2459-2467, (2016).
153 Keener, J. E. *et al.* Chemical Additives Enable Native Mass Spectrometry Measurement of Membrane Protein Oligomeric State within Intact Nanodiscs. *Journal of the American Chemical Society* **141**, 1054-1061, (2019).

154 Hoi, K. K., Robinson, C. V. & Marty, M. T. Unraveling the Composition and Behavior of Heterogeneous Lipid Nanodiscs by Mass Spectrometry. *Analytical Chemistry* **88**, 6199-6204, (2016).
155 Kostelic, M. M., Ryan, A. M., Reid, D. J., Noun, J. M. & Marty, M. T. Expanding the Types of Lipids Amenable to Native Mass Spectrometry of Lipoprotein Complexes. *Journal of The American Society for Mass Spectrometry* **30**, 1416-1425, (2019).
156 Marty, M. T. Nanodiscs and mass spectrometry: Making membranes fly. *International Journal of Mass Spectrometry* **458**, 116436, (2020).
157 Carlson, M. L. *et al.* The Peptidisc, a simple method for stabilizing membrane proteins in detergent-free solution. *eLife* **7**, e34085, (2018).
158 Leney, A. C. *et al.* Picodiscs for Facile Protein-Glycolipid Interaction Analysis. *Analytical Chemistry* **87**, 4402-4408, (2015).
159 Li, J. *et al.* Detecting Protein–Glycolipid Interactions Using CaR-ESI-MS and Model Membranes: Comparison of Pre-loaded and Passively Loaded Picodiscs. *Journal of the American Society for Mass Spectrometry* **29**, 1493-1504, (2018).
160 Ravula, T., Hardin, N. Z. & Ramamoorthy, A. Polymer nanodiscs: Advantages and limitations. *Chemistry and Physics of Lipids* **219**, 45-49, (2019).
161 Overduin, M. & Esmaili, M. Native Nanodiscs and the Convergence of Lipidomics, Metabolomics, Interactomics and Proteomics. *Applied Sciences* **9** (2019).
162 Hellwig, N. *et al.* Native mass spectrometry goes more native: investigation of membrane protein complexes directly from SMALPs. *Chemical Communications* **54**, 13702-13705, (2018).
163 Hesketh, S. J. *et al.* Styrene maleic-acid lipid particles (SMALPs) into detergent or amphipols: An exchange protocol for membrane protein characterisation. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1862**, 183192, (2020).
164 Quetschlich, D. *et al.* NaViA: a program for the visual analysis of complex mass spectra. *Bioinformatics* **37**, 4876-4878, (2021).
165 Steinberg, R. H., Fisher, S. K. & Anderson, D. H. Disc morphogenesis in vertebrate photoreceptors. *Journal of Comparative Neurology* **190**, 501-518, (1980).
166 Carter-Dawson, L. D. & Lavail, M. M. Rods and cones in the mouse retina. I. Structural analysis using light and electron microscopy. *Journal of Comparative Neurology* **188**, 245-262, (1979).
167 Arikawa, K., Molday, L. L., Molday, R. S. & Williams, D. S. Localization of peripherin/rds in the disk membranes of cone and rod photoreceptors: relationship to disk membrane morphogenesis and retinal degeneration. *Journal of Cell Biology* **116**, 659-667, (1992).
168 Patil, H. *et al.* Selective loss of RPGRIP1-dependent ciliary targeting of NPHP4, RPGR and SDCCAG8 underlies the degeneration of photoreceptor neurons. *Cell Death & Disease* **3**, e355-e355, (2012).
169 Chuang, J.-Z., Zhao, Y. & Sung, C.-H. SARA-Regulated Vesicular Targeting Underlies Formation of the Light-Sensing Organelle in Mammalian Rods. *Cell* **130**, 535-547, (2007).
170 Burgoyne, T. *et al.* Rod disc renewal occurs by evagination of the ciliary plasma membrane that makes cadherin-based contacts with the inner segment. *Proceedings of the National Academy of Sciences* **112**, 15922-15927, (2015).
171 Volland, S. *et al.* Three-dimensional organization of nascent rod outer segment disk membranes. *Proceedings of the National Academy of Sciences* **112**, 14870-14875, (2015).

172 Ding, J.-D., Salinas, R. Y. & Arshavsky, V. Y. Discs of mammalian rod photoreceptors form through the membrane evagination mechanism. *Journal of Cell Biology* **211**, 495-502, (2015).
173 Young, R. W. & Bok, D. PARTICIPATION OF THE RETINAL PIGMENT EPITHELIUM IN THE ROD OUTER SEGMENT RENEWAL PROCESS. *Journal of Cell Biology* **42**, 392-403, (1969).
174 Young, R. W. & Bok, D. Autoradiographic Studies on the Metabolism of the Retinal Pigment Epithelium. *Investigative Ophthalmology & Visual Science* **9**, 524-536, (1970).
175 Kevany, B. M. & Palczewski, K. Phagocytosis of Retinal Rod and Cone Photoreceptors. *Physiology* **25**, 8-15, (2010).
176 Young, R. W. The daily rhythm of shedding and degradation of rod and cone outer segment membranes in the chick retina. *Investigative Ophthalmology & Visual Science* **17**, 105-116, (1978).
177 O'Day, W. T. & Young, R. W. Rhythmic daily shedding of outer-segment membranes by visual cells in the goldfish. *Journal of Cell Biology* **76**, 593-604, (1978).
178 Young, R. W. The daily rhythm of shedding and degradation of cone outer segment membranes in the lizard retina. *Journal of Ultrastructure Research* **61**, 172-185, (1977).
179 Johnson, N. F. Phagocytosis in the normal and ischaemic retinal pigment epithelium of the rabbit. *Experimental Eye Research* **20**, 97-107, (1975).
180 Fisher, S. K., Pfeffer, B. A. & Anderson, D. H. Both rod and cone disc shedding are related to light onset in the cat. *Investigative Ophthalmology & Visual Science* **24**, 844-856, (1983).
181 LaVail, M. M. Rod Outer Segment Disk Shedding in Rat Retina: Relationship to Cyclic Lighting. *Science* **194**, 1071-1074, (1976).
182 Rakshit, T. & Park, P. S. H. Impact of Reduced Rhodopsin Expression on the Structure of Rod Outer Segment Disc Membranes. *Biochemistry* **54**, 2885-2894, (2015).
183 Whited, A. M. & Park, P. S. H. Nanodomain organization of rhodopsin in native human and murine rod outer segment disc membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1848**, 26-34, (2015).
184 Papermaster, D. S., Schneider, B. G., Zorn, M. A. & Kraehenbuhl, J. P. Immunocytochemical localization of opsin in outer segments and Golgi zones of frog photoreceptor cells. An electron microscope analysis of cross-linked albumin-embedded retinas. *Journal of Cell Biology* **77**, 196-210, (1978).
185 Moritz, O. L. & Molday, R. S. Molecular cloning, membrane topology, and localization of bovine rom-1 in rod and cone photoreceptor cells. *Investigative Ophthalmology & Visual Science* **37**, 352-362, (1996).
186 Goldberg, A. F. X. & Molday, R. S. Subunit Composition of the Peripherin/rds−Rom-1 Disk Rim Complex from Rod Photoreceptors: Hydrodynamic Evidence for a Tetrameric Quaternary Structure. *Biochemistry* **35**, 6144-6149, (1996).
187 Loewen, C. J. R. & Molday, R. S. Disulfide-mediated Oligomerization of Peripherin/Rds and Rom-1 in Photoreceptor Disk Membranes: IMPLICATIONS FOR PHOTORECEPTOR OUTER SEGMENT MORPHOGENESIS AND DEGENERATION *. *Journal of Biological Chemistry* **275**, 5370-5378, (2000).
188 Kevany, B. M. *et al.* Structural and Functional Analysis of the Native Peripherin-ROM1 Complex Isolated from Photoreceptor Cells *. *Journal of Biological Chemistry* **288**, 36272-36284, (2013).

189 Zulliger, R., Conley, S. M., Mwoyosvi, M. L., Al-Ubaidi, M. R. & Naash, M. I. Oligomerization of Prph2 and Rom1 is essential for photoreceptor outer segment formation. *Human Molecular Genetics* **27**, 3507-3518, (2018).
190 Milstein, M. L. *et al.* Multistep peripherin-2/rds self-assembly drives membrane curvature for outer segment disk architecture and photoreceptor viability. *Proceedings of the National Academy of Sciences* **117**, 4400-4410, (2020).
191 Quazi, F., Lenevich, S. & Molday, R. S. ABCA4 is an N-retinylidene-phosphatidylethanolamine and phosphatidylethanolamine importer. *Nature Communications* **3**, 925, (2012).
192 Quazi, F. & Molday, R. S. ATP-binding cassette transporter ABCA4 and chemical isomerization protect photoreceptor cells from the toxic accumulation of excess 11-cis-retinal. *Proceedings of the National Academy of Sciences* **111**, 5024-5029, (2014).
193 Beharry, S., Zhong, M. & Molday, R. S. Retinylidene-phosphatidylethanolamine Is the Preferred Retinoid Substrate for the Photoreceptor-specific ABC Transporter ABCA4 (ABCR) *Journal of Biological Chemistry* **279**, 53972-53979, (2004).
194 Park, P. S. H. Supramolecular organization of rhodopsin in rod photoreceptor cell membranes. *Pflügers Archiv - European Journal of Physiology* **473**, 1361-1376, (2021).
195 Herrick-Davis, K. *et al.* Native Serotonin 5-HT2C Receptors Are Expressed as Homodimers on the Apical Surface of Choroid Plexus Epithelial Cells. *Molecular Pharmacology* **87**, 660, (2015).
196 Fotiadis, D. *et al.* Rhodopsin dimers in native disc membranes. *Nature* **421**, 127-128, (2003).
197 Liang, Y. *et al.* Organization of the G Protein-coupled Receptors Rhodopsin and Opsin in Native Membranes *. *Journal of Biological Chemistry* **278**, 21655-21662, (2003).
198 Liang, Y. *et al.* Rhodopsin Signaling and Organization in Heterozygote Rhodopsin Knockout Mice
Journal of Biological Chemistry **279**, 48189-48196, (2004).
199 Fotiadis, D. *et al.* The G protein-coupled receptor rhodopsin in the native membrane. *FEBS Letters* **564**, 281-288, (2004).
200 Buzhynskyy, N., Salesse, C. & Scheuring, S. Rhodopsin is spatially heterogeneously distributed in rod outer segment disk membranes. *Journal of Molecular Recognition* **24**, 483-489, (2011).
201 Baylor, D. A., Lamb, T. D. & Yau, K. W. Responses of retinal rods to single photons. *The Journal of Physiology* **288**, 613-634, (1979).
202 Lambert, N. A. & Javitch, J. A. CrossTalk opposing view: Weighing the evidence for class A GPCR dimers, the jury is still out. *The Journal of Physiology* **592**, 2443-2445, (2014).
203 Mansoor, S. E., Palczewski, K. & Farrens, D. L. Rhodopsin self-associates in asolectin liposomes. *Proceedings of the National Academy of Sciences* **103**, 3060-3065, (2006).
204 Zhang, T. *et al.* Dimerization of visual pigments in vivo. *Proceedings of the National Academy of Sciences* **113**, 9093-9098, (2016).
205 Zhao, D. Y. *et al.* Cryo-EM structure of the native rhodopsin dimer in nanodiscs. *Journal of Biological Chemistry* **294**, 14215-14230, (2019).
206 Palczewski, K. *et al.* Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **289**, 739-745, (2000).

207 Poo, M.-m. & Cone, R. A. Lateral diffusion of rhodopsin in the photoreceptor membrane. *Nature* **247**, 438-441, (1974).
208 Chabre, M. & le Maire, M. Monomeric G-Protein-Coupled Receptor as a Functional Unit. *Biochemistry* **44**, 9395-9403, (2005).
209 Whorton, M. R. *et al.* Efficient Coupling of Transducin to Monomeric Rhodopsin in a Phospholipid Bilayer*. *Journal of Biological Chemistry* **283**, 4387-4394, (2008).
210 Ernst, O. P., Gramse, V., Kolbe, M., Hofmann, K. P. & Heck, M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proceedings of the National Academy of Sciences* **104**, 10859-10864, (2007).
211 Banerjee, S., Huber, T. & Sakmar, T. P. Rapid Incorporation of Functional Rhodopsin into Nanoscale Apolipoprotein Bound Bilayer (NABB) Particles. *Journal of Molecular Biology* **377**, 1067-1081, (2008).
212 Cone, R. A. Rotational Diffusion of Rhodopsin in the Visual Receptor Membrane. *Nature New Biology* **236**, 39-43, (1972).
213 Chabre, M. X-ray diffraction studies of retinal rods. I. Structure of the disc membrane, effect of illumination. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **382**, 322-335, (1975).
214 Saibil, H., Chabre, M. & Worcester, D. Neutron diffraction studies of retinal rod outer segment membranes. *Nature* **262**, 266-270, (1976).
215 Roof, D. J. & Heuser, J. E. Surfaces of rod photoreceptor disk membranes: integral membrane components. *Journal of Cell Biology* **95**, 487-500, (1982).
216 Gao, Y. *et al.* Isolation and structure-function characterization of a signaling-active rhodopsin-G protein complex. *Journal of Biological Chemistry* **292**, 14280-14289, (2017).
217 Gao, Y. *et al.* Structures of the Rhodopsin-Transducin Complex: Insights into G-Protein Activation. *Molecular Cell* **75**, 781-790.e783, (2019).
218 Chen, Q. *et al.* Structures of rhodopsin in complex with G-protein-coupled receptor kinase 1. *Nature* **595**, 600-605, (2021).
219 Kang, Y. *et al.* A structural snapshot of the rhodopsin–arrestin complex. *The FEBS Journal* **283**, 816-821, (2016).
220 Kang, Y. *et al.* Cryo-EM structure of human rhodopsin bound to an inhibitory G protein. *Nature* **558**, 553-558, (2018).
221 Tsukamoto, H., Sinha, A., DeWitt, M. & Farrens, D. L. Monomeric Rhodopsin Is the Minimal Functional Unit Required for Arrestin Binding. *Journal of Molecular Biology* **399**, 501-511, (2010).
222 Ikuta, T. *et al.* Structural insights into the mechanism of rhodopsin phosphodiesterase. *Nature Communications* **11**, 5605, (2020).
223 Jastrzebska, B. *et al.* Functional Characterization of Rhodopsin Monomers and Dimers in Detergents. *Journal of Biological Chemistry* **279**, 54663-54675, (2004).
224 Calvert, P. D. *et al.* Membrane protein diffusion sets the speed of rod phototransduction. *Nature* **411**, 90-94, (2001).
225 Liebman, P. A. & Entine, G. Lateral Diffusion of Visual Pigment in Photoreceptor Disk Membranes. *Science* **185**, 457-459, (1974).
226 Kukura, P., McCamant, D. W., Yoon, S., Wandschneider, D. B. & Mathies, R. A. Structural Observation of the Primary Isomerization in Vision with Femtosecond-Stimulated Raman. *Science* **310**, 1006-1009, (2005).

227 Yoshizawa, T. r. & Shichida, Y. in *Methods in Enzymology* Vol. 81 333-354 (Academic Press, 1982).
228 Kliger, D. S. & Lewis, J. W. Spectral and Kinetic Characterization of Visual Pigment Photointermediates. *Israel Journal of Chemistry* **35**, 289-307, (1995).
229 Epps, J., Lewis, J. W., Szundi, I. & Kliger, D. S. Lumi I → Lumi II: The Last Detergent Independent Process in Rhodopsin Photoexcitation†. *Photochemistry and Photobiology* **82**, 1436-1441, (2006).
230 Szundi, I., Lewis, J. W. & Kliger, D. S. Two Intermediates Appear on the Lumirhodopsin Time Scale after Rhodopsin Photoexcitation. *Biochemistry* **42**, 5091-5098, (2003).
231 Szundi, I., Lewis, J. W. & Kliger, D. S. Deriving reaction mechanisms from kinetic spectroscopy. Application to late rhodopsin intermediates. *Biophysical Journal* **73**, 688-702, (1997).
232 Szundi, I. *et al.* Proton Transfer Reactions Linked to Rhodopsin Activation. *Biochemistry* **37**, 14237-14244, (1998).
233 Tachibanaki, S., Imai, H., Terakita, A. & Shichida, Y. Identification of a new intermediate state that binds but not activates transducin in the bleaching process of bovine rhodopsin. *FEBS Letters* **425**, 126-130, (1998).
234 Ernst, O. P. *et al.* Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chemical Reviews* **114**, 126-163, (2014).
235 Bartl, F. J. & Vogel, R. Structural and functional properties of metarhodopsin III: Recent spectroscopic studies on deactivation pathways of rhodopsin. *Physical Chemistry Chemical Physics* **9**, 1648-1658, (2007).
236 Vogel, R., Siebert, F., Zhang, X.-Y., Fan, G. & Sheves, M. Formation of Meta III during the Decay of Activated Rhodopsin Proceeds via Meta I and Not via Meta II. *Biochemistry* **43**, 9457-9466, (2004).
237 Vuong, T. M., Chabre, M. & Stryer, L. Millisecond activation of transducin in the cyclic nucleotide cascade of vision. *Nature* **311**, 659-661, (1984).
238 Heck, M. & Hofmann, K. P. Maximal rate and nucleotide dependence of rhodopsin-catalyzed transducin activation: initial rate analysis based on a double displacement mechanism. *Journal of Biological Chemistry* **276**, 10000-10009, (2001).
239 Irwin, M. J. *et al.* The molecular architecture of photoreceptor phosphodiesterase 6 (PDE6) with activated G protein elucidates the mechanism of visual excitation. *Journal of Biological Chemistry* **294**, 19486-19497, (2019).
240 Mendez, A. *et al.* Rapid and Reproducible Deactivation of Rhodopsin Requires Multiple Phosphorylation Sites. *Neuron* **28**, 153-164, (2000).
241 Doan, T., Mendez, A., Detwiler, P. B., Chen, J. & Rieke, F. Multiple Phosphorylation Sites Confer Reproducibility of the Rod's Single-Photon Responses. *Science* **313**, 530-533, (2006).
242 Gurevich, V. V., Hanson, S. M., Song, X., Vishnivetskiy, S. A. & Gurevich, E. V. The functional cycle of visual arrestins in photoreceptor cells. *Progress in Retinal and Eye Research* **30**, 405-430, (2011).
243 Berry, J. *et al.* Effect of Rhodopsin Phosphorylation on Dark Adaptation in Mouse Rods. *The Journal of Neuroscience* **36**, 6973, (2016).
244 Zang, J. & Neuhauss, S. C. F. The Binding Properties and Physiological Functions of Recoverin. *Frontiers in Molecular Neuroscience* **11** (2018).

245 Hsieh, C.-L., Yao, Y., Gurevich, V. V. & Chen, J. Arrestin Facilitates Rhodopsin Dephosphorylation in Vivo. *The Journal of Neuroscience* **42**, 3537, (2022).
246 Lee, K. A., Nawrot, M., Garwin, G. G., Saari, J. C. & Hurley, J. B. Relationships among Visual Cycle Retinoids, Rhodopsin Phosphorylation, and Phototransduction in Mouse Eyes during Light and Dark Adaptation. *Biochemistry* **49**, 2454-2463, (2010).
247 Palczewski, K. & Kiser, P. D. Shedding new light on the generation of the visual chromophore. *Proceedings of the National Academy of Sciences* **117**, 19629-19638, (2020).
248 Hubbard, R. & Wald, G. Cis-trans isoforms of Vitamin A and retinene in the rhodopsin system. *Journal of General Physiology* **36**, 269-315, (1952).
249 Hao, W., Chen, P. & Fong, H. K. W. in *Methods in Enzymology* Vol. 316 413-422 (Academic Press, 2000).
250 Choi, E. H., Daruwalla, A., Suh, S., Leinonen, H. & Palczewski, K. J. J. o. l. r. Retinoids in the visual cycle: role of the retinal G protein-coupled receptor. **62** (2021).
251 Zhang, J. *et al.* Photic generation of 11-cis-retinal in bovine retinal pigment epithelium. *Journal of Biological Chemistry* **294**, 19137-19154, (2019).
252 Morshedian, A. *et al.* Light-Driven Regeneration of Cone Visual Pigments through a Mechanism Involving RGR Opsin in Müller Glial Cells. *Neuron* **102**, 1172-1183.e1175, (2019).
253 Maeda, A., Golczak, M., Maeda, T. & Palczewski, K. Limited Roles of Rdh8, Rdh12, and Abca4 in all-trans-Retinal Clearance in Mouse Retina. *Investigative Ophthalmology & Visual Science* **50**, 5435-5443, (2009).
254 Sahu, B. & Maeda, A. Retinol Dehydrogenases Regulate Vitamin A Metabolism for Visual Function. *Nutrients* **8** (2016).
255 Groenendijk, G. W. T., Jacobs, C. W. M., Bonting, S. L. & Daemen, F. J. M. Dark Isomerization of Retinals in the Presence of Phosphatidylethanolamine. *European Journal of Biochemistry* **106**, 119-128, (1980).
256 Shichi H Fau - Somers, R. L. & Somers, R. L. Possible involvement of retinylidene phospholipid in photoisomerization of all-trans-retinal to 11-cis-retinal. (1974).
257 Kaylor, J. J. *et al.* Blue light regenerates functional visual pigments in mammals through a retinyl-phospholipid intermediate. *Nature Communications* **8**, 16, (2017).
258 Suzuki T Fau - Callender, R. H. & Callender, R. H. Primary photochemistry and photoisomerization of retinal at 77 degrees K in cattle and squid rhodopsins. (1981).
259 Boesze-Battaglia, K. & Albert, A. D. Phospholipid distribution among bovine rod outer segment plasma membrane and disk membranes. *Experimental Eye Research* **54**, 821-823, (1992).
260 Fliesler, A. J. & Anderson, R. E. Chemistry and metabolism of lipids in the vertebrate retina. *Progress in Lipid Research* **22**, 79-131, (1983).
261 Boesze-Battaglia, K. & Albert, A. D. Fatty acid composition of bovine rod outer segment plasma membrane. *Experimental Eye Research* **49**, 699-701, (1998).
262 Albert, A. D., Young Je Fau - Paw, Z. & Paw, Z. Phospholipid fatty acyl spatial distribution in bovine rod outer segment disk membranes. (1998).
263 Martin, R. E., Elliott, M. H., Brush, R. S. & Anderson, R. E. Detailed Characterization of the Lipid Composition of Detergent-Resistant Membranes from Photoreceptor Rod Outer Segment Membranes. *Investigative Ophthalmology & Visual Science* **46**, 1147-1154, (2005).

264 Sander, C. L. *et al.* Nano-scale resolution of native retinal rod disk membranes reveals differences in lipid composition. *Journal of Cell Biology* **220**, e202101063, (2021).
265 Boesze-Battaglia, K., Hennessey T Fau - Albert, A. D. & Albert, A. D. Cholesterol heterogeneity in bovine rod outer segment disk membranes. (1989).
266 Boesze-Battaglia, K., Fliesler Sj Fau - Albert, A. D. & Albert, A. D. Relationship of cholesterol content to spatial distribution and age of disc membranes in retinal rod outer segments. (1990).
267 Teague, W. E. *et al.* Elastic properties of polyunsaturated phosphatidylethanolamines influence rhodopsin function. *Faraday Discussions* **161**, 383-395, (2013).
268 Alves, I. D. *et al.* Phosphatidylethanolamine Enhances Rhodopsin Photoactivation and Transducin Binding in a Solid Supported Lipid Bilayer as Determined Using Plasmon-Waveguide Resonance Spectroscopy. *Biophysical Journal* **88**, 198-210, (2005).
269 Bennett, M. P. & Mitchell, D. C. Regulation of Membrane Proteins by Dietary Lipids: Effects of Cholesterol and Docosahexaenoic Acid Acyl Chain-Containing Phospholipids on Rhodopsin Stability and Function. *Biophysical Journal* **95**, 1206-1216, (2008).
270 Botelho, A. V., Gibson, N. J., Thurmond, R. L., Wang, Y. & Brown, M. F. Conformational Energetics of Rhodopsin Modulated by Nonlamellar-Forming Lipids. *Biochemistry* **41**, 6354-6368, (2002).
271 Mitchell, D. C., Niu, S.-L. & Litman, B. J. Optimization of Receptor-G Protein Coupling by Bilayer Lipid Composition I: KINETICS OF RHODOPSIN-TRANSDUCIN BINDING *. *Journal of Biological Chemistry* **276**, 42801-42806, (2001).
272 Niu, S.-L., Mitchell, D. C. & Litman, B. J. Optimization of Receptor-G Protein Coupling by Bilayer Lipid Composition II: FORMATION OF METARHODOPSIN II-TRANSDUCIN COMPLEX *. *Journal of Biological Chemistry* **276**, 42807-42811, (2001).
273 Endress, E., Heller, H., Casalta, H., Brown, M. F. & Bayerl, T. M. Anisotropic Motion and Molecular Dynamics of Cholesterol, Lanosterol, and Ergosterol in Lecithin Bilayers Studied by Quasi-elastic Neutron Scattering. *Biochemistry* **41**, 13078-13086, (2002).
274 Pitman, M. C., Grossfield, A., Suits, F. & Feller, S. E. Role of Cholesterol and Polyunsaturated Chains in Lipid−Protein Interactions: Molecular Dynamics Simulation of Rhodopsin in a Realistic Membrane Environment. *Journal of the American Chemical Society* **127**, 4576-4577, (2005).
275 Albert, A. D. & Boesze-Battaglia, K. The role of cholesterol in rod outer segment membranes. *Progress in Lipid Research* **44**, 99-124, (2005).
276 Soubias, O., Niu, S.-L., Mitchell, D. C. & Gawrisch, K. Lipid−Rhodopsin Hydrophobic Mismatch Alters Rhodopsin Helical Content. *Journal of the American Chemical Society* **130**, 12465-12471, (2008).
277 Mitchell, D. C., Straume, M. & Litman, B. J. Role of sn-1-saturated,sn-2-polyunsaturated phospholipids in control of membrane receptor conformational equilibrium: effects of cholesterol and acyl chain unsaturation on the metarhodopsin I .tautm. metarhodopsin II equilibrium. *Biochemistry* **31**, 662-670, (1992).
278 Mitchell, D. C., Niu, S.-L. & Litman, B. J. DHA-Rich phospholipids optimize G-Protein coupled signaling. *The Journal of Pediatrics* **143**, 80-86, (2003).
279 Mitchell, D. C., Lawrence, J. T. R. & Litman, B. J. Primary Alcohols Modulate the Activation of the G Protein-coupled Receptor Rhodopsin by a Lipid-mediated Mechanism. *Journal of Biological Chemistry* **271**, 19033-19036, (1996).

280 Grossfield, A., Feller, S. E. & Pitman, M. C. A role for direct interactions in the modulation of rhodopsin by ω-3 polyunsaturated lipids. *Proceedings of the National Academy of Sciences* **103**, 4888-4893, (2006).
281 Soubias, O. & Gawrisch, K. Docosahexaenoyl Chains Isomerize on the Sub-Nanosecond Time Scale. *Journal of the American Chemical Society* **129**, 6678-6679, (2007).
282 Athanasiou, D. *et al.* The molecular and cellular basis of rhodopsin retinitis pigmentosa reveals potential strategies for therapy. *Progress in Retinal and Eye Research* **62**, 1-23, (2018).
283 Ferretti, L., Karnik, S. S., Khorana, H. G., Nassal, M. & Oprian, D. D. Total synthesis of a gene for bovine rhodopsin. *Proceedings of the National Academy of Sciences* **83**, 599-603, (1986).
284 Franke, R. R., Sakmar, T. P., Graham, R. M. & Khorana, H. G. Structure and function in rhodopsin. Studies of the interaction between the rhodopsin cytoplasmic domain and transducin. *Journal of Biological Chemistry* **267**, 14767-14774, (1992).
285 Hubbell, W. L., Altenbach, C., Hubbell, C. M. & Khorana, H. G. in *Advances in Protein Chemistry* Vol. 63 243-290 (Academic Press, 2003).
286 Oprian, D. D., Molday, R. S., Kaufman, R. J. & Khorana, H. G. Expression of a synthetic bovine rhodopsin gene in monkey kidney cells. *Proceedings of the National Academy of Sciences* **84**, 8874-8878, (1987).
287 Ovchinnikov, Y. A. Rhodopsin and bacteriorhodopsin: structure—function relationships. **148**, 179-191, (1982).
288 Hargrave, P. A. *et al.* The structure of bovine rhodopsin. *Biophysics of structure and mechanism* **9**, 235-244, (1983).
289 Nathans, J. & Hogness, D. S. Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell* **34**, 807-814, (1983).
290 Dixon, R. A. F. *et al.* Cloning of the gene and cDNA for mammalian β-adrenergic receptor and homology with rhodopsin. *Nature* **321**, 75-79, (1986).
291 Fredriksson, R., Lagerström, M. C., Lundin, L.-G. & Schiöth, H. B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Molecular Pharmacology* **63**, 1256, (2003).
292 Ballesteros, J. A. *et al.* Activation of the Beta-Adrenergic Receptor Involves Disruption of an Ionic Lock between the Cytoplasmic Ends of Transmembrane Segments 3 and 6 *. *Journal of Biological Chemistry* **276**, 29171-29177, (2001).
293 Angel, T. E., Chance, M. R. & Palczewski, K. Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proceedings of the National Academy of Sciences* **106**, 8555-8560, (2009).
294 Angel, T. E., Gupta, S., Jastrzebska, B., Palczewski, K. & Chance, M. R. Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. *Proceedings of the National Academy of Sciences* **106**, 14367-14372, (2009).
295 Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature* **477**, 549-555, (2011).
296 Jastrzebska, B. *et al.* Rhodopsin–transducin heteropentamer: Three-dimensional structure and biochemical characterization. *Journal of Structural Biology* **176**, 387-394, (2011).

297 Jastrzebska, B., Ringler, P., Palczewski, K. & Engel, A. The rhodopsin-transducin complex houses two distinct rhodopsin molecules. *Journal of Structural Biology* **182**, 164-172, (2013).
298 Jastrzebska, B., Orban, T., Golczak, M., Engel, A. & Palczewski, K. Asymmetry of the rhodopsin dimer in complex with transducin. *The FASEB Journal* **27**, 1572-1584, (2013).
299 Cottet, M. *et al.* BRET and time-resolved FRET strategy to study GPCR oligomerization: from cell lines toward native tissues. *Front Endocrinol* **3** (2012).
300 Park, P. S. H. & Palczewski, K. Diversifying the repertoire of G protein-coupled receptors through oligomerization. *Proceedings of the National Academy of Sciences* **102**, 8793-8794, (2005).
301 Audet, M. & Bouvier, M. Restructuring G-protein-coupled receptor activation. *Cell* **151**, 14-23, (2012).
302 Farrens, D. L. & Khorana, H. G. Structure and Function in Rhodopsin: MEASUREMENT OF THE RATE OF METARHODOPSIN II DECAY BY FLUORESCENCE SPECTROSCOPY (*). *Journal of Biological Chemistry* **270**, 5073-5076, (1995).
303 Blazynski, C. & Ostroy, S. E. Dual pathways in the photolysis of rhodopsin: Studies using a direct chemical method. *Vision Research* **21**, 833-841, (1981).
304 Blazynski, C. & Ostroy, S. E. Pathways in the hydrolysis of vertebrate rhodopsin. *Vision Research* **24**, 459-470, (1984).
305 Heck, M. *et al.* Signaling States of Rhodopsin: FORMATION OF THE STORAGE FORM, METARHODOPSIN III, FROM ACTIVE METARHODOPSIN II*. *Journal of Biological Chemistry* **278**, 3162-3169, (2003).
306 Jastrzebska, B. *et al.* Functional and Structural Characterization of Rhodopsin Oligomers *Journal of Biological Chemistry* **281**, 11917-11922, (2006).
307 Gulati, S. *et al.* Photocyclic behavior of rhodopsin induced by an atypical isomerization mechanism. *Proceedings of the National Academy of Sciences* **114**, E2608-E2615, (2017).
308 Marty, M. T. *et al.* Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Analytical Chemistry* **87**, 4370-4376, (2015).
309 Hutchins, P. D., Russell, J. D. & Coon, J. J. LipiDex: An Integrated Software Package for High-Confidence Lipid Identification. *Cell Systems* **6**, 621-625.e625, (2018).
310 Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395, (2010).
311 Schey, K. L., Papac, D. I., Knapp, D. R. & Crouch, R. K. Matrix-assisted laser desorption mass spectrometry of rhodopsin and bacteriorhodopsin. *Biophysical Journal* **63**, 1240-1243, (1992).
312 Whitelegge, J. P., Gundersen, C. B. & Faull, K. F. Electrospray-ionization mass spectrometry of intact intrinsic membrane proteins. *Protein Science* **7**, 1423-1430, (1998).
313 Norris, C. E. Native mass spectrometry reveals the simultaneous binding of lipids and zinc to rhodopsin. *Int. J. Mass Spectrom.* **460** (2021).
314 Sakmar, T. P., Franke, R. R. & Khorana, H. G. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proceedings of the National Academy of Sciences* **86**, 8309-8313, (1989).

315 Kiser, P. D. & Palczewski, K. Pathways and disease-causing alterations in visual chromophore production for vertebrate vision. *Journal of Biological Chemistry* **296** (2021).
316 Hilger, D., Masureel, M. & Kobilka, B. K. Structure and dynamics of GPCR signaling complexes. *Nature Structural & Molecular Biology* **25**, 4-12, (2018).
317 Weis, W. I. & Kobilka, B. K. The Molecular Basis of G Protein–Coupled Receptor Activation. *Annual Review of Biochemistry* **87**, 897-919, (2014).
318 Lee, Y. *et al.* Molecular basis of β-arrestin coupling to formoterol-bound β1-adrenoceptor. *Nature* **583**, 862-866, (2020).
319 Zhang, M. *et al.* Cryo-EM structure of an activated GPCR–G protein complex in lipid nanodiscs. *Nature Structural & Molecular Biology* **28**, 258-267, (2021).
320 Van Eps, N. *et al.* Conformational equilibria of light-activated rhodopsin in nanodiscs. *Proceedings of the National Academy of Sciences* **114**, E3268-E3275, (2017).
321 Xu, P. *et al.* Structural insights into the lipid and ligand regulation of serotonin receptors. *Nature* **592**, 469-473, (2021).
322 Baker, B. Y. *et al.* Crystallization of proteins from crude bovine rod outer segments. *Methods Enzymol* **557**, 439-458, (2015).
323 Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1** (2006).
324 Getter, T. *et al.* Stereospecific modulation of dimeric rhodopsin. *The FASEB Journal* **33**, 9526-9539, (2019).
325 Getter, T., Kemp, A., Vinberg, F. & Palczewski, K. Identification of small-molecule allosteric modulators that act as enhancers/disrupters of rhodopsin oligomerization. *Journal of Biological Chemistry* **297** (2021).
326 Wang, X., Plachetzki, D. C. & Cote, R. H. The N termini of the inhibitory γ-subunits of phosphodiesterase-6 (PDE6) from rod and cone photoreceptors differentially regulate transducin-mediated PDE6 activation. *J. Biol. Chem.* **294** (2019).
327 Gulati, S. & Palczewski, K. New focus on regulation of the rod photoreceptor phosphodiesterase. *Curr. Opin. Struct. Biol.* **69** (2021).
328 Soubias, O. & Gawrisch, K. Probing Specific Lipid−Protein Interaction by Saturation Transfer Difference NMR Spectroscopy. *Journal of the American Chemical Society* **127**, 13110-13111, (2005).
329 Apweiler, R., Hermjakob, H. & Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database11Dedicated to Prof. Akira Kobata and Prof. Harry Schachter on the occasion of their 65th birthdays. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1473**, 4-8, (1999).
330 Wohlschlager, T. *et al.* Native mass spectrometry combined with enzymatic dissection unravels glycoform heterogeneity of biopharmaceuticals. *Nature Communications* **9**, 1713, (2018).
331 Struwe, W. B. & Robinson, C. V. Relating glycoprotein structural heterogeneity to function – insights from native mass spectrometry. *Current Opinion in Structural Biology* **58**, 241-248, (2019).
332 Han, L. *et al.* Protein–Glycolipid Interactions Studied in Vitro Using ESI-MS and Nanodiscs: Insights into the Mechanisms and Energetics of Binding. *Analytical Chemistry* **87**, 4888-4896, (2015).

333 Han, L. *et al.* Neoglycolipids as Glycosphingolipid Surrogates for Protein Binding Studies Using Nanodiscs and Native Mass Spectrometry. *Analytical Chemistry* **92**, 14189-14196, (2020).
334 Boeri Erba, E., Signor, L. & Petosa, C. Exploring the structure and dynamics of macromolecular complexes by native mass spectrometry. *Journal of Proteomics* **222**, 103799, (2020).
335 Leney, A. C., Rafie, K., van Aalten, D. M. F. & Heck, A. J. R. Direct Monitoring of Protein O-GlcNAcylation by High-Resolution Native Mass Spectrometry. *ACS Chemical Biology* **12**, 2078-2084, (2017).
336 Dixit, S. M., Polasky, D. A. & Ruotolo, B. T. Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Current Opinion in Chemical Biology* **42**, 93-100, (2018).
337 Yang, Y. *et al.* Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nature Communications* **7**, 13397, (2016).
338 Čaval, T., Tian, W., Yang, Z., Clausen, H. & Heck, A. J. R. Direct quality control of glycoengineered erythropoietin variants. *Nature Communications* **9**, 3342, (2018).
339 Franc, V., Yang, Y. & Heck, A. J. R. Proteoform Profile Mapping of the Human Serum Complement Component C9 Revealing Unexpected New Features of N-, O-, and C-Glycosylation. *Analytical Chemistry* **89**, 3483-3491, (2017).
340 Rudd, P. M. *et al.* Oligosaccharide sequencing technology. *Nature* **388**, 205-207, (1997).
341 Struwe, W. B. & Rudd, P. M. Aminoquinolines as fluorescent labels for hydrophilic interaction liquid chromatography of oligosaccharides. *Biological Chemistry* **393**, 757-765, (2012).
342 Grünewald, S., Matthijs, G. & Jaeken, J. Congenital Disorders of Glycosylation: A Review. *Pediatric Research* **52**, 618-624, (2002).
343 Biffi, S. *et al.* Carbohydrate-deficient transferrin (CDT) as a biochemical tool for the screening of congenital disorders of glycosylation (CDGs). *Clinical Biochemistry* **40**, 1431-1434, (2007).
344 Jaeken, J. *et al.* Sialic acid-deficient serum and cerebrospinal fluid transferrin in a newly recognized genetic syndrome. *Clinica Chimica Acta* **144**, 245-247, (1984).
345 Faid, V., Chirat, F., Seta, N., Foulquier, F. & Morelle, W. A rapid mass spectrometric strategy for the characterization of N- and O-glycan chains in the diagnosis of defects in glycan biosynthesis. *PROTEOMICS* **7**, 1800-1813, (2007).
346 Shakin-Eshleman, S. H., Spitalnik, S. L. & Kasturi, L. The Amino Acid at the X Position of an Asn-X-Ser Sequon Is an Important Determinant of N-Linked Core-glycosylation Efficiency (∗). *Journal of Biological Chemistry* **271**, 6363-6366, (1996).
347 Mellquist, J. L., Kasturi, L., Spitalnik, S. L. & Shakin-Eshleman, S. H. The Amino Acid Following an Asn-X-Ser/Thr Sequon Is an Important Determinant of N-Linked Core Glycosylation Efficiency. *Biochemistry* **37**, 6833-6837, (1998).
348 Hang, H. C. & Bertozzi, C. R. The chemistry and biology of mucin-type O-linked glycosylation. *Bioorganic & Medicinal Chemistry* **13**, 5021-5034, (2005).
349 Nwosu, C. C. *et al.* Simultaneous and Extensive Site-specific N- and O-Glycosylation Analysis in Protein Mixtures. *Journal of Proteome Research* **10**, 2612-2624, (2011).

350 Ko, B. J. & Brodbelt, J. S. Comparison of glycopeptide fragmentation by collision induced dissociation and ultraviolet photodissociation. *International Journal of Mass Spectrometry* **377**, 385-392, (2015).
351 Windwarder, M. & Altmann, F. Site-specific analysis of the O-glycosylation of bovine fetuin by electron-transfer dissociation mass spectrometry. *Journal of Proteomics* **108**, 258-268, (2014).
352 Lin, Y.-H., Zhu, J., Meijer, S., Franc, V. & Heck, A. J. R. Glycoproteogenomics: A Frequent Gene Polymorphism Affects the Glycosylation Pattern of the Human Serum Fetuin/α-2-HS-Glycoprotein [S]. *Molecular & Cellular Proteomics* **18**, 1479-1490, (2019).
353 Bunn, H. F. Erythropoietin. *Cold Spring Harbor Perspectives in Medicine* **3** (2013).
354 Erbayraktar, S. *et al.* Asialoerythropoietin is a nonerythropoietic cytokine with broad neuroprotective activity <em>in vivo</em&gt. *Proceedings of the National Academy of Sciences* **100**, 6741, (2003).
355 Walsh, G. & Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nature Biotechnology* **24**, 1241-1252, (2006).
356 Fukuda, M. N., Sasaki, H., Lopez, L. & Fukuda, M. Survival of Recombinant Erythropoietin in the Circulation: The Role of Carbohydrates. *Blood* **73**, 84-89, (1989).
357 Solá, R. J. & Griebenow, K. Glycosylation of Therapeutic Proteins. *BioDrugs* **24**, 9-21, (2010).
358 Falck, D. *et al.* Affinity purification of erythropoietin from cell culture supernatant combined with MALDI-TOF-MS analysis of erythropoietin N-glycosylation. *Scientific Reports* **7**, 5324, (2017).
359 Cowper, B. *et al.* Comprehensive glycan analysis of twelve recombinant human erythropoietin preparations from manufacturers in China and Japan. *Journal of Pharmaceutical and Biomedical Analysis* **153**, 214-220, (2018).
360 Bui, D. T. *et al.* Mass Spectrometry-Based Shotgun Glycomics Using Labeled Glycan Libraries. *Analytical Chemistry* **94**, 4997-5005, (2022).
361 Sagi, D., Peter-Katalinic, J., Conradt, H. S. & Nimtz, M. Sequencing of tri- and tetraantennary N-glycans containing sialic acid by negative mode ESI QTOF tandem MS. *Journal of the American Society for Mass Spectrometry* **13**, 1138-1148, (2002).
362 Treuheit, M. J., Costello, C. E. & Halsall, H. B. Analysis of the five glycosylation sites of human α1-acid glycoprotein. *Biochemical Journal* **283**, 105-112, (1992).
363 Dias, A. M. *et al.* Metabolic control of T cell immune response through glycans in inflammatory bowel disease. *Proceedings of the National Academy of Sciences* **115**, E4651, (2018).
364 Dias, A. M. *et al.* Dysregulation of T cell receptor N-glycosylation: a molecular mechanism involved in ulcerative colitis. *Human Molecular Genetics* **23**, 2416-2427, (2014).
365 Zhang, M. *et al.* Altered β1,6-GlcNAc and bisecting GlcNAc-branched N-glycan on integrin β1 are associated with early spontaneous miscarriage in humans. *Human Reproduction* **30**, 2064-2075, (2015).
366 Hochepied, T., Berger, F. G., Baumann, H. & Libert, C. α1-Acid glycoprotein: an acute phase protein with inflammatory and immunomodulating properties. *Cytokine & Growth Factor Reviews* **14**, 25-34, (2003).

367 Van Dijk, W. & Poland, D. C. W. in *Glycobiology and Medicine.* (ed John S. Axford) 251-256 (Springer US).
368 Williams, J. P. *et al.* α1-Acid glycoprotein reduces local and remote injuries after intestinal ischemia in the rat. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **273**, G1031-G1035, (1997).
369 Imberty, A. & Pérez, S. Structure, Conformation, and Dynamics of Bioactive Oligosaccharides: Theoretical Approaches and Experimental Validations. *Chemical Reviews* **100**, 4567-4588, (2000).
370 Raph, S. M., Bhatnagar, A. & Nystoriak, M. A. Biochemical and physiological properties of K+ channel-associated AKR6A (Kvβ) proteins. *Chemico-Biological Interactions* **305**, 21-27, (2019).
371 Jan, L. Y. & Jan, Y. N. Voltage-gated potassium channels and the diversity of electrical signalling. *The Journal of Physiology* **590**, 2591-2599, (2012).
372 Bushey, D., Huber, R., Tononi, G. & Cirelli, C. Drosophila Hyperkinetic Mutants Have Reduced Sleep and Impaired Memory. *The Journal of Neuroscience* **27**, 5384, (2007).
373 Cirelli, C. *et al.* Reduced sleep in Drosophila Shaker mutants. *Nature* **434**, 1087-1092, (2005).
374 McCormack, K. *et al.* Genetic Analysis of the Mammalian K+ Channel β Subunit Kvβ2 (Kcnab2)∗. *Journal of Biological Chemistry* **277**, 13219-13228, (2002).
375 Douglas, C. L. *et al.* Sleep in Kcna2 knockout mice. *BMC Biology* **5**, 42, (2007).
376 Donlea, J. M. *et al.* Recurrent Circuitry for Balancing Sleep Need and Sleep. *Neuron* **97**, 378-389.e374, (2018).
377 Ni, J. D. *et al.* Differential regulation of the Drosophila sleep homeostat by circadian and arousal inputs. *eLife* **8**, e40487, (2019).
378 Kempf, A., Song, S. M., Talbot, C. B. & Miesenböck, G. A potassium channel β-subunit couples mitochondrial electron transport to sleep. *Nature* **568**, 230-234, (2019).
379 Xu, P., Cox, K. H. & Takahashi, J. S. A Hyperkinetic Redox Sensor Drives Flies to Sleep. *Trends in Neurosciences* **42**, 514-517, (2019).
380 Chouinard, S. W., Wilson, G. F., Schlimgen, A. K. & Ganetzky, B. A potassium channel β subunit related to the aldo-keto reductase superfamily is encoded by the Drosophila hyperkinetic locus. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 6763-6767, (1995).
381 Long Stephen, B., Campbell Ernest, B. & MacKinnon, R. Voltage Sensor of Kv1.2: Structural Basis of Electromechanical Coupling. *Science* **309**, 903-908, (2005).
382 Doyle, D. A. *et al.* The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **280**, 69-77, (1998).
383 Gulbis, J. M., Mann, S. & MacKinnon, R. Structure of a Voltage-Dependent K+ Channel β Subunit. *Cell* **97**, 943-952, (1999).
384 Gulbis, J. M., Zhou, M., Mann, S. & MacKinnon, R. Structure of the Cytoplasmic β Subunit--T1 Assembly of Voltage-Dependent K+ Channels. *Science* **289**, 123-127, (2000).
385 Edwards, G. & Weston, A. H. The role of potassium channels in excitable cells. *Diabetes Research and Clinical Practice* **28**, S57-S66, (1995).

386 Rhodes, K. J. *et al.* Association and Colocalization of the Kvβ1 and Kvβ2 β-Subunits with Kv1 α-Subunits in Mammalian Brain K⁺Channel Complexes. *The Journal of Neuroscience* **17**, 8246, (1997).
387 Shi, G. *et al.* βSubunits Promote K+ Channel Surface Expression through Effects Early in Biosynthesis. *Neuron* **16**, 843-852, (1996).
388 Nagaya, N. & Papazian, D. M. Potassium Channel ä and β Subunits Assemble in the Endoplasmic Reticulum. *Journal of Biological Chemistry* **272**, 3022-3027, (1997).
389 Uebele, V. N., England, S. K., Chaudhary, A., Tamkun, M. M. & Snyders, D. J. Functional Differences in Kv1.5 Currents Expressed in Mammalian Cell Lines Are Due to the Presence of Endogenous Kvβ2.1 Subunits (*). *Journal of Biological Chemistry* **271**, 2406-2412, (1996).
390 Rehm, H. & Lazdunski, M. Purification and subunit structure of a putative K+-channel protein identified by its binding properties for dendrotoxin I. *Proceedings of the National Academy of Sciences* **85**, 4919-4923, (1988).
391 Parcej, D. N., Scott, V. E. S. & Dolly, J. O. Oligomeric properties of .alpha.-dendrotoxin-sensitive potassium ion channels purified from bovine brain. *Biochemistry* **31**, 11084-11088, (1992).
392 Scott, V. E. *et al.* Primary structure of a beta subunit of alpha-dendrotoxin-sensitive K+ channels from bovine brain. *Proceedings of the National Academy of Sciences* **91**, 1637-1641, (1994).
393 Weng, J., Cao, Y., Moss, N. & Zhou, M. Modulation of Voltage-dependent Shaker Family Potassium Channels by an Aldo-Keto Reductase. *Journal of Biological Chemistry* **281**, 15194-15200, (2006).
394 Perkowski, J. J. & Murphy, G. G. Deletion of the Mouse Homolog of KCNAB2, a Gene Linked to Monosomy 1p36, Results in Associative Memory Impairments and Amygdala Hyperexcitability. *The Journal of Neuroscience* **31**, 46, (2011).
395 Hoshi, T., Zagotta William, N. & Aldrich Richard, W. Biophysical and Molecular Mechanisms of Shaker Potassium Channel Inactivation. *Science* **250**, 533-538, (1990).
396 Xu, J. & Li, M. Kvβ2 Inhibits the Kvβ1-mediated Inactivation of K+ Channels in Transfected Mammalian Cells*. *Journal of Biological Chemistry* **272**, 11728-11735, (1997).
397 Rettig, J. *et al.* Inactivation properties of voltage-gated K+ channels altered by presence of β-subunit. *Nature* **369**, 289-294, (1994).
398 Hoog, S. S., Pawlowski, J. E., Alzari, P. M., Penning, T. M. & Lewis, M. Three-dimensional structure of rat liver 3 alpha-hydroxysteroid/dihydrodiol dehydrogenase: a member of the aldo-keto reductase superfamily. *Proceedings of the National Academy of Sciences* **91**, 2517-2521, (1994).
399 Barski, O. A., Tipparaju, S. M. & Bhatnagar, A. The Aldo-Keto Reductase Superfamily and its Role in Drug Metabolism and Detoxification. *Drug Metabolism Reviews* **40**, 553-624, (2008).
400 Lee, S. H. & Blair, I. A. Characterization of 4-Oxo-2-nonenal as a Novel Product of Lipid Peroxidation. *Chemical Research in Toxicology* **13**, 698-702, (2000).

401 Tipparaju, S. M., Barski, O. A., Srivastava, S. & Bhatnagar, A. Catalytic Mechanism and Substrate Specificity of the β-Subunit of the Voltage-Gated Potassium Channel. *Biochemistry* **47**, 8840-8854, (2008).
402 McCormack, T. & McCormack, K. Shaker K+ channel β subunits belong to an NAD(P)H-dependent oxidoreductase superfamily. *Cell* **79**, 1133-1135, (1994).
403 Bringmann, H. Sleep-Active Neurons: Conserved Motors of Sleep. *Genetics* **208**, 1279-1289, (2018).
404 Murphy, Michael P. How mitochondria produce reactive oxygen species. *Biochemical Journal* **417**, 1-13, (2008).
405 Wong, Y. C., Peng, W. & Krainc, D. Lysosomal Regulation of Inter-mitochondrial Contact Fate and Motility in Charcot-Marie-Tooth Type 2. *Developmental Cell* **50**, 339-354.e334, (2019).
406 Balaban, R. S., Nemoto, S. & Finkel, T. Mitochondria, Oxidants, and Aging. *Cell* **120**, 483-495, (2005).
407 Parvez, S., Long, M. J. C., Poganik, J. R. & Aye, Y. Redox Signaling by Reactive Electrophiles and Oxidants. *Chemical Reviews* **118**, 8798-8888, (2018).
408 Sies, H. & Jones, D. P. Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nature Reviews Molecular Cell Biology* **21**, 363-383, (2020).
409 Higdon, A., Diers, Anne R., Oh, Joo Y., Landar, A. & Darley-Usmar, Victor M. Cell signalling by reactive lipid species: new concepts and molecular mechanisms. *Biochemical Journal* **442**, 453-464, (2012).
410 Czapski, G. A., Czubowicz, K., Strosznajder, J. B. & Strosznajder, R. P. The Lipoxygenases: Their Regulation and Implication in Alzheimer's Disease. *Neurochemical Research* **41**, 243-257, (2016).
411 Hofstadler, S. A. & Sannes-Lowery, K. A. Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. *Nature Reviews Drug Discovery* **5**, 585-595, (2006).
412 Tipparaju, S. M., Liu, S.-Q., Barski, O. A. & Bhatnagar, A. NADPH binding to β-subunit regulates inactivation of voltage-gated K+ channels. *Biochemical and Biophysical Research Communications* **359**, 269-276, (2007).
413 Vacher, H. *et al.* Cdk-mediated phosphorylation of the Kvβ2 auxiliary subunit regulates Kv1 channel axonal targeting. *Journal of Cell Biology* **192**, 813-824, (2011).
414 Alka, K., Dolly, J. O., Ryan, B. J. & Henehan, G. T. M. New inhibitors of the Kvβ2 subunit from mammalian Kv1 potassium channels. *The International Journal of Biochemistry & Cell Biology* **55**, 35-39, (2014).
415 Gauglitz, J. M. *et al.* Enhancing untargeted metabolomics using metadata-based source annotation. *Nature Biotechnology* (2022).
416 Esser, T. K. *et al.* Mass-selective and ice-free electron cryomicroscopy protein sample preparation via native electrospray ion-beam deposition. *PNAS Nexus* **1**, pgac153, (2022).
417 Esser, T. K. *et al.* Cryo-EM samples of gas-phase purified protein assemblies using native electrospray ion-beam deposition. *Faraday Discussions* (2022).
418 Schopper, S. *et al.* Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nature Protocols* **12**, 2391-2410, (2017).

419 Reading, E. *et al.* Interrogating Membrane Protein Conformational Dynamics within Native Lipid Compositions. *Angewandte Chemie International Edition* **56**, 15654-15657, (2017).
420 Calabrese, A. N., Ault, J. R., Radford, S. E. & Ashcroft, A. E. Using hydroxyl radical footprinting to explore the free energy landscape of protein folding. *Methods* **89**, 38-44, (2015).
421 McKenzie-Coe, A., Montes, N. S. & Jones, L. M. Hydroxyl Radical Protein Footprinting: A Mass Spectrometry-Based Structural Method for Studying the Higher Order Structure of Proteins. *Chemical Reviews* **122**, 7532-7561, (2022).
422 Lu, Y. *et al.* Fast Photochemical Oxidation of Proteins Maps the Topology of Intrinsic Membrane Proteins: Light-Harvesting Complex 2 in a Nanodisc. *Analytical Chemistry* **88**, 8827-8834, (2016).
423 Kempa, E. E. *et al.* Coupling Droplet Microfluidics with Mass Spectrometry for Ultrahigh-Throughput Analysis of Complex Mixtures up to and above 30 Hz. *Analytical Chemistry* **92**, 12605-12612, (2020).