Understanding novel reconstitution systems for native membrane protein biophysical characterisation

Peter Andrew Fox, BSc

Submitted in accordance with the requirement for the degree of

Master of Science by Research

University of Leeds

Faculty of Biological Sciences

School of Molecular and Cellular Biology

April 2022

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement

The right of Peter Andrew Fox to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

I. Acknowledgments

This thesis is the result of much hard work and collaboration of a number of people from different institutions and differing specialisms. I would like to start by thanking Dr Carine De Marcos Lousa (Leeds Beckett University) and Dr Vincent Postis (University of Dundee) for helping develop the project. Their continued support from my time at Leeds Beckett university onwards through my degree at university of Leeds leaves me forever indebted to them.

As part of Leeds Beckett university was also Mahmoud Eldahshoury who provided support, insight and expertise during this project and for that I would like to show my appreciation.

I would like to display my gratitude to University of Leeds, The Astbury Centre for Structural Molecular Biology for hosting this postgraduate project. Furthermore, extending this gratitude to Professor Frank Sobott for supervising me, providing support, constructive feedback throughout and opportunities. This facilitated my development as a researcher.

Sincere thanks towards all others associated with this project particularly Dr Anna Higgins allowing me to work with her, which was a delight and taught me much that will follow me through life as a research scientist.

In addition, many thanks go to collaborators who provided materials and insight that facilitated this project Such as Dr Stephen Muench, Dr Leonhard Urner (university of oxford) and Dr Manuela Zoonens (University of Paris).

Finally, my gratitude towards my partner Grace Armstrong, BA for sharing this journey and for her continued support throughout my studies.

II. Abstract

Membrane Proteins (MPs) constitute the majority of developed drug targets, and our ability to satisfactorily investigate membrane proteins is paramount to a variety of areas within science. This focus has put a strong emphasis on understanding how best to imitate the local environment of membrane proteins to maintain a native-like conformation of target MPs for biophysical analysis. In particular, mass spectrometry has started to be used to characterise membrane protein-lipid interactions. Most commonly used detergent based MP reconstitution methods have been shown to remove some of these crucial lipids, altering MP characteristics. Because of this, reconstitution systems have been developed to improve flexibility in membrane protein research. However, a lack of understanding as to how well each of them replicates the native environment of membrane proteins hinders progress. Here, novel reconstitution systems (C_6-C_2-50, C_8-C_0-50 and G1 OGD modular detergent) are compared against current systems (SMA, DDM and A8-35) to provide understanding on how well each replicates the native membrane environment and how well they apply to different membrane environments. A model membrane protein, Bacteriorhodopsin (HbR), was expressed in its native host (*H. salinarium*) and heterologously expressed in *E. coli* for comparison. After membrane isolation, HbR and *E. coli* bacteriorhodopsin (EbR) was purified using a variety of detergents and nanoparticles. These preparations were analysed using native mass spectrometry and mass photometry to identify retained lipids and oligomerisation states. Our tested reconstitution systems demonstrated a mixture of advantages and disadvantages described in this thesis, that will be valuable in structural biology. Exploiting these findings will help further understand current limitations in membrane proteins structural investigations.

III. Table of Tables

IV. Table of Figures

Chapter 1: Introduction

1.1 Background: Membrane Protein (MP) investigations

Despite the genome coding for 20-30% of MPs, there is still little information on both structural and functional characteristics of many MPs (Hoi et al., 2021; Olerinyova et al., 2021). Furthermore, their vital function of maintaining cellular homeostasis and consequently creating physiological homeostasis presents our ability to investigate MPs as vitally important in scientific research of disease. More so, given that MPs comprise a large proportion of drug targets (Marconnet et al., 2020). Although this is the case, it has become clear that the discrepancy in MP knowledge is mainly due to difficulties investigating them.

1.2 Membrane protein purification

The insolubility of MPs meant biophysical characterisation is difficult without techniques for MP solubilisation. It also quickly became clear that outside of the native cellular environment, MPs are unstable and require a support structure to maintain their structural characteristics. The development of detergent-based isolation techniques tackled both these problems and quickly became a preferred method of MP isolation. Detergent based isolation techniques have therefore allowed for the application of modern characterisation methods, particularly facilitating the use of many characterisation methods such as Nuclear magnetic resonance (NMR), Cryogenic electron microscopy (CryoEM), mass spectrometry (MS), X-ray crystallography among others.

Detergent based isolation were the original method for investigating membrane proteins. Detergent stabilisation of MPs were originally thought to replicate the lipidic environment of the membrane due to the structural similarity to lipids. Upon consistent use of detergents for many MPs it became clear that there were functional and structural alterations when using detergents to stabilise membrane proteins. MPs from here on have been increasingly shown to rely on the native environment for both structural stability and functional characteristics with examples such as lipid dependant gating of potassium channels (Zheng et al., 2011). This poses a problem when removing them from the native cellular environment for biophysical characterisation using these popular detergents such as n-dodecyl-β-D-maltopyranoside (DDM). This turned the focus of structural biology on retaining bulk lipids to allow for a more native-like environment.

1.3 'Nativeomics'

The new focus on retention of endogenous lipids of membrane proteins has gained popularity and has given rise to a new 'omics'. Described as a mixture of lipidomics, proteomics and metabolomics; it focuses on retaining endogenous ligands and lipids during characterisation of membrane proteins. To this aim, novel techniques have been developed such as sonicated native proteoliposomes for probing native MP confirmation (Chorev et al., 2018). Termed SoLVe-MS or sonicated lipovesicle mass spectrometry; native membranes (outer and inner membranes) containing proteins are sonicated before using nMS for characterisation (Chorev et al., 2020). This is an exciting development being able to investigate MPs directly from native membranes and has shown some promise being applied to membrane protein enriched extracellular vesicles from mammalian cells. It appears that there is limited use of this with bacterial based expression. Chorev et al., (2018) demonstrated the ability to attribute MS spectrum peaks to proteins known to be present in the native membrane of *E. coli* using SoLVe-MS. However, this study sparked controversy due to the lack of specificity to identified proteins. This was because the proteins weren't purified using a protein specific technique such as Immobilisation metal affinity chromatography during purification and were identified solely on their mass.

Another technique that has gained much attention for this application is the use of synthetic lipid liposomes. The reconstitution of the target membrane protein into liposomes, such as DMPC (for the target protein of this investigation) or POPC. This allowed the incorporation of the native lipids within the final proteosome. Furthermore, the target membrane protein can then be solubilised directly from these liposomes using styrene maleic acid (SMA)(Hoi et al., 2021). The detergent-free nature and the ability to identify lipid adducts within nMS spectra has made this an attractive technique.

These techniques have proven to be significant developments within the nativeomics field and highlight the need to understand how best to retain native fold and the surrounding molecules of membrane proteins. As of currently, there are a variety of novel molecules that can be used to provide native-like environment for solubilisation/isolation of membrane proteins. Despite this, our understanding on how these novel reconstitution molecules affect membrane protein dynamics is minimal.

1.4 Membrane protein reconstitution systems

As mentioned above there are a range of current and novel reconstitution methods that aim to retain endogenous lipids of target membrane proteins. Given the very recent

synthesis and application of some of these reconstitution methods it is not clear how they compare to each other or how well they would apply to biophysical characterisation methods. Before underpinning the desired outcome of this investigation, each membrane protein stabilisation method should be introduced to govern their theoretical application within structural biology.

1.4.1 Modular detergents (MD)

As already mentioned, detergents can be detrimental to membrane protein structure and function due to the stripping of endogenous ligands. Despite this, they are still used as a common form of membrane protein isolation. This is primarily due to the large yields associated with detergent isolation. Recently, synthesised modular detergents have been used to isolate membrane proteins demonstrating a higher yield to commonly used DDM detergent (Urner et al., 2020).

It was thought the derivation of these molecules from cellular detergents lends themselves to providing a flexible environment for membrane proteins. These MDs or oligoglycerol detergents (OGD) are a subclass of detergents called dendrons. It has been demonstrated that maintaining a hydroxyl linker group and altering the head and tail groups (Figure 1) of OGDs provides flexibility in using these for a variety applications in membrane protein biology (Urner et al., 2021). In particular, the G1 OGD have proven to resolve lipid adducts in nMS, suggesting the preservation of membrane protein-lipid interactions. This flexibility could permit a standard solubilisation system for biophysical characterisation of a large variety of membrane proteins in their native folds.

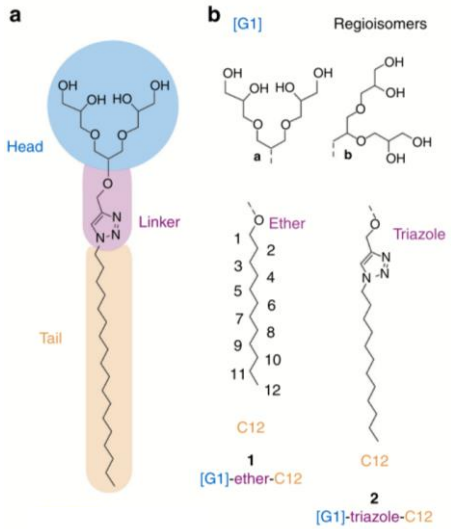

*Figure 1 - **Modular detergent structure simplified** from Urner et al (2020)*

1.4.2 SMALPs (styrene maleic acid lipoparticles)

Styrene maleic acid lipoparticles were a considerable development within biochemistry. A study in 2009 originally demonstrated the ability of SMA (Figure 2) to reconstitute membrane proteins (Lee, S. C. et al., 2016). This was an introduction to detergent free reconstitution of membrane proteins. Moreover, the modality of SMA appears to be direct interaction with surrounding lipids to provide a 'cookie cutter' effect by removing the target MP with its annular lipids (Stroud et al., 2018). MP activity investigations of SMA isolated MPs started to demonstrate the importance of the native membrane environment in MP characterisation studies (Hesketh et al., 2020; Lee, S. C. et al., 2016; Stroud et al., 2018).

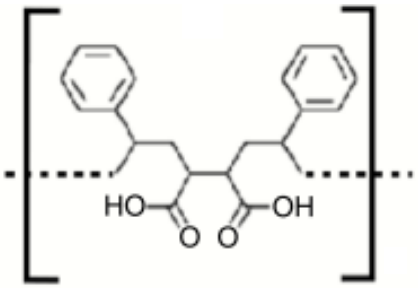

*Figure 2- **SMA structure** from Lee et al (2016)*

SMA is still considered to be the gold standard for isolation of native MPs. The use of SMALPs has gained traction and has become more ubiquitous for native membrane protein analysis. Despite this, SMA does not apply well to certain biophysical characterisation methods such as nMS and has only recently been applied to cryoEM of AcrB (Parmar et al., 2018)and YnaI mechanosensitive channel (Catalano et al., 2021) among others.

1.4.3 Amphipols

Amphipols are a class of polymers that have allowed isolation of a variety of membrane proteins. Despite the variety of amphipols, A8-35 appears to be the most commonly used for biophysical characterisation. A8-35 has been used to solubilise a variety of membrane proteins such bacteriorhodopsin and OmpF (Tribet et al., 1996) and many more discussed in depth by Zoonens & Popot (2014). A8-35 was seen as a good alternative to detergent-based isolation with its ability to retain native confirmations. The development of A8-35 was particularly useful for CryoEM eliminating background noise (Ratkeviciute et al., 2021). It has also been applied to mass spectrometry showing an improvement over the use of detergent-based reconstitution (Calabrese et al., 2015) both

these findings making it an attractive approach for biophysical characterisation of MPs as it does not limit methodology.

Membrane protein isolation using A8-35 previously required use of detergent to reconstitute MPs (Gohon et al., 2008; Ratkeviciute et al., 2021; Tribet et al., 1996). In the search for a more flexible amphipol, they were modified to include cyclic hydrocarbon groups in place of linear groups (Figure 3) giving rise to two novel reconstitution systems (C_6-C_2-50 and C_8-C_0-50). These alterations were similar in structure to the aromatic side chains of SMA. It was hoped that this similarity with SMA would provide the cyclic amphipols with SMA-like mode of action; retaining native lipids and providing a native like environment. All the while being more amenable to biophysical characterisation methods such as mass spectrometry and CryoEM (Marconnet et al., 2020).

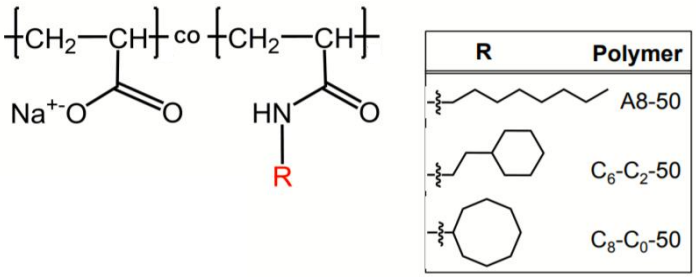

*Figure 3 - **Cyclic amphipols structure** modified from Marconnet et al (2020). Showing structure of cyclic groups of both novel cyclic amphipols*

Cyclic amphipols are a novel endeavour and as of currently have shown promise their application to membrane protein analysis. Both C_6-C_2-50 and C_8-C_0-50 have recently been used to solubilise membrane proteins directly from membranes for biophysical analysis using CryoEM, providing a high resolution (3.2Å) structure of AcrB (Higgins et al., 2021). Additionally, lipids were extracted from novel amphipol purified MPs by precipitating the proteins and amphipols. The lipids confirmed by separating them via thin layer, a liquid phase lipid separation technique. The novel amphipol isolated protein samples therefore demonstrated to have co-purified lipids. This shows promise of the application of cyclic amphipols to characterisation of native membrane proteins with their native lipids.

1.5 Biophysical characterisation of complex biomolecules

1.5.1 Introduction

The understanding of MP structure and function has led to further scientific understanding in a variety of areas allowing for molecular target identification and much more. This vast array of applications places proteomics at the forefront of many of the current research areas. The importance of understanding protein dynamics meant that

a large variety of techniques have been developed to analyse membrane proteins. Current techniques that have been subject to much investigation as of present are X-ray crystallography, CryoEM, mass spectrometry and NMR. These techniques allow protein scientists to uncover both structural and functional data of target MPs. It is also apparent that these techniques come with a variety of advantage and disadvantages, especially in the light of the topic of interest. Xray crystallography is still a trustworthy technique for structural investigation this does however require extensive purification, high concentrations and crystallization typically using detergents. This workflow has been improved with the introduction of CryoEM, with smaller volumes and lower concentrations being required for acquisition of high-resolution protein structures. Both these techniques have been used to investigate protein-lipid and protein-protein interactions. However, both these techniques demonstrate difficulties in determining lipid specific protein-lipid interactions. This is primarily to do with the workflow and purification steps associated with each technique. Furthermore, CryoEM struggles to acquire structures for smaller molecules around 100kDa. Because of these disadvantages of using individual techniques, it increasingly becomes a target of scientists to use a multidisciplinary approach to investigate these target proteins. Structural biology is an example of an area that benefits from an interdisciplinary approach using a variety of complementary techniques. Not only does this provide reproducibility but also provides detailed conclusions. Because of this, it is common to use two or more of these techniques together to provide a detailed understanding of the target MP. As mentioned previously certain reconstitution methods are not amenable to all biophysical characterisation methods. Therefore, it is intended within this project to identify reconstitution systems that could contend with current systems or develop strategies to use current solubilisation molecules to produce the desired data. This investigation will benefit from the use of mass spectrometry and mass photometry for a multitude of reasons detailed below.

1.5.2 Mass spectrometry

Mass spectrometry (MS), illustrated in Figure 4, applies technology to transfer a molecule into the gas-phase which can be achieved using a variety of techniques. Uptake of the charged biomolecule by the mass spectrometer is facilitated by a potential difference, the potential difference is facilitated by the ionisation of sample (+ve) and the charge of the sample intake (-ve). Upon intake of the sample ion, it is accelerated and manipulated towards a mass analyser under a high vacuum. Many current mass spectrometers also include some sort of collision cell for MS/MS applications creating fragment ions for

additional characterisation capabilities. Many current mass spectrometers are also equipped with a mass analysers before fragmentation for precursor ion identification. The post fragmentation mass analyser then allows identification of subcomplexes or fragments that have been dissociated from the precursor ions. This is known as true MS/MS or MS^2. After detection by a mass analyser, this is then interpreted as a mass/charge ratio (m/z). Further data, on abundance and charge state is represented as m/z on a spectrum.

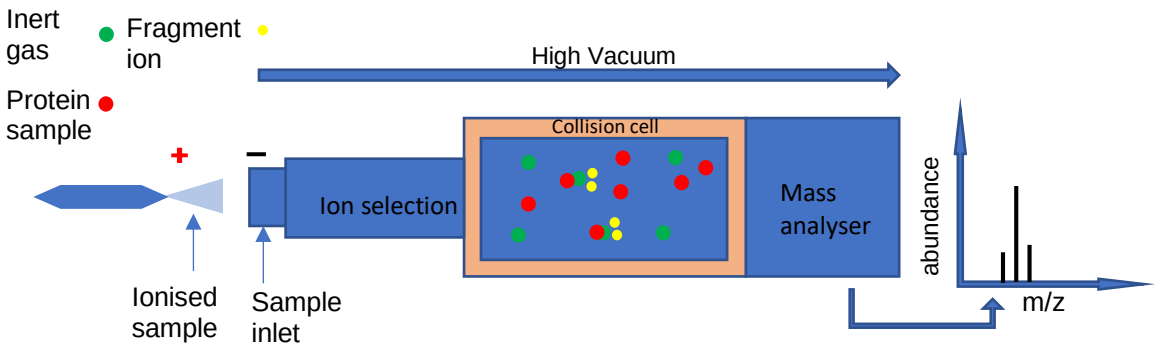

Figure 4 - ***basic principles of Mass spectrometry*** *showing typical mass spectrometer with ion selection and collision cell for the application of collision induced dissociation coupled with a mass analyser*

The recent focus of structural biology to elucidate biological structures in a native-like conformation has highlighted the need for methodologies that can investigate whole protein complexes and molecule-specific interactions. In particular, mass spectrometry has attracted a lot of attention due to its sensitivity, minimal sample usage and ability to retain native structure while providing detailed information on the protein fold. Mass spectrometry is a field in its own, with a variety of mass spectrometers that can be applied to a multitude of different investigations such as top-down (protein complexes) and bottom-up (peptide focus) investigations. Providing that the remit of this investigation is to understand how well each purification method retains native structure, it will focus on the use of native mass spectrometry for ligand identification and conformation of the target protein.

1.5.3 Native Mass Spectrometry (nMS)

The defining concept of native mass spectrometry is the retention of native fold of the target molecule complete with non-covalent interactions. To achieve the retention of native conformations, a volatile buffer is used to replace unwanted non-volatile molecules while also providing a preferable environment to retain native-like

confirmation. This requires a buffer exchange process which is carried out extensively to remove salt and other undesirable molecules. Ammonium acetate is usually the buffer of choice for these investigations, holding a desirable pH with the already mentioned volatility to facilitate gas-phase analysis.

The key requirement before analysis is that the target molecule has to be charged to be analysed. There are a variety of techniques that can be used to apply a charge to an uncharged sample. Certain ionisation methods provide too harsh conditions to provide intact macromolecular complexes during acquisition. This means native mass spectrometry relies heavily on the use of soft ionisation techniques such as Matrix assisted laser desorption ionisation (MALDI) and in particular Electrospray ionisation (Pimlott & Konermann, 2021; Zubarev & Makarov, 2013). The extensive use of Electrospray ionisation makes this the most relevant technique currently, because of this investigation intends to use this as the ionisation method of choice.

1.5.4 Electrospray Ionisation

Electrospray ionisation (ESI) was developed for mass spectrometry of intact proteins (Rohner et al., 2004), and has emerged as the primary ionisation technique used for native mass spectrometry. This technique sprays membrane proteins within a solvent via an electrostatically charged needle towards the mass spectrometer inlet.

Currently there are three proposed models for how this technique produces charged ions; Chain ejection model (CEM), Ion evaporation model (IEM) and Charge residue model (CRM)(Pimlott & Konermann, 2021; Raab et al., 2021). Despite these three differing mechanisms it appears that the IEM model is rarely included in discussion primarily because it is thought to apply to mainly small ions such as salts. Because of this, CRM and CEM are usually compared.

It is suggested that unfolded protein follows the CEM model in which the peptide tails are slowly ejected from the droplet due to redistribution of H+ ions, leaving a highly charged molecule behind (Pimlott & Konermann, 2021; Raab et al., 2021). In contrast there is an impression that native MS in particular favours the CRM model, in which the buffer evaporates, leaving behind a low-charged molecule. This is especially with recent data favouring this model (Pimlott & Konermann, 2021).

Despite this data there currently isn't yet an agreed ESI theory for all types of analytes. The subject of ionisation mechanism has become somewhat controversial, but now the CRM and CEM models are generally accepted to apply to ionization of proteins and

complexes. ESI has demonstrated to be a trusted technique to produce charged macromolecular structures, as shown by the consistent observation of lower charge states associated with native conformations using this technique (Pimlott & Konermann, 2021; Tamara et al., 2021). The main theory behind the reduced charged state of correctly folded membrane proteins is that there are fewer available protonation sites than if it was unfolded (Raab et al., 2021). Because of this the application of ESI has become ubiquitous throughout the structural biology community and within native MS.

1.5.5 High resolution native mass spectrometry

High resolution native mass spectrometry has been developed over many years. It has evolved from alterations/modifications of previous mass spectrometers. Initially the primary mass spectrometer to be used was the time-of-flight mass spectrometer. Before, inclusion of the quadrupole for directing/ selecting precursor ions to give rise to quadrupole time of flight mass spectrometry (Q-ToF). The Q-ToF was used for nMS due to its large m/z range, making it ideal for mass spectrometry of large protein complexes (Tamara et al., 2021). Despite this, it proved to come with drawbacks, mainly the reduced resolving power. This was 'fixed' with the introduction of Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS). FTICR also provided an introduction of mass spectrometry including an ion trap and rather than the steady stream previously used in the Q-ToF(Eliuk & Makarov, 2015; Tamara et al., 2021). This meant rather than using the time of flight to interpret mass/charge, the oscillation frequency could now be interpreted from a signal, significantly increasing resolution (Eliuk & Makarov, 2015; Zubarev & Makarov, 2013). FTICR did however also come with drawbacks one of which was the limited detection of larger macromolecules and large footprint and cost (Zubarev & Makarov, 2013), features that were improved with the introduction of orbitrap mass spectrometry.

1.5.5.1 Orbitrap mass spectrometry

Orbitrap mass spectrometry was derived from FTICR MS. Like the FTICR, the orbitrap was designed to trap generated ions in an orbit. However , the orbitrap maintained an orbit of ions around a central spindle and interprets ion oscillations into a signal (Figure 5)(Eliuk & Makarov, 2015; Zubarev & Makarov, 2013). This can then be translated via Fourier transform calculation into peaks assigned based on their mass-to-charge (m/z) ratio.

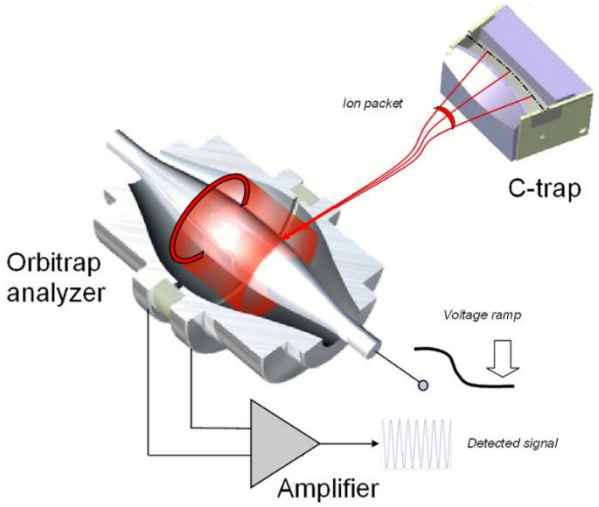

*Figure 5 - **Orbitrap mass analyser function** modified from Zubarev & Makarov, (2013). Showing captured ions being controlled into an orbit around the orbitrap central spindle to derive a signal for mass calculation*

The orbitrap was an improvement over FTICR as it has larger range of mass detection opening the door to high resolution analysis of large protein complexes. Particularly with the inclusion of in-source trapping, this allowed temporary collection and desolvation of generated ions before injection into the orbitrap mass spectrometer (Zubarev & Makarov, 2013). The implementation of orbitrap mass analysers into mass spectrometers gave birth to a high resolution table-top tandem mass spectrometer such as the Q-Exactive UHMR.

1.5.5.2 Q-Exactive UHMR

The Q-Exactive UHMR orbitrap mass spectrometer (Figure 6) is a hybrid mass spectrometer and a significant improvement on the previous models. It was designed to take the advantages of both the Q-ToF and FTICR MS and implement them into one mass spectrometer. It included the quadrupole for directing ions and the orbitraps derived from FTICR improving both ion selection and mass accuracy. Furthermore, the inclusion of the bent flatapole reduced the amount of neutral ions being transmitted to the mass analyser, therefore reducing noise. This is what provides the Q-Exactive with its characteristic ability to provide high resolution spectra.

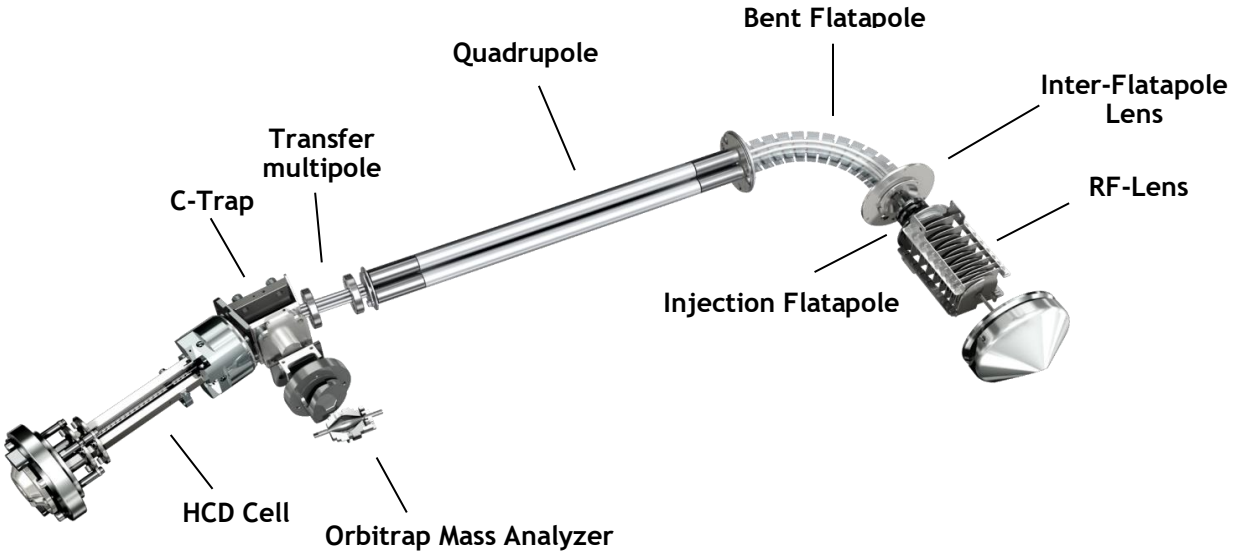

*Figure 6 –Q-**exactive orbitrap quadrupole mass spectrometer** modified from Thermo (2019)*

The Q-ToF however did include a collision cell, which was used for dissociation experiments. This contained inert gas to increase collisions with target ions providing collision induced unfolding (CIU) or collision induced dissociation (CID). The collisions of the target protein ions balanced with higher voltages can provide a harsh enough environment and can cause the dissociation of subcomplexes, this is called as CID. CIU has a similar concept, however the environment provided causes the unfolding of the target protein. As mentioned previously, this can be used in conjunction with precursor ion selection of the quadrupole facilitating protein mass analysis, subcomplex identification and identification of other associated molecules. Identification of subcomplexes and associated ligands can particularly be inferred by comparing the change in mass of a target ion before and during the application of CID.

1.5.5.3 HCD cell

The inclusion of the HCD cell in the Q-exactive instrument facilitated pseudo MS/MS applications such as CID. By increasing collision gas pressure or voltage, the HCD cell allows the increase of ion activation within the mass spectrometer, leading to dissociation of ligands or subcomplexes. In particular, it can be used to probe for posttranslational modification (PTMs) of proteins. PTMs cannot be dissociated as they are covalent modifications. This ability has also been recently used to probe MP-ligand interactions.

1.5.5.4 Probing protein-lipid interaction using nMS

Ultimately the pairing of the orbitrap mass analyser and soft ionisation techniques has allowed the probing of complex biomolecules without affecting quaternary structure as

these reach the orbitrap mass analyser intact (Rose et al., 2012). More specifically, for this investigation, the sensitivity permits the ability to detect small ligands associated with membrane proteins, such as lipids (Gupta et al., 2018). Furthermore, the control of the HCD cell allows the dissociation of associated ligands and molecular interactions to differentiate between PTMs and MP-lipid interactions.

The merging of lipid mass spectrometry and native mass spectrometry has spurred the experimentation throughout the community in developing methodologies for determining membrane protein- lipid thermodynamics, most of which have used high resolution native mass spectrometry. These methodologies look at protein oligomerisation studies after supplementation with specific lipids (Cong et al., 2016). Recently however, it has become apparent that high resolution native mass spectrometry can facilitate the probing of native membrane proteins with native lipids directly after purification. These lipids are shown as adducts within the spectrum associated with a parent ion. By interpreting the change in mass from the parent ion, these adducts provide a mass that is indicative of an additional molecule. These molecules can be characterised based on their m/z informed by known masses in the literature and lipidMAPS database. LC-MS can further supplement nMS to determine unknown sequences and PTMs (Bender & Schmidt, 2019).

1.5.5.5 MS of Solubilised MPs

Use of SMALPs has demonstrated to be valuable for membrane protein research. The SMALP toolbox was recently expanded with the application of SMA based exchange techniques (Hesketh et al., 2020). The hope for this workflow was to retain native conformation and tightly bound lipids/molecules without limiting biophysical characterisation methods. This however, does highlight a need of a standardised workflow for biophysical characterisation of native MPs.

Despite SMALP solubilised membrane proteins being considered a gold standard membrane mimetic system, it has not yet lent itself to use with Native MS. This is particularly evident with a singular paper undertaking nMS of bR after SMA reconstitution (Hoi et al., 2021). This is due to difficulty acquiring a clear spectrum linked with the size and heterogeneity of SMALP particles. Despite the difficulty of using SMA for nMS of membrane proteins, it is apparent that well resolved spectrum containing lipids can be acquired, however this requires the use of synthetic lipids to facilitate the reconstitution in SMA nano discs (Hoi et al., 2021).

The use of Laser Induced Liquid Bead Ion Desorption-MS (LILBIDMS) of SMALP membrane proteins has produced interesting results (Hellwig et al., 2018), showing lower charge states associated with native conformations. This technique is interesting as it was demonstrated to detect both monomeric MP dissociated from the SMALP complex and protein retaining the SMALP complex. The Q-Exactive has been used to produce a similar mode of action to that of LILBID-MS; releasing the target membrane proteins from the SMALP complex. Hoi et al (2021) hypothesised a similar event using the Q-Exactive by using the injection flatapole to liberate the MP from the SMALP and using a gentle potential gradient of the interflatapole lens and bent flatapole to retain the conformation and lipids before detection at the orbitrap mass analyser. The same potential gradient also was used by Urner et al (2020) to produce a native spectrum with well resolved lipid adducts. This study clearly demonstrates the reproducibility of this method and asks the question whether this can be applied to other reconstitution systems.

1.6 Mass photometry

Mass photometry, previously named interferometric scattering mass spectrometry, is a liquid phase analytical technique in its infancy (Li, Y. et al., 2020; Soltermann et al., 2020; Sonn-segev et al., 2020). Mass photometry provides a molecular mass based on landing events of particles on a glass slip (Figure 7). These events are detected based on light scattering of the molecule on the glass surface, presented as a mass distribution in a histogram.

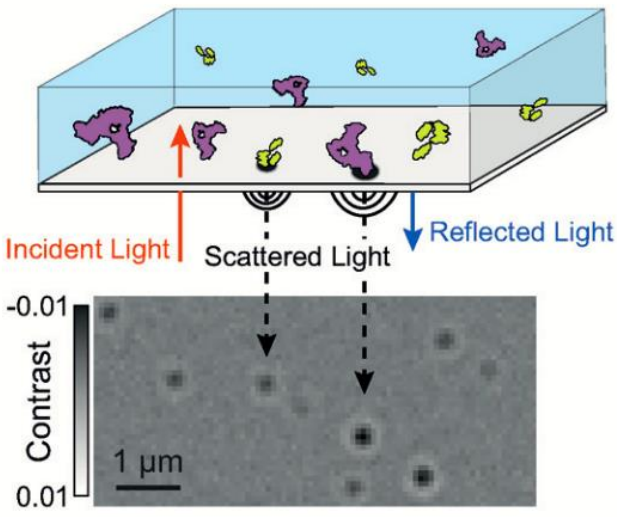

Figure 7 - ***Mass photometry****:* ***detection based on light scattered by proteins*** *modified from Solterman et al (2020). Depicting the interpretation of a mass of a target molecule after landing on a glass slide and being visualised using a specialised camera*

Recently, mass photometry has been used to analyse large proteins solubilized using a variety of isolation systems, such as amphipols, SMALPs and detergents (Olerinyova et al., 2021). Additionally, it has demonstrated value for determining sample heterogeneity

for CryoEM (Sonn-segev et al., 2020). This suggests that mass photometry could become a relatively quick tool to analyse oligomerisation states of complex mixtures, possibly becoming a tool to improve workflows within structural biology. However, mass photometry has not yet been used extensively in this way and so little information is available on how it could apply to the variety of membrane proteins being researched. Because of this, the application of mass photometry will supplement native mass spectrometry findings to provide detailed oligomerisation data of model membrane proteins. It will also help to evaluate the efficacy of mass photometry within structural biology.

1.7 Bacteriorhodopsin as a model membrane protein

1.7.1 Structure and function

Bacteriorhodopsin (HbR) is a light sensitive 27 kDa integral membrane protein with seven transmembrane helices (Figure 8A). HbR is associated with the genus of *H. salinarium*, a halophilic archaebacterium and is a light sensitive proton pump (Bratanov et al., 2015; Gohon et al., 2008). HbR has been the target of much of investigation, thus providing a detailed snapshot of its photocycle associated with its light dependant function. One of the key identifying features of HbR is the all trans retinal group bound to lys216 position via Schiff base (N-H) linkage (Hasegawa et al., 2018; Nango et al., 2016). Upon photoexcitation, conversion of the retinal chromophore to a 13-cis conformation (Figure 8A) sequentially causes the transfer of a proton to asp85. The transfer of a proton within bacteriorhodopsin cause structural changes producing various HbR intermediates. The initial proton exchange between the Schiff base and asp85 has been termed as transition from L-to-M intermediates (Figure 8C). it is accepted that from the M intermediate a proton release group releases this proton into the extracellular space during which the Schiff base is protonated again. A variety of other states of this protein have been identified, leading to further understanding of the HbR photocycle.

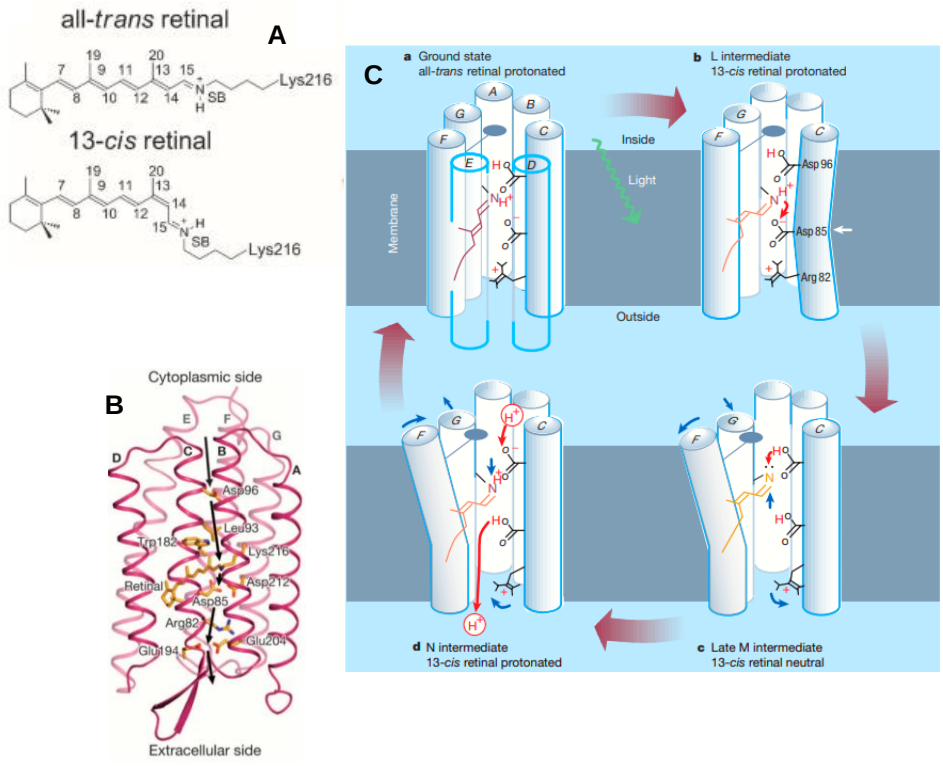

Figure 8 - ***Bacteriorhodopsin photocycle and structural implications****. A) structure of all trans and 13 cis retinal chromophore from (Nango et al., 2016) B)direction of proton transfer to Asp85* (Bondar, A. N. et al., 2004) C) accepted conformational changes of HbR during photocycle from (Kühlbrandt, 2000).

1.7.2 Spectrophotometric quantification of HbR

The identification of the bacteriorhodopsin photocycle has allowed the probing of HbR conformations using spectrophotometric methodologies such as raman spectroscopy (Braimen & Matheis, 1982; Tahara, S. et al., 2019). The spectrophotometric characterisation of HbR during the photocycle revealed a variety of additional intermediates (intermediates K, L, M, N &O) including an early (M1) and late M (M2) intermediates (Figure 9)(Wickstrand et al., 2015) some of which have led to controversy within mechanical transitions between them and retinal confirmations. These intermediates have been reproducibly confirmed using raman spectroscopy and crystallographic studies (Braimen & Matheis, 1982; Tahara, S. et al., 2019; Wickstrand et al., 2015). These investigations do however inform the native characteristics of HbR, being that ground state of HbR contains an all-trans conformation of the retinal chromophore. This in turn, providing the purple colour and 560 nm peak absorbance HbR is associated with.

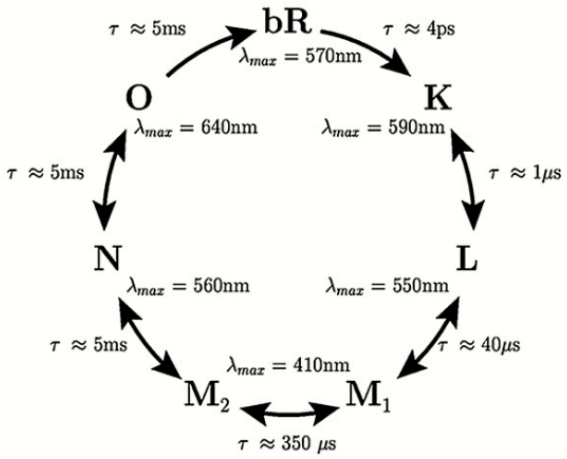

Figure 9 - ***Bacteriorhodopsin photocycle, interrmediates absorbances*** *from Gohon et al (2008) Intermediates K, L, M1, M2, N an O during the photocycle of bacteriorhodopsin and the wavelengths associated with these intermediate structures*

HbRs light sensitive characteristic has allowed a simpler method for direct quantification and characterisation of HbR. As mentioned, the retinal in its typical all trans conformation with associated lipids group gives a strong absorbance as 560 nm. This allows the quantification of HbR based on the absorbance using it's extinction Coefficient of 63000 $M^{-1}cm^{-1}$ (Jeganathan et al., 2019; Shiu et al., 2014).

1.7.3 Purple membrane lipid environment

It has been demonstrated that HbR is tightly associated with its native lipids within the purple membrane of *H. salinarium* (Hoi et al., 2021; Reichow & Gonen, 2009). This makes it suitable as a model of membrane protein native MP investigations.

It has been established that the local lipids aid the stability of the bacteriorhodopsin trimer in these harsh hypersaline environment *H. salinarium* grows (Jawurek et al., 2018; Yamagami et al., 2019). These lipids have been found in crystal structures and even after delipidation suggesting a tight association with HbR. Because of this, HbR has been the subject of a variety of protein- lipid interaction investigations. These investigations introduced a variety of lipids thought to be within the purple membrane. Structures of each lipid were used to inform expected mass for mass spectrometry during this investigation.

As expected, a mixture of glycolipids and phospholipids are present within the PM. Each lipid possesses an archaeol fatty acid group making the hydrophobic fatty acid tails for each PM lipid. The variable head group is what gives PM lipids their uniqueness. The main lipids mentioned within the literature appear to be STGA-1 and PGP-Me.

In order to interpret HbR native MS data, an understanding of current knowledge on how lipids are organised around HbR has to be established. Despite the consistent investigation of HbR and the PM, an agreed abundance of PM lipids had yet to be established (table 1).

Table 1 - purple membrane lipid composition in the literature

Paper	Year	Lipid	Quantity: HbR monomer	Technique(s)	Reconstitution system
Renner et al	2005	PGP-me	4.4	NMR	Triton x 100, SDS, Deutereated SDS
		PG	~0.5		
		BPG	~1.5		
		STGD-1	1.9		
		Glycolipids	2		
		Lipid chains	17		
Krebs et al (cited kates et al (1986) and Grigoreiff et al (1996))	2000	Phospholipids	6-7	Xray crystallography	
		Sulfoglycolipids	2-3		
Essen et al	1998	STGA-1	3/trimer	Xray crystallography	
Angelini et al	2009	BPG		MALDI-TOF/MS	
		PGP-me			
		PGS			
		PS			
		STGD-1			
		STGD-1-PA			
Essen et al	1998	PGP-me		MS	beta octyl glucopyranoside
		PGS		MS	
		PG		MS	
		STGA-1	2	MS, Xray crystallography	
Hoi et al	2020	PGP-me	1	nMS	SMA

Inada et al	2020	STGA-1	6 / trimer	spectroflurometry (proton influx)	Triton X 100
		PGP-me	24/ trimer		
		STGA-1	7.3/ trimer	CD spectroscopy	
		PGP-me	25.5/trimer		
Reichow & gonen	2009	STGA-1	3 / trimer	Xray crystallography	

The problem with establishing an abundance of lipids surrounding bacteriorhodopsin appears to be due to abundance of certain lipids within the PM being low particularly glycolipids (STGD-1 and STGA-1).

HbR associated lipid composition of the PM have primarily been investigated using X-ray crystallography of the HbR trimer. These investigations provide a snapshot of lipids directly associated with the HbR protein.

Other investigations, most notably Renner et al (2005), Angelini et al (2010) and Essen et al (1998) investigated the purple membrane directly with use of lipid mass spectrometry paired with TLC and NMR to identify unknown lipids and providing m/z values for each lipid (Table 2).

Throughout these investigations, PGP-me has been identified in nMS spectrum after solubilisation using SMA from HbR-DMPC proteosomes (Hoi et al., 2021) and it is clear from data presented in table 1 that PGP-me is the most abundant lipid within the PM. Considering this, the retention of this lipid after reconstitution directly from the PM could suggest a comparable amount of lipid retention as the methodology demonstrated by Hoi et al (2021). This could also therefore suggest a similar ability to their methodology for providing crucial lipids for retention of membrane protein native fold after reconstitution. Charge state distribution can provide another measure of this by comparing the resulting charge state distribution with that of Hoi et al (2021). Despite these data, nMS of HbR has not appeared to provide information on any other lipids that are closely associated.

Table 2 -PM lipid and small molecule masses and expected M/z values for nMS

Lipid				Experimental mass (m/z)	
Abbreviated name	Full name	Chemical formula	Expected mass		
PGP-me	2,3-di-O-phytanyl-sn-glycero-1-phospho-(3'-sn-glycerol-1'-methyl phosphate)	$C_{47}H_{98}O_{11}P_2$	901.2*	900.65	
BPG	2,3-Bisphosphoglycerate	$C_3H_3O_{10}P_2$		1520.2	
PG	Phosphatidalglycerol			805.7	805.6
PGS	Phosphatidylglycerol sulfate		744.18	885.6	
STGD-1		$C_{58}H_{18}O_{183}S$	1098.79	1218.1	
STGA-1			1181.51		1216
	Squalene	$C_{30}H_{47}$	407.69		

*From https://pubchem.ncbi.nlm.nih.gov/compound/PGP-Me structure

1.7.4 Lipid associated function of HbR

Of particular interest recently is the STGA-1 glycolipid. This is because it has been identified within HbR crystal structures to be associated with HbR monomer interfaces within the trimer (Essen et al., 1998; Reichow & Gonen, 2009). This lead to speculation whether STGA-1 was responsible for HbR trimer formation. Inada et al (2019) went on to demonstrate that trimer formation and HbR activity is highly reliant on the presence of STGA-1. They did this by optimising HbR function with STGA-1 lipid using Circular dichroism and proton influx spectrophotometry. Circular dichromism analysed the HbR photocycle based on the STGA-1 concentration. Proton influx spectrophotometry facilitated the measuring of proton pumping activity of HbR based on the STAG-1 concentration within a DMPC liposome. This data allowed them to calculate on the amount of lipids per HbR molecules and demonstrated the requirement of this lipid in trimerization and the photocycle.

Another investigation looked at the effect of PGP-Me on HbR functionality showing a reliance of HbR on the PGP-Me headgroup for sufficient function (Cui et al., 2015). It was also noted that reduced activity was demonstrated in delipidated HbR. These findings suggest the importance of lipids in the function of HbR. This also demonstrates that variability in lipid retention could be demonstrated by spectrophometric characterisation mentioned previously.

1.7.5 Native MS of bacteriorhodopsin

The presented information has demonstrated how HbR can and has provided a reliable model MP with the wealth of investigations and structural knowledge. It is now left to this section to demonstrate the current understanding of native MS to provide an idea of what constitutes a native conformation of HbR using nMS.

The desired retention of native conformation requires that this investigation has a method to characterise a native fold associated with each reconstitution system. It has already been suggested that lower charge states are characteristic of native MS, this is because of the lack of protonation points associated with a correctly folded MPs. This lack of protonation points is partially due to some of the transmembrane regions providing less solvent accessibility due to its 3D structure, this means certain regions of the protein can't be accessed by the ammonium acetate and cannot be ionised during ESI. Ultimately reducing the observe charge state when correctly folded seen compared to its unfolded charge state. Charge state distribution therefore, could present as a good measure of the retention of membrane protein folding associated with each reconstitution system. Both Hoi et al (2021) and Shannon et al (2016) would provide a good standard both demonstrating the abundant +9 and +8 charge states when investigating HbR using nanodisc and detergent based reconstitution systems using native MS. Hoi et al (2021) also demonstrated a lower +5 charge state of monomeric HbR using the previously mentioned SMALP reconstitution method. This suggests that the +5 charge state is closer to a native conformation than that of nano disc and detergent isolation and should be used as a bench mark. This lower charge state could also be due to solvent accessibility (ammounium acetate) provided by the SMALP during ESI and has yet to be reproduced. Many different investigations have provided a comparison of reconstitution systems and their effect on target proteins. However, this investigation has a strong focus on understanding the relationship between reconstitution system used ; retention of endogenous ligands and native fold. It should be then tested on different lipid environments to demonstrate the reliability of each system on a variety of membrane

compositions. To do this, this investigation employs the use of an *E. coli* expression system to provide a contrast from the unique and tough lipid environment of *H. salinarium*.

1.7.6 Heterologous expression of EbR

Heterologous expression of membrane proteins is a commonly used technique to express and investigate membrane proteins. Despite this, heterologous expression of functional HbR in *E. coli* had proven to be difficult. This was unexpected as other members of the rhodopsin family demonstrated the ability to express in *E. coli* (Bratanov et al., 2015). It had been determined that trouble expression HbR in *E. coli* was because of degradation of much of the EbR immediately after synthesis. This is what made it initially hard to express in a foreign cell species such as *E. coli*. It has been proven that cell free synthesis methods could prove useful in *E. coli* based bacteriorhodopsin expression (Nekrasova et al., 2010). This investigation was further replicated by Bratanov et al (2015). This investigation identified regions of the HbR mRNA that had unstable free folding energies. Implying that the synthesis of HbR in *E. coli* would not satisfy the 'positive inside' rule. These sections were then modified in an attempt to improve expression and insertion of EbR into the *E. coli* membrane. These modifications of HbR included the replacement of sections of the first and second transmembrane helices. These sections were replaced with corresponding sections of helices from a homolog rhodopsin called sensory rhodopsin II (SRII) (Figure 10). This subsequently demonstrated stable expression of EbR.

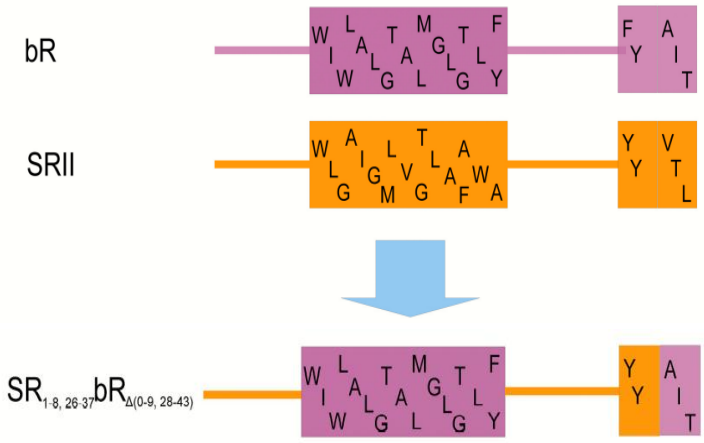

Figure 10- **Modification of HbR1st and 2nd transmembrane helices for expression within E. coli** modified from Bratanov et al (2015) resulting amino acid differences after modification of first and second transmembrane for expression in E. coli

Despite the modification, the resulting EbR demonstrated a remarkable resemblance to its wild-type counterpart. The wavelength at which it absorbs remains 560 nm. This is also beneficial for future bacteriorhodopsin investigations in providing quicker expression times of 2-3 days, compared to HbR expression of 7-14 days depending on your protocol.

1.7.7 *E. coli* lipid environment

The *E. coli* membrane environment provides a good contrast to the purple membrane not only due to the drastic differences in lipid structure and composition, but also because the equal extensive investigation it has undergone. It has been established that *E. coli* lipid membranes are primarily composed of phosphatidylethanolamine(PE) (75%), with the rest composing of Phosphatidyl glycerol (PG) and cardiolipid (CL)(Sohlenkamp & Geiger, 2016). It should be noted despite this, *E. coli* has also been shown to vary its lipid production and composition based on the environment. Give the frequent use of *E. coli* based expression systems, the lipid structure has been investigated using a variety of techniques. This has allowed the development of the lipidMAPS database that was designed to provide lipid identification tools. In particular a database exists for nMS based analysis of lipids providing expected m/z values of membrane lipids to allow for the desired workflow presented.

1.8 Aims and objectives

In light of the topic of this investigation, the presented information provides a tool for a different direction of study. The ability to express a membrane protein in two cellular systems with different cellular membrane compositions. Allowing us to elucidate how differing reconstitution methods affect structure, oligomerisation and associated ligand retention of the target MPs. Furthermore, providing a model as to how well each reconstitution system applies to different membrane environments. To address this objective we will:

- express bacteriorhodopsin in both *H. salinarium* and *E. coli*
- optimise reconstitution of EbR and HbR with SMA, DDM, A8-35 and novel amphipols.
- Use the presented native mass spectrometry methodology to determine protein masses, charge state and associated ligands, this data can inform the ability of each reconstruction system to retain native protein fold and native lipids.
- Use spectrophotometry to determine mass spectrometry data and inform structural changes associated with each reconstitution system and closely associated lipids
- Implement mass photometry to supplement masses seen within mass spectrums and investigate oligomerisation state of both EbR and HbR after reconstitution with SMA, DDM, A8-35, novel amphipols and G1 OGD

This is with the intention of collating this data to inform future implementation of these novel reconstitution molecules within membrane protein structural biology. Additionally,

presenting methodologies to overcome current challenges associated with using these molecules for membrane protein characterisation.

Chapter 2: Materials and Methodology

2.1 LB and LB agar preparation

LB was prepared by adding 2.5 g Yeast extract, 5 g tryptone, 5 g sodium chloride to a final volume of 1 L double distilled water. After fully dissolving the components, the media was sterilised by autoclaving at 120°C for 15 mins. LB-agar was prepared by the further addition of 7.5 g/L of bacteriological agent before autoclaving, antibiotic (kanamycin 50 µg/mL) was added and mixed before pouring and allowing to set in a sterile environment.

2.2 ZYP-5052 media preparation

The ZYP-5052 autoinduction media was used and prepared as in Bratanov et al., 2015 for heterologous expression of EbR. briefly, stock solutions 50X 5052 (25 % (w/v) glycerol, 2.5 % (w/v) glucose, 10 % (w/v) α-lactose monohydrate), 20X P (0.5M Na_2HPO_4, 1 M KH_2PO_4, 1 M NH_4 Cl) were prepared. 20XP was further adjusted to pH 7.0 before autoclaving. stock solution of 1M $MgSO_4$ was also prepared, this and 50X 5052 was sterilised using a 0.22 µm microfilter.

ZYP-5052 media was then prepared from stock solutions as required using amounts described in Table 3 (excluding 50x 5052 component) before autoclaving at 120 °C for 15 mins. 50x 5052 (volume described in table 3) was then added in a sterile fashion before inoculation of the media.

Table 3 - ZYP-5052 media prepared from 1M $MgSO_4$, 50x5052 and 20x P stocks

Solution/substance	**Amount per 1L**
1 M $MgSO_4$	1 mL
50x 5052	20 mL
20x P	50 mL
Yeast extract	20 g
Tryptone	32 g

2.3 Modified Basal Media preparation

Basal Media preparation followed Jeganathan et al (2019) using a modified basal media to improve HbR expression. Basal media was prepared by mixing, 10 g peptone, 3 g sodium tricitrate, 20 g $MgSO_4 \cdot H_2O$, 250 g NaCl in 900 mL of double distilled water. This was then adjusted to pH 7.2 using NaOH before sterilising by autoclaving at 120°C for 15 mins.

2.4 Molecular biology

2.4.1 Plasmid construct

EbR Plasmid constructs were kindly provided by Dr Vincent Postis, Leeds Beckett University. They consisted of a modified Bacteriorhodopsin construct ligated into a pET vector. To create a stock of usable plasmid construct, pET-EbR was transformed into omnnimax 2 T2 chemically competent cells with pRARE plasmid as in section 3.6. After plating on LB-kanamycin (50 µg/mL) and incubating overnight at 37 °C. A colony was suspended in 10 mL LB containing 10 µL kanamycin (50 mg/mL), The culture was then incubated 37°C in a shaker incubator overnight. Plasmid was then purified using a monarch plasmid miniprep kit (NEB) as described in the kit protocol

2.4.2 DNA Screening

To ensure the EbR gene was of desired size in the used plasmid. It was screened using DNA gel electrophoresis. It was prepared first for restriction digestion to excise the EbR gene form the pET -EbR plasmid (Figure 11). In short, 0.3 µL of XbaI and XhoI was added to pET-EbR in 1 X cutsmart™ buffer (Invitrogen) solution for a final volume of 8 µL. The mixture was then incubated at 37°C for 1 hour before addition of 2 µL of 6 x loading dye (Invitrogen) for electrophoresis on a 1% agarose gel made with 1x TAE solution (40 mM tris acetate, 1mM EDTA) and run in 1xTAE gel electrophoresis buffer. This was run at 100 V for 25 mins before visualisation.

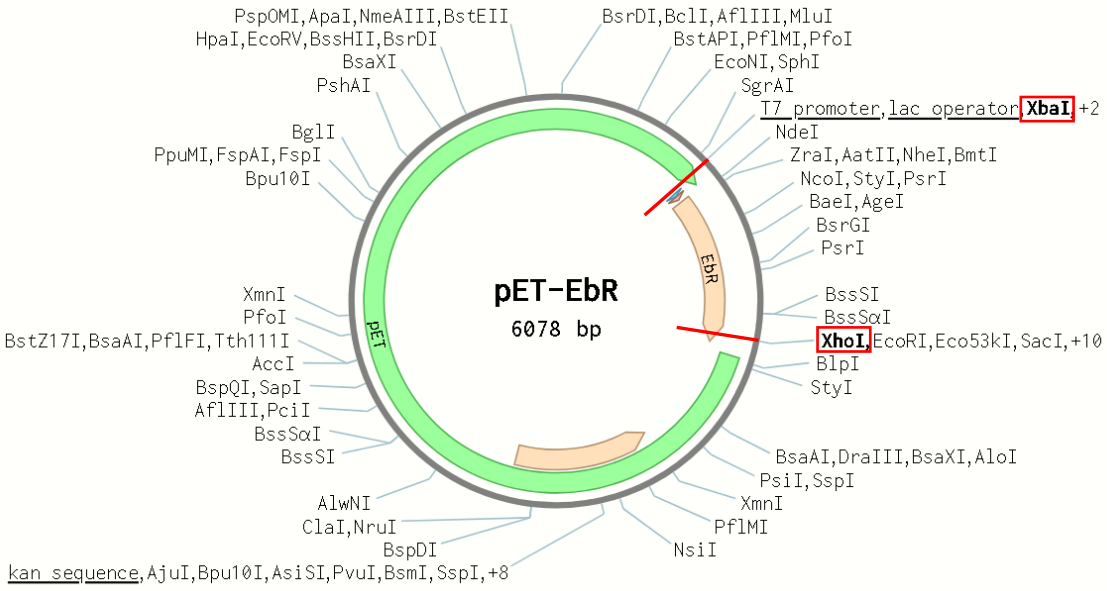

Figure 11- **pET EbR plasmid map** with highlighted restriction sites (red boxes) flanking the EbR gene for DNA screening with red lines showing the position of them on the plasmid

DNA minipreps were also quantified using Nanodrop-spectrophotometry to determine DNA concentration. DNA concentration (~100 ng/µL) informed the use of DNA sequencing using the primer 5'-TAATACGACTCACTATAGGG-3' (T7 promoter) for further confirmation of the EbR gene.

2.5 Bacterial transformations

E. coli BL21(DE3) gold containing pRARE plasmid was transformed with pET-EbR construct. 1 µL of purified vector was mixed with thawed *E. coli* BL21 (DE3) cells containing pRARE plasmid. This was incubated on ice for 45 mins before heat shocking at 42 °C for 45 seconds. The cells were removed promptly and left to cool on ice for 5 mins. 450 µL of LB was added; this suspension was incubated at 37 °C for 1 hour. Finally, 50 µL of the transformed cells were plated on LB agar containing 50 µg/mL kanamycin.

2.6 Protein expression

2.6.1 EbR

Expression of modified HbR was as described in Bratanov et al (2015). briefly, one colony of *E. coli* BL21(DE3) + pRARE with pET-EbR was used to inoculate in LB media with kanamycin (50 µg/ml) and incubated at 37 °C overnight. After measuring the OD600 of the overnight culture, it was used to inoculate prepared ZYP-5052. This was incubated at 37 °C with agitation at 200 RPM to an OD600 of 1.0-1.2, Sequentially adding 400 µL of 50 mM all-trans retinal (sigma) in ethanol. The culture was further incubated overnight at 20 °C before pelleting and isolating *E. coli* membranes.

2.6.2 Overexpression HbR

ATCC Halobacterium salinarum (Harrison and Kennedy) Elazari-Volcani (ATCC 33170) was purchased and activated according to the company's provided instructions. Overexpression of HbR was informed by data from Jeganathan et al (2019). In short, a *H. salnarium* colony was used to inoculate 50 mL of ATCC 217 (1 g/L K_2HPO_4, 0.5 g/L $MgSO_4$, 10 g/L Yeast extract, pH 7.0-7.2) media and cultivated to for 7 days. 25 mL of culture was used to inoculate basal media supplemented with 0.2 g/L L-arginine and cultivated for a further 6 days at 37 °C with agitation (150 rpm). Cells were harvested at 8000 rpm for 1 hour before purple membrane preparation.

2.7 Membrane protein purification

2.7.1 *E. coli* membrane isolation

For the membrane preparation of *E. coli* membranes expressing EbR, 5 x 400 mL were induced and pelleted. This pellet was resuspended in 100-150 mL of 20 mM Tris – HCL, 0.5 mM EDTA, pH 7.5 before adding protease inhibitor cocktail. This suspension was

homogenised using a mechanical homogeniser before running through a cell disruptor at 30 kpsi twice. The disrupted cells were subjected to centrifugation at 14000 g for 40 mins at 4 °C.

The subsequent supernatant was collected and centrifuged at 131000 g for 2 hours. After discarding the supernatant, the pellet was washed by resuspending in 2 mL Tris -HCl pH 7.5 and centrifuged for 2 hours at 131000 g. This step was repeated a further three times before flash freezing into pellets and storing at -80 °C.

2.7.2 Purple membrane isolation

Isolation of the purple membrane from *H. salinarium* was carried out as in Shiu et al (2013) with minor modifications. In short, 75 mL *H. salinarium* culture was collected by centrifuging at 8000 g for 1 hr. This was resuspended in 30 mL containing 75 µL DNAse solution and incubated at room temperature for 30 mins. Disrupted cells were spun at 24000 RPM for 30 mins at 4 °C. The pellet was resuspended in 30mL DDH_2O and pelleted again. The purple pellet was resuspended in a minimal volume of distilled water and layered over a sucrose gradient of three layers (20 %, 40 % and 60 %) described in Hoi et al (2021) . The purple band was collected and dialysed against 2 L DDH_2O using a 12,000-14,000 MWCO dialysis membrane. The dialysed sample was collected and designated as Purple membrane.

2.7.3 EbR DDM IMAC isolation

Flash frozen *E. coli* membranes were weighed, suspended and homogenised in solubilisation buffer (20 mM Tris-HCl pH 8, 10 % glycerol, 250 mM NaCl) to 1 mg/mL before adding DDM to 1 %. This was incubated for 2 hours at 4°C with gentle agitation. The solubilisation was centrifuged at 100,000 g for 45 mins to remove insoluble material. 1 mL of IMAC cobalt resin beads (500 µL bed volume) was washed (3x DDH_2O and 3 x Wash buffer) and pelleted before resuspended in the soluble material mixture. The suspension was incubated overnight at 4°C in the dark with gentle agitation before running through a gravity purification column (biorad). After, four washes of 10mL wash buffer (20 mM Tris-HCl pH 7.5, 250 mM NaCl, 10 % glycerol, 0.05 % DDM, 20 mM imidazole). EbR was eluted with elution buffer (20 mM Tris HCl pH 8, 250 mM NaCl, 10 % glycerol, 0.025 % DDM, 300 mM imidazole) as 10 x 1 mL fractions and incubating at 4°C with agitation each time before eluting. The samples were protected from light upon elution. Samples were dialysed to remove imidazole and concentrated using a 10,000 MWCO centrifugal concentrator and then spun at 20,000 g for 20 mins to remove aggregates.

2.7.4 EbR SMA and amphipol isolation

Flash frozen *E. coli* membrane pellets were prepared for solubilisation as above. SMA was added to 2 % or A8-35 was added to 0.2 % (w/v) before incubating at room temperature with gentle agitation for 2 hours. The insoluble fraction was removed by centrifugation at 100,000 g for 45 mins. Protein was then purified using gravity flow column (biorad), washing with wash buffer (20 mM Tris-HCl pH 8, 250 mM NaCl, 10 % glycerol, 20 mM imidazole) three times and eluting 10 x 1 mL (10 x 500 µL for novel amphipols) fraction with elution buffer (20 mM Tris HCl pH 8, 250 mM NaCl, 10 % glycerol, 300 mM imidazole). All pure protein samples were subsequently pooled and dialysed in solubilisation buffer overnight to remove imidazole. Dialysed samples were concentrated using a 10,000 MWCO centrifugal concentrator and then spun at 20,000 g for 20 mins to remove aggregates.

2.7.5 EbR modular detergent isolation

Modular detergent G1 oligoglycerol detergent (OGD) were synthesised and kindly provided by Dr Leanohard Urner (University of Oxford). Isolation of EbR using these detergents were carried out as described in Urner et al (2020) with minor modification to volumes informed by Leonhard Urner (personal communication). In short, flash frozen membranes were homogenised in 1.5 mL solubilisation buffer (20 mM Tris, 100 mM NaCl, 20 % glycerol pH 7.4) to 1 mg/mL. The G1 OGD stock solution (10 %) was added to a concentration of 1 % and incubated at 4 °C for 16 hrs with light agitation. 500 µL bed volumes of Cobalt agarose beads were washed (3 x 500 µL water) and loaded in an empty chromatography column (sigma). The column was washed with 1 x 500 µL wash buffer (50 mM, 200 mM NaCl, 20 mM imidazole, pH 8) ; 1x 500 µL elution buffer (50 mM Tris, 100 mM NaCl, 500 mM imidazole, 10 % glycerol, pH 8) and 5x 500 µL wash buffer. After pelleting cell debris from solubilisation mixture (21000 g, 40 mins, 4 °C), the beads were resuspended in protein solution and incubated at 4°C for 30 mins with light agitation. The flow through was collected before washing the column with 5 x 500 µL wash buffer, 2 x Buffer mixture (wash/elution buffer, v/v, 9/1). Protein was eluted using 5 x550 µL elution buffer. Protein solutions were pooled and concentrated in a centrifugal concentrator (10 MWKO), washed five times with elution buffer (without imidazole) before concentrating to ~1mg/ml.

2.7.6 IMAC purification optimisation

Optimisations were carried out to improve the capture of solubilised EbR by Cobalt beads and therefore improving final yield from purification. These were carried out by

solubilising EbR at 1 mg/ml in a final volume of 1.5 mL. Washed cobalt beads with a bed volume of 75 µL (150 µL of slurry) were added and incubating at 4 °C with agitation. Samples (500 µL) were taken during incubation with the Cobalt beads at time points 30 mins, 1 hr, 2 hr, 3 hr and overnight. These samples were centrifuged at 20,000 g for 1 hr and the supernatant was prepped and western blotted as in section 2.9.2.

2.7.7 HbR solubilisation

Enriched PM in water was diluted to 2 mg/mL in 200 µL and pelleted at 20,000 g for 1 hour before resuspending in solubilisation buffer (20 mM Tris- HCl, pH 7.5, 250 mM – 4 M NaCl,10 % glycerol). The solubilisation agent was then added to the desired percentage as in table 4. The solubilisation mixture was then incubated in the dark at 4 °C (25 °C for amphipols) for 2 hours. Insoluble material was removed by centrifuging at 20,000 g for 1 hour, the supernatant was collected in a fresh tube for analysis.

Table 4 – Concentration of reconstitution systems for solubilisation of HbR from the PM

Reagent	Used concentration (w/v)
DDM	1 %
SMA	2 %
A8-35	0.2 %
C_6-C_2-50	0.1 %
C_8-C_0-50	0.5 %
G1 OGD	1 %

Concentration of solubilised HbR was quantified using nanodrop spectrophotometry or interpreted by SDS-PAGE by comparing the A560 pre solubilisation and post centrifugation and expressed as a percentage.

2.8 MP characterisation

2.8.1 SDS-PAGE

Two different gels were used throughout this investigation to facilitate the analysis of both proteins at various stages.

EbR was mixed with SDS sample buffer to a 1x concentration and incubated at room temperature for ~2 mins before loading on a 4-12 % Bolt™ Tris-glycine gel with Bolt™ MES running buffer (Invitrogen™). 4-12 % gels were stained by soaking in instant blue coomasie stain (abcam) with gentle agitation for 1 hour before viewing.

HbR in SMA, DDM, A8-35, C_6-C_2-50 and C_8-C_0-50 was run on a prepared 15% acrylamide gel, samples were prepared the same as EbR before running using SDS electrophoresis buffer. 15 % prepared gels were soaked with prepared Coomasie R-250 blue stain for 30 mins before destaining using multiple washes of destaining solution (10 % acetic acid, 40 % methanol) before viewing.

2.9 Western blot

Western blot was used to screen for expression of EbR in *E. coli,* solubilisation from whole cell lysates and for protein identification from protein samples or binding assays.

2.9.1 Whole cell lysates

For protein expression screening, whole cell lysates were prepared. In short 400 µL of induced cells were spun down as 13400 RPM in an Eppendorf table-top centrifuge for 5 mins. These were then resuspended in 100 µL of lysis buffer (50 mM HEPES, 25 % Sucrose, 5 mM $MgCl_2$, 1% Triton X-100). This was incubated at 22°C with agitation (1000 RPM) for 35 mins. Whole cell lysates (10 µL) were mixed with 2 µL 6x sample buffer and run on a prepared 10% gel.

The gel was removed transferred to a Watman paper and nitrocellulose sandwich previously incubated in transfer buffer (20% methanol, Tris 3.05 g/L, glycine 14.4 g/L, SDS 1 g/L). The sandwich was then run in a transfer buffer for 1hr on ice and with a mixing flea.

2.9.2 Protein samples

Purified protein samples (15 µL) were mixed with 3 µL 6x sample buffer (0.1M Tris HCl, pH 8, 0.68 mg/mL glycerol, 0.1 g/mL SDS, 1.79 mg/ml EDTA 0.6 mg/ml Y-pyronin), before loading. On a Precast 4-12 % Bolt™ Tris-glycine gel in MES running buffer as in section 2.8.1. The gel was then separated from the cast and transferred to a trans blot turbo MIDI transfer pack (biorad) transferred using the preset MIDI gel function on a transblot turbo (biorad).

Both the whole cell lysate and protein sample gel were blotted the same after transfer. The nitrocellulose was then washed with PBS-T (1 X PBS, 1 % Tween) by soaking for 5 mins with agitation, this step was repeated two more times before incubating with blocking buffer (5 % dried milk powder in PBS-T) at 4 °C with agitation overnight.

The blocking buffer was then discarded and the nitrocellulose washed with PBS-T three times, incubating at RT for 5 mins and discarding after each time.

The nitrocellulose was incubated in anti-His tag antibody (1:5000) for 1 hour before removing and washing with PBS-T three times with 5 mins incubation time per wash with agitation. ECL was added to the nitrocellulose for 1 min before viewing.

2.10 MP concentration determination

2.10.1 HbR

Two methods were used for determining HbR and EbR concentration. HbR in purple membrane was determined using a nanodrop spectrophotometer (Denovix), the absorbance at 560 nm (A560) was measured and the concentration was determined using HbR extinction coefficient at 560 nm (63000 $M^{-1}CM^{-1}$). Measurements were taken in triplicates to ensure accuracy. Where shifts in absorbances were observed A280 was used to determine concentration.

2.10.2 EbR

EbR concentration in isolated membranes was determined using BCA assay kit (thermofisher) as detailed by the kit instructions.

2.11 Biophysical characterisation

2.11.1 Native Mass Spectrometry

Samples were buffer exchanged into 200 mM ammonium acetate buffer before loading into the electrospray needle. Samples were run on a Q-Exactive UHMR Plus Quadrupole-orbitrap mass spectrometer. Potential gradient was optimised for membrane proteins (Hoi et al., 2021; Urner et al., 2020). In short, spray voltage 1.2 kV, in-source trapping was set at 15 V- 125 V depending on the sample, injection flatapole 7.9 V, interflatapole lens 6.94 V, bent flatapole 5.9 V, transfer multipole 4 V, C trap entrance lens 5.8 V. To differentiate covalent modification, adducts were dissociated using HCD cell (0-250 V). Spectra were visualised and analysed using Masslynx V4.2. Protein peaks were assigned based on expected masses, EbR mass based on sequence mass and HbR was based on mass within the literature (27000 Da). Lipid adducts were identified based masses found within the literature mentioned in section 1.7.3.

2.11.2 SMA chemical dissociation

For chemical dissociation of SMA from HbR-SMA samples, the sample was buffer exchanged into 200mM Ammonium acetate once. 3 µL of buffer exchanged sample was then mixed with 1 µL Dissociation solution (200 mM Ammonium acetate, 40 µM Magnesium acetate) before immediately loading into the ESI needle. nMS settings were set as above (section 2.11.1) before acquiring mass spectrum and characterised using HCD cell for activation experiments.

2.11.3 Mass photometry

Concentrated samples were diluted in respective solubilisation buffer or PBS to ~100nM as necessary before analysing. Glass slides were washed DDH_2O, then isopropanol, before drying via nitrogen stream. After calibrating the PBS droplet (16 µL), 4 µL of diluted sample was added and homogenised. A video (60 seconds) was acquired on refynMP (Refyn ltd) and analysed using DiscoveryMP (Refyn ltd) to provide a mass distribution histogram.

Chapter 3: Results

Given the reproducibility of HbR expression and purification in the literature, it is first desirable to determine whether EbR expression would be successful. This was first done by ensuring the pET-EbR plasmid was as required for protein expression.

3.1 EbR

3.1.1 Confirmation of pET-EbR plasmid

pET-EbR was analysed to ensure the plasmid provided contained the desired EbR gene. This was done using restriction digestion flanking the EbR gene. After digesting concomitantly with XbaI and XhoI, the DNA gel was expected to yield 886 bp (EbR gene) and 519 2bp (plasmid) bands (Figure 12A). This digestion was compared to a virtual digestion using a benchling, an open source software used for molecular biology (Figure 11B).

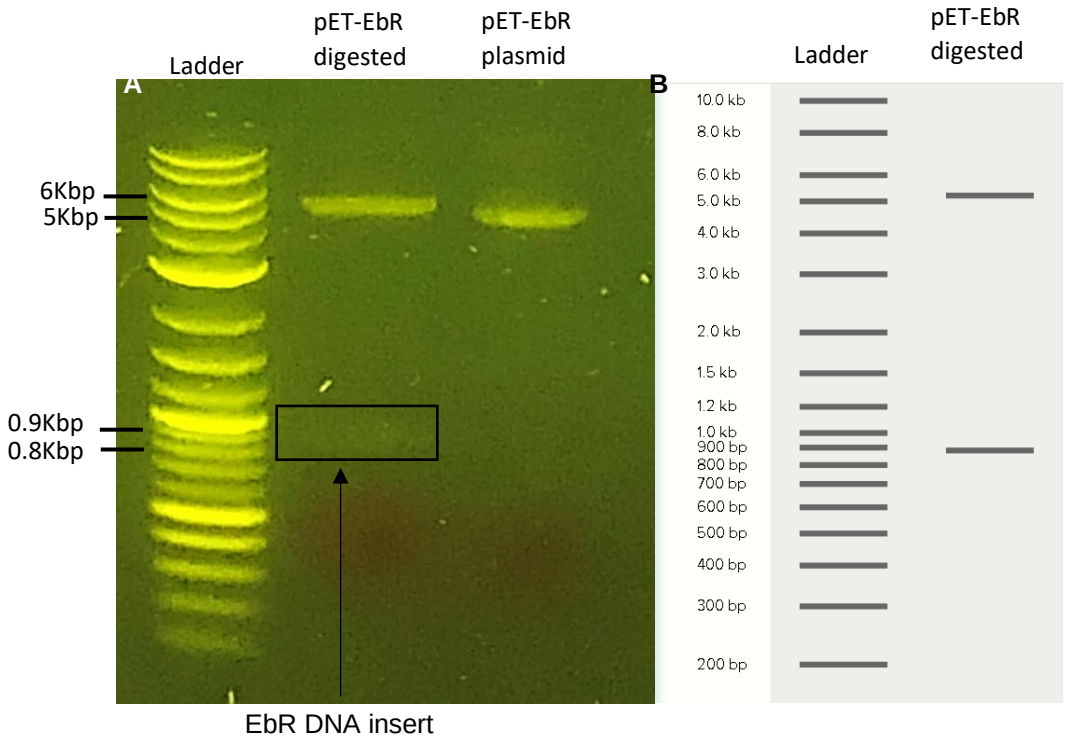

Figure 12- ***DNA gel electrophoresis of pET-EbR plasmid after digestion*** *with XhoI and XbaI showing expected Plasmid 886 bp and 5192 bp bands conforming plasmid to be desired construct. Ladder is 1 kbp plus DNA ladder (NEB)*

The digestion confirmed the presence of the gene within the plasmid. The EbR gene was further sequenced using T7 promoter primer just before the EbR gene achieving full coverage of the EbR gene. The sequencing confirmed no mutation was present as compared to the sequence of 4XXJ (PDB) using benchling software. This permitted that transformation and expression of the plasmid could be carried out.

3.1.2 EbR expression

The plasmid was then transformed into *E. coli* BL21 (DE3) + pRARE for expression. Protein expression was carried out as in section 2.6.1. A control was also included which had not been induced by not adding the 50x5052 component of the ZYP-5052 media, required for expression. Whole cell lysates were run on SDS-PAGE to determine if expression had been achieved (Figure 13).

The whole cell lysates were prepared and run on SDS-PAGE with two different sample volumes (5 µL and 9 µL). This was because of uncertainty associated with the protein content of samples. The desired size of protein for this expression was ~24 kDa. There was no observed difference between the none induced (ctrl) and the induced samples (Ind). There are a few reasons for this inconclusive result, one of which being the absence of expression and the other being the presence of minimal expression. From this the band thickness of the 9µL samples were sufficient for analysis by SDS-PAGE.

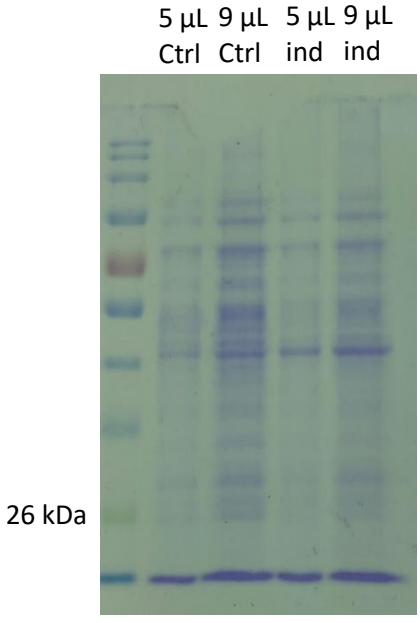

Figure 13- ***SDS-PAGE of cell lysate before and after induction*** *no difference between control and induced cells were observed suggesting minimal or no expression of EbR.*

It was decided that these results could possibly be due to the previously mentioned problem with expression of EbR seen in Bratanov et al (2015). This required a more sensitive technique be used to investigate this. Given that EbR is a His-tagged protein, a technique such as western blot can be employed (Figure 14). The expression (Well 1) was compared to a non-induced control sample (Well 3) and an induced *E. coli* +pRARE not containing the pET-EbR plasmid (well 3). Given the clarity observed previously using 9 µL of whole cell lysates, this volume was run on SDS-PAGE and then western blotted with an anti-His-tag antibody. The same 24 kDa His-tagged band was expected for this experiment also. Well 1 showed a strong expression of the expected ~24 kDa His-tagged protein with an extra band below thought to be degradation during loading. Well 2 showed the same band but fainter than the expression. This implied the presence 'leaky' expression when not induced. It was also confirmed that the expression wasn't related to the pRARE plasmid with no bands visible from induced *E. coli*+ pRARE (well 3) .

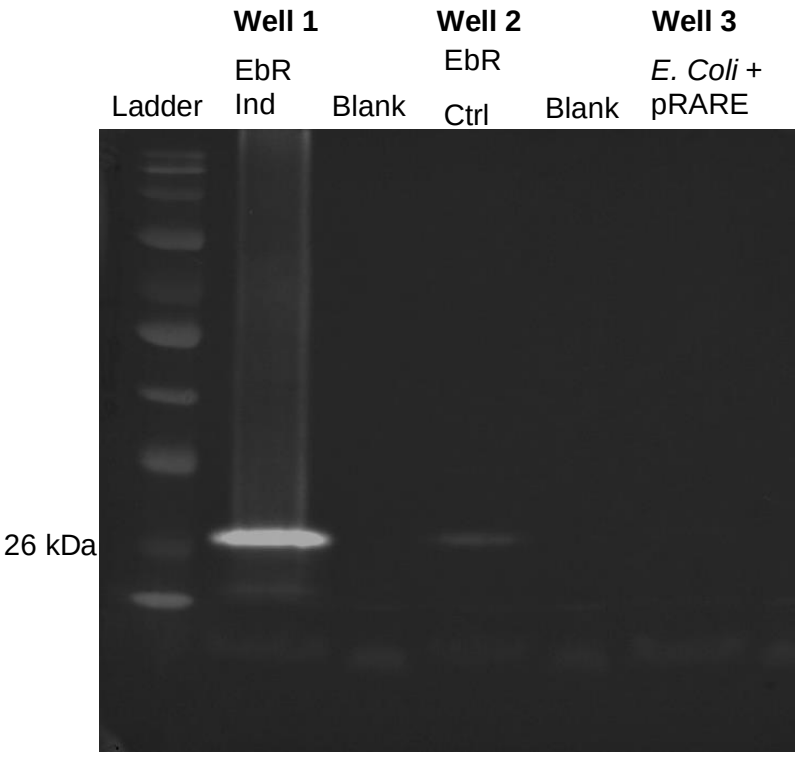

Figure 14- ***Western blot with anti-His Tag anti body*** *after expression of EbR showing expression of a his tagged protein at desired size of 24 kDa and possible degradation with the band below (well 1). Well 2 also implied 'leaky' expression in non-induced cultures. Well 3 also showing expression is not associated with pRARE plasmid.*

This showed that a His-tagged protein was being expressed at the desired size (~24 kDa) implying the successful expression of EbR. This means that a large scale expression could be carried out for *E. coli* membrane isolation and subsequent purification of EbR using the different reconstitution systems.

3.1.3 Solubilisation of EbR

After EbR expression, the *E. coli* cell pellet were disrupted and the subsequent membranes were washed as in 2.7.1. Subsequent membranes were flash frozen and designated as *E. coli* membranes. An initial small scale solubilisation of EbR using the various agents could be carried. This would aid optimisation of EbR solubilisation. Typically, DDM is used at 1 % (w/v), SMA is used at 2 % (w/v) and A8-35 is used at 0.2 % (w/v). Because of this, these concentrations were used for this solubilisation. It was not known which concentration of C_6-C_2-50 and C_8-C_0-50 would be optimal and so two concentrations were used for both (0.1 % and 0.5 %), this would also provide information on concentration dependant solubilisation from the novel amphipols. G1 family of oligoglycerol detergents (G1OGD) is usually used at 1% by Urner et al (2021) and so this was the chosen concentration for this assay. These solubilisations were carried out on homogenised membranes and so western blot would elucidate comparison of total EbR present (T) and solubilised EbR (S) (Figure 15). Solubilised EbR was prepared post solubilisation by centrifuging the mixture at high speed to remove insoluble material.

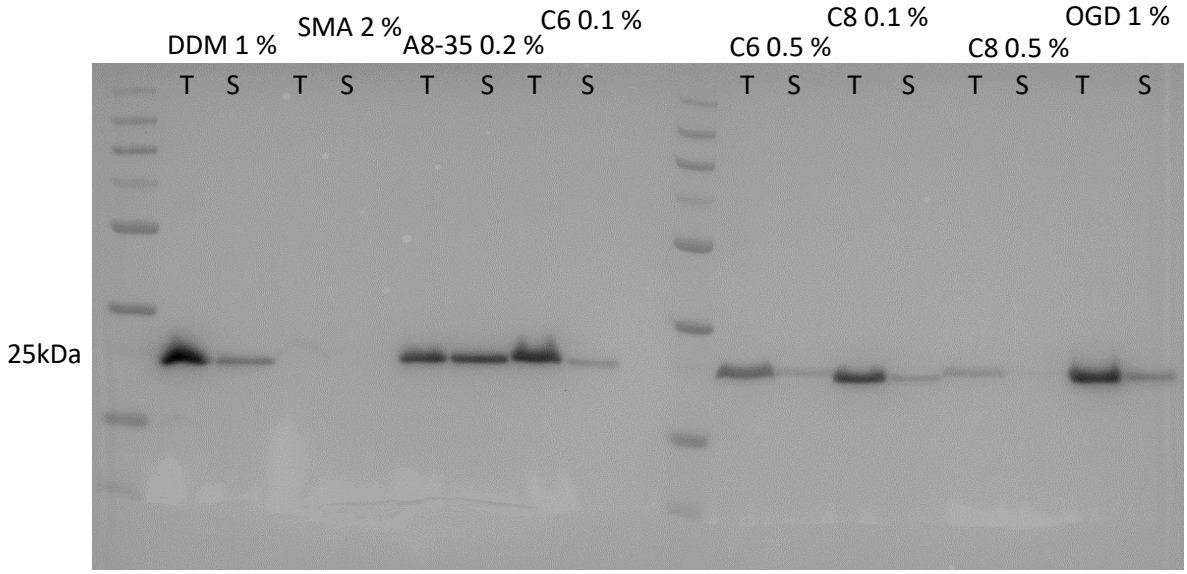

Figure 15- ***Western blot of EbR after solubilisation from E. coli membranes****. T= Total (soluble + insoluble) S=solubilised material. C6 demonstrated optimal solubilisation at 0.1 %; C8 appeared to show improved solubilisation at 0.5 %. OGD solubilisation was similar to that of DDM at the same concentration.*

To provide a semi-quantitative analysis of the solubilisation efficiency shown in this western blot; For each sample, Total (T) and soluble material (S) was compared using imageJ. The negative image of Figure 15 allowed the calculation of a solubilisation

efficiency (Figure 16). From this we can see A8-35 and C_6-C_2-50 (C6) outperformed all the other agents. In comparison also, DDM at 1 % appeared to have a lower solubilisation efficiency compared to G1 OGD at 1% directly from *E. coli* membranes. This showed the use of cyclic amphipols C_6-C_2-50 at 0.5 % and C_8-C_0-50 at 0.1 % would be sufficient for solubilisation of EbR using the novel amphipols. This is because of the higher solubilisation efficiency found in these concentrations compared to the other concentrations used. It was also determined that DDM, SMA and A8-35 concentrations shown were more than sufficient for purification and characterisation of EbR. SMA was not included in this measurement as the western blot bands appeared to be extremely faint would not sufficiently represent the solubilisation efficiency of SMA.

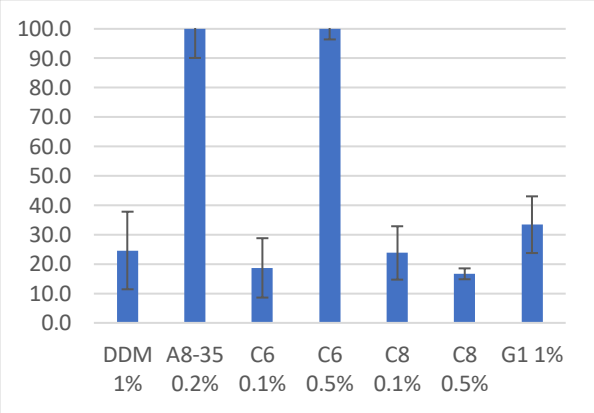

Figure 16- **Percentage solubilisation of EbR using various agents** quantified using band brightness (an average of three measurements) using imageJ A8-35 0.2 % and C6 0.5 % had to be capped at 100 %.

This data showed that DDM 1 %, C_6-C_2-50 0.5 %, C_8-C_0-50 0.1 %, G1 OGD 1 % would be sufficient for IMAC purification of EbR and so were used individually for optimisation and further understanding of the samples observed.

3.1.4 EbR DDM purification optimisation

Given the previous isolation of EbR using DDM by Bratanov et al (2015), this would provide as a comparison for other isolation systems. It was noted however that the protocol they used for isolation could be harsh on the membrane protein given the pH changes and that precipitation observed during dialysis. This required that the protocol be changed to provide gentle conditions during purification to prevent the precipitation and pH changes. A purification was implemented that used 250 mM NaCl at pH 8 for the solubilisation, wash and elution buffer. This meant that only the imidazole concentration varied between them. Initial IMAC purification of DDM-EbR yielded a well solubilised protein of the desired size (not shown). These results also showed proteins in the larger kDa range. To improve the quality of isolation an extra wash step was added before eluting (Figure 17 left panel). This extra step did remove more contaminant proteins. Due

to the purity, elution fractions 8-10 were pooled, dialysed and concentrated. The result showed a high abundance of EbR over contaminant proteins (Figure 17 right panel) which is suitable for mass spectrometry

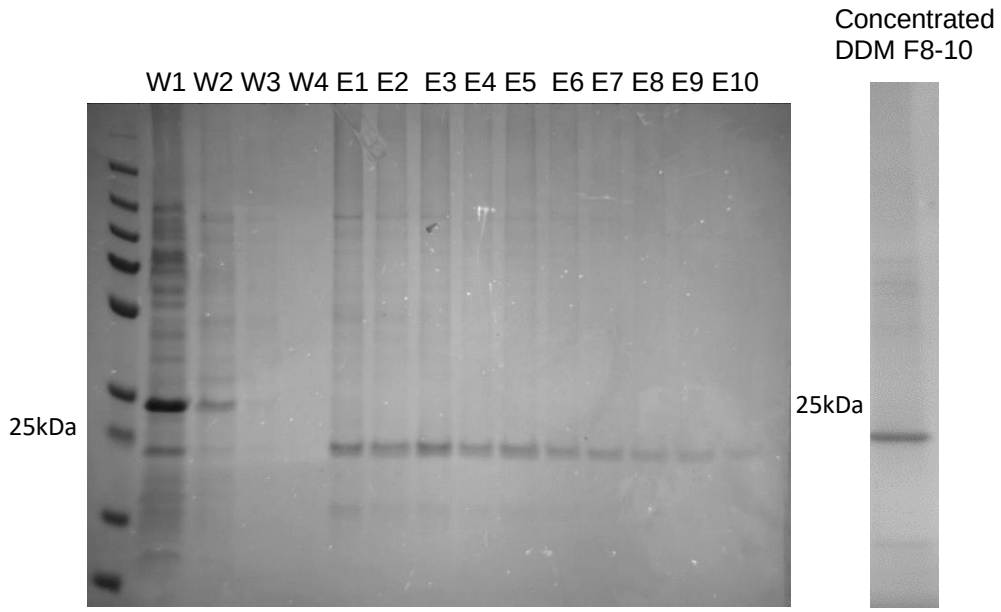

*Figure 17 – **SDS-PAGE of DDM purified EbR** using gravity flow column Wash and elution of EbR using affinity chromatography after solubilisation using 1 % DDM. W = Wash E=Elution right panel) showing concentrated sample of EbR-DDM fraction(F) 8-10 with minimal contaminant proteins in the high kDa range*

3.1.5 EbR G1 OGD

The modular detergents from the G1 family of oligoglycerol detergents (G1 OGD) described in section 1.4.1 were also tested to purify EbR. When first using the G1 OGD, the protocol was informed by Dr Leonhard Urner (Technische Universitat Dortmund). The protocol was as described in his earlier work in Urner et al (2020) but volumes were significantly reduced from a 5 mL of membranes to 1.5mL of homogenised membranes. After purification, flow through (FT), washes (W) 1-4 and elutions (E) 1-5 were run on SDS-PAGE to examine purity and presence of desired size band (~24 kDa) (Figure 18). It should be noted that there are seven wash steps for this protocol, the last three were not shown on this gel. It does not appear that any protein was eluted during the wash steps, implying the capture and elution of all EbR present. The elutions from this purification of EbR by G1 OGD yielded bands at the desired size (red box) with no other bands present. This implies the elution fractions are relatively pure with no contaminant proteins and therefore can be pooled, dialysed and concentrated.

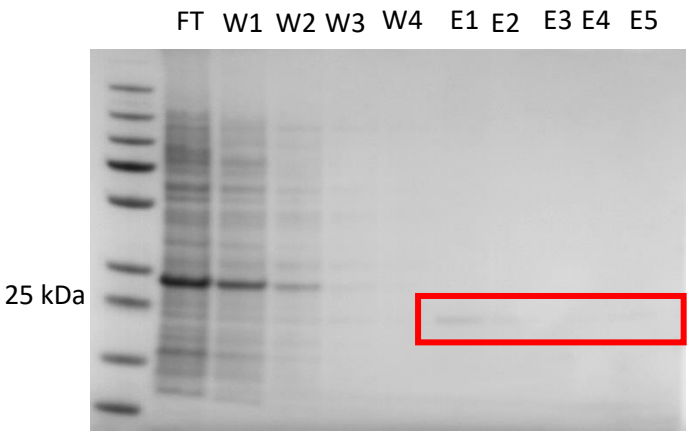

*Figure 18**- EbR G1 IMAC purification SDS-PAGE** not showing wash (W) 5-7 showing protein elution (E) samples 1-4 and Flow through (FT). With no eluted EbR in the FT and washes which shows capture of all EbR present. Elutions showed only one protein at ~24 kDa (red box) implying the successful purification of EbR without contaminant protein, sufficient for mass spectrometry.*

From this it was determined that these samples were suitable for mass spectrometry given the presence of a singular band at the desired size.

3.1.6 EbR SMA purification optimisation

The absence/faintness of SMA bands on the solubilisation assay western blot (figure 15) made the observed solubilisation efficiency (Figure 16) questionable. To further assess solubilisation efficiency and optimise IMAC purification, an investigation was carried out. This was to investigate the required incubation time of solubilised EbR solution with the cobalt beads, this would further ensure that all EbR-SMA was captured and subsequently eluted, improving yield. Solubilisation sample was prepared as before and samples were taken before (well 1) and after solubilisation in SMA (well 2&3) (Figure 19). After the addition of cobalt beads and subsequent incubation, samples were taken at time points 30 mins, 1 hr, 2 hr, 3 hr and overnight (ON) and centrifuged at high speed (well 4-8) to remove insoluble material and beads. The supernatants were then western blotted (Figure 18A). If the binding of the protein to the beads was successful, the protein will be pelleted with the beads and it will disappear from the supernatant, as compared to well 3. Well 3 represents the SMA solubilisation of EbR directly from membranes, where the solubilised sample (without beads) was centrifuged at high speed to remove of insoluble material representing the amount of solubilised EbR (well 3). This data demonstrated that most if not all of the solubilised EbR SMA was captured by the cobalt beads after 30 minutes incubation due the absence of any His-tagged protein band (Figure 18 well 4 compared to well 3).

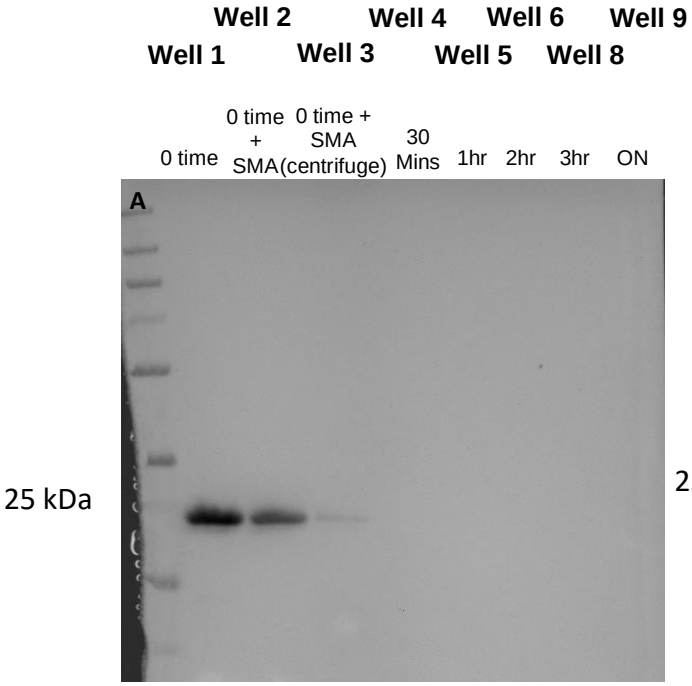

*Figure 19- **Western blot of EbR SMA IMAC optimisation** including centrifuged post solubilisation sample this confirmed the solubilisation of EbR by SMA and demonstrated that incubation of the reconstituted samples with cobalt resin for 30 minutes was sufficient to capture all reconstituted EbR.*

To provide a better representation of solubilisation efficiency of SMA, this western blot was analysed with Image J. The negative image was used to create a semiquantitative analysis of band brightness to inform a solubilisation efficiency. Image J was used to create a peak signal based on the band brightness and then the peak area was plotted in a bar graph (Figure 20). When comparing the total amount of material (total) to the amount solubilised (soluble) this equates to ~8 % solubilisation. Significantly lower than previously observed, but more conclusive given the clear presence of both total and soluble material bands in the western blot.

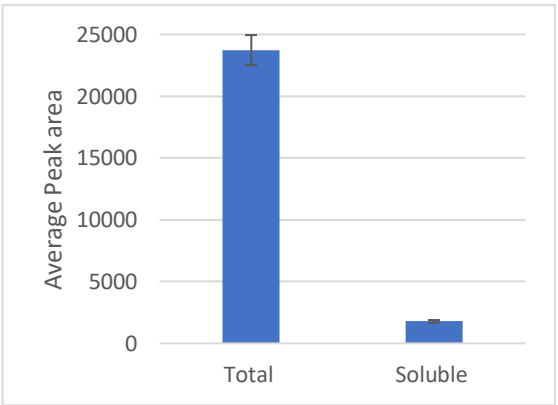

Figure 20– **Image J quantitation of band brightness to inform solubilisation efficiency** showing The average of three peak area measurements (based on band brightness) processed by image J displaying ~8 % solubilisation efficiency

This demonstrates that the solubilisation is not as efficiency as previously thought but still allowed purification of EbR using SMA.

3.1.7 EbR A8-35 IMAC purification

Previously A8-35 showed to solubilise 100 % of EbR present used at 0.2 % and so this is the concentration used for purification of EbR by IMAC. The elutions from this purification were then run on SDS-PAGE to determine purity and identity (Figure 21). These elutions only contained one protein band just below 25 kDa, indicative of EbR as previously seen. no other protein bands were present suggesting no contaminant proteins were present in this sample.

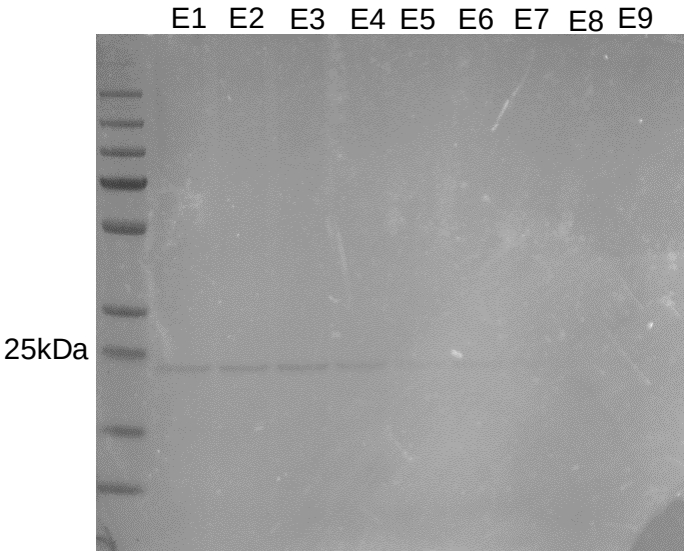

*Figure 21**- EbR A8-35 IMAC purification**. Showing elutions (E) 1-9. EbR is present in all elution fractions with no contaminant proteins shown.*

All the fractions containing EbR were pooled, dialysed and concentrated to yield an A8-35 EbR stock. This could then be used as comparison for the novel amphipols to determine limitations of these polymers in membrane protein investigations.

3.1.8 Novel amphipol EbR IMAC purification

IMAC purification of EbR using the novel amphipols C_6-C_2-50 and C_8-C_0-50 was carried out at 0.5 % and 0.1 % respectively and then run on SDS-PAGE (Figure 22). These concentration for solubilisation were used because of the solubilisation efficiency previously seen in the solubilisation assay (

Figure 15). The washes in these preparations appeared to be clear, although containing small amount of EbR that are being eluted during these stages. Furthermore, The EbR protein bands in the elutions (~24kDa) were not as intense compared to the previous DDM purification suggesting less protein has been purified with C_6-C_2-50 and C_8-C_0-50 than DDM. An abundant copurified protein band at ~35 kDa was also observed within the elutions with both amphipols. It was discussed that removing the co-purified protein from this preparation would be difficult even using methods such as size exclusion chromatography due to the similarity in size.

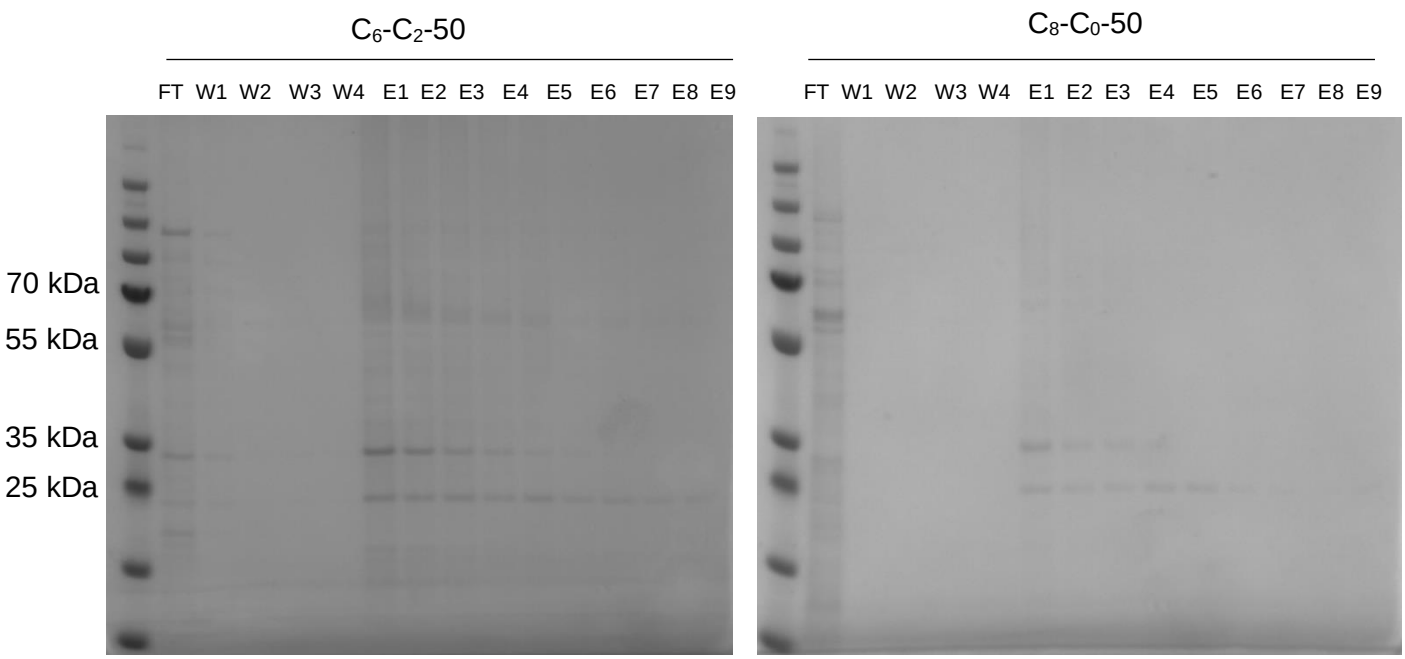

Figure 22- ***EbR IMAC purification using C_6-C_2-50, C_8-C_0-50*** *abundant band of EbR and copurified protein (OmpA) FT = Flow through W = Wash E = Elution*

It was clear that the abundance of other contaminant proteins would not pose a problem during mass spectrometry. However, the copurified 35 kDa protein would indeed show within the MS spectrum. We decided to identify the contaminant by submitting the sample to the University of Leeds MS facility for protein identification using trypsin digestion and LC-MS. The copurified protein digestion was compared to the uniprot database and found to be outer membrane protein A (OmpA).

3.1.9 EbR C_6-C_2-50 characterisation

The presented purifications all demonstrated a suitability for native mass spectrometry and so it was decided to characterise these with nMS starting with the EbR C_6-C_2-50. The identification of the abundant co-purified protein as OmpA allowed the identification of a protein mass during mass spectrometry. OmpA has a monomeric mass of 37200.76 Da (uniprot) derived from its protein sequence and therefore the determination of any oligomers seen within the mass spectrum would be easily spotted. EbR monomeric mass derived from its sequence is 29113.03 Da, this meant that dimeric EbR would be 58226.03 Da. The solubilised samples were prepared for nMS by buffer exchanging once into 200mM ammonium acetate and settings were set as in 2.11.1. The processed mass spectrometry data of C_6-C_2-50 EbR (Figure 23 top panel) was compared to the smoothed data (Figure 23 bottom panel) to ensure no peaks were removed during the processing. This showed abundant peaks at 66639.13 Da with the charge state (z) z=+16 (Series A).

The expected mass of a EbR dimer being 58226.03 Da, this size is suggesting of the EbR dimer would be present in combination with ligands that could be either lipids or polymers. The unexpected mass of the main protein means that these additional molecules can not be identified from this spectrum. Furthermore, there does appear to be the presence of another protein series here with low abundance around 6000 m/z, the mass of could not be assigned due to the low abundance. Because of the lack of clarity behind the masses of oligomers present it was determined that using the CID capabilities of the Q-Exactive UHMR would facilitate characterisation of the sample and a gentler approach might show the other oligomers with lower abundances.

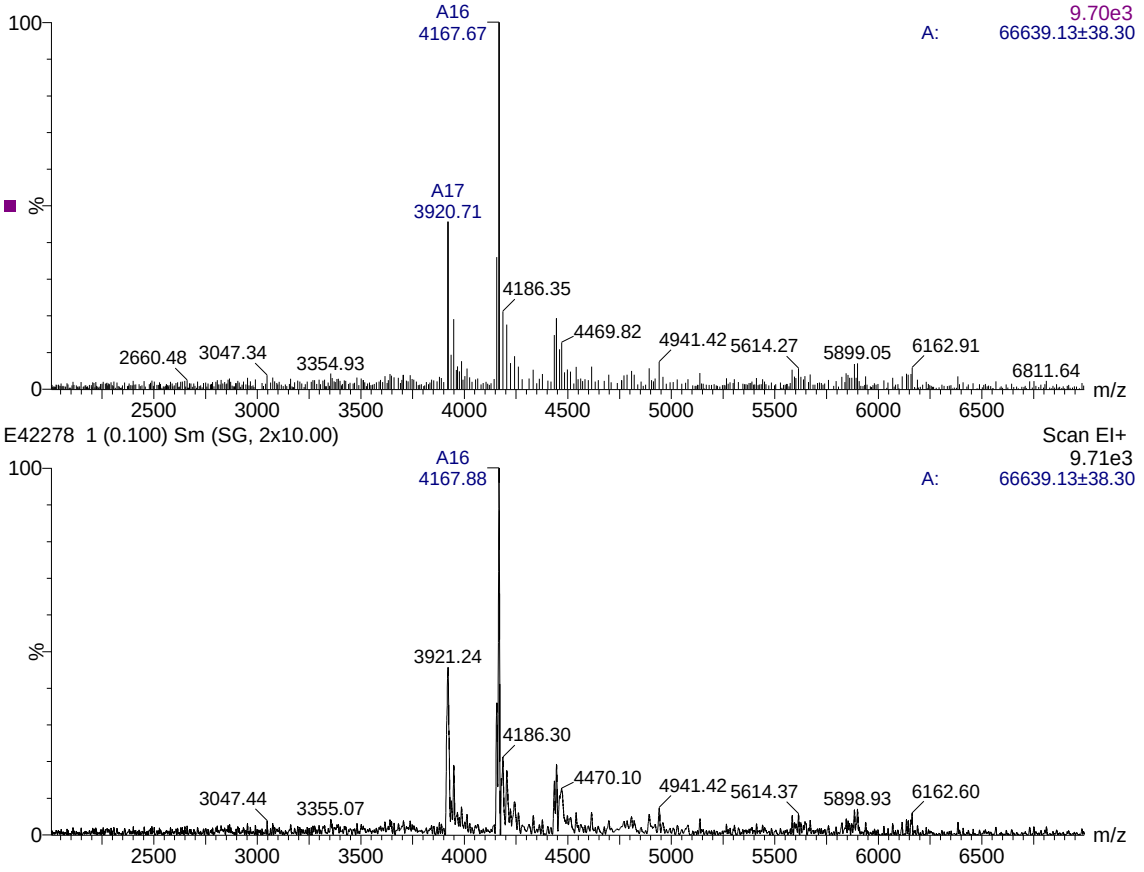

*Figure 23- **Mass spectrometry EbR in C_6-C_2-50** showing unexpected masses not attributed to EbR or OmpA.*

It was hoped the activation experiment would increase abundance of EbR monomers and dissociate associated ligands to show stripped protein giving an accurate mass (Figure 24). During mass spectrometry with no additional activation (Figure 24 bottom panel) a similar main protein series with a mass 66818.73 Da z=+16 (Series B) was observed with another protein series at a mass of 10249.88 Da z=+20 (Series C). Series B appeared to come with additional ligands (series B1 and B2). To determine the mass

of these adducts however, the main protein series mass had to be taken into account. This is because after using the HCD cell at 190V, the voltage-based dissociation platform of the Q-Exactive UHMR, to increase ion activation both main protein series masses decreased (Figure 23 top panel). This suggests that without the use of the HCD cell, the settings retained molecules of masses 270.85 Da (series B) and 100.8 Da (series C). Furthermore, previously results using a higher desolvation voltage produced a protein series of a similar mass at 66639.13 Da (Figure 23) suggesting this is likely to be the true mass of the observed molecules. The additional

molecules, series B1 and B2, also dissociated with the use of the HCD cell implying the presence of a protein-lipid complex. These adduct masses of 67076.01 kDa (series B1) and 67920.49 kDa (series B2) were used to calculate suspected lipid masses by subtracting the main protein series mass (66547.88 Da) . This implied possible lipid masses of 528.13 Da (series B1) and 1372.62 Da (Series B1). The Increase in HCD cell voltage (Figure 24 top panel) also showed to be an increase in B series and a decrease in the C series suggesting the 102949.88 Da series is being dismantled producing more 66818.73 Da molecules implying that these molecules are derived from a similar protein complex.

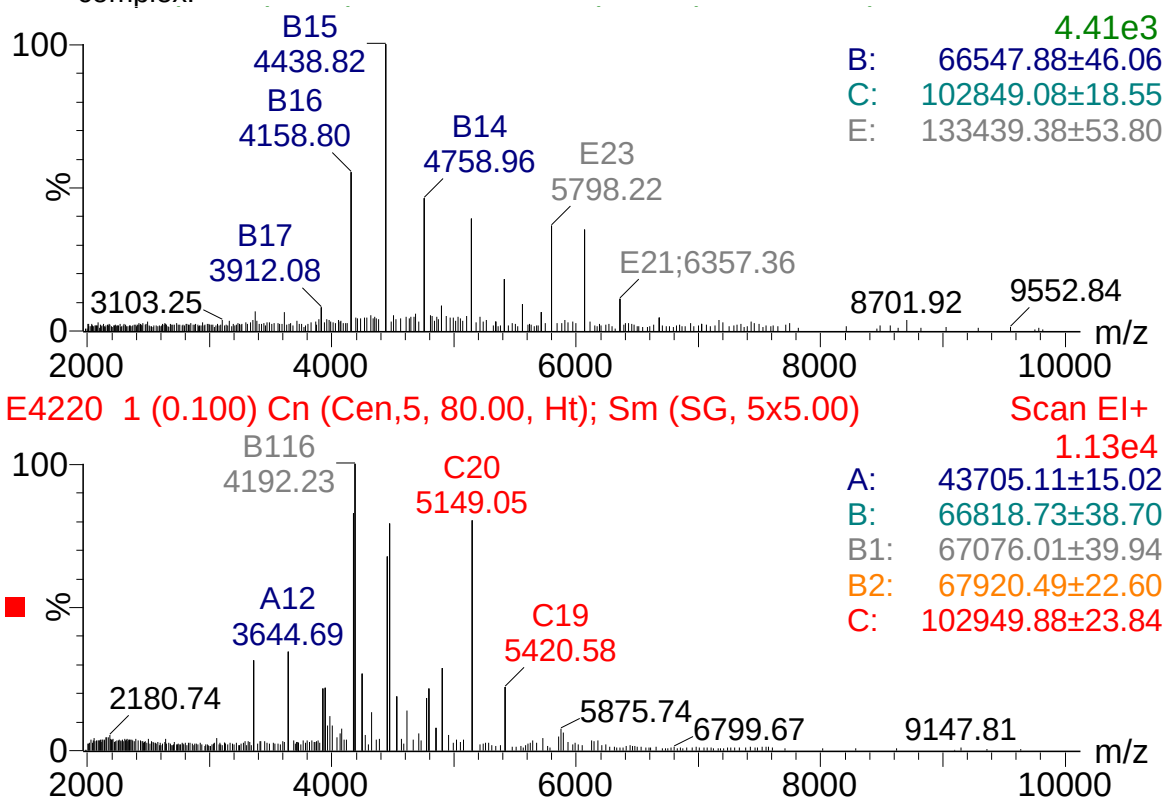

*Figure 24– **nMS activation experiment of C_6-C_2-50 preparations** showing the sample 66.8kDa protein series with 102.8kDa with unknown identity*

During nMS of the C_6-C_2-50 isolated EbR sample (Figure 23) adducts were found with masses 528.13 Da and 1372.62 Da. As previously mentioned, they are derived from the protein mass that is expected to be the true mass of the protein series observed (66547.88 Da). Due to the variability of *E. coli* membrane lipids the lipidMAPs database was used to determine the possible identity of the lipids based on the observed masses (Figure 25). Both of these results suggesting phosphatidylinositol triphosphate (PIP3) lipids with differing lipid chain lengths. Adducts B1 suggested a lipopolysaccharide with a potassium ion associated and PIP3 molecules with sodium ions associated, possibly from residual NaCl from the solubilisation buffer. Adducts B2 suggest the lipid came associated with an ammonium ion, presumably from the ammonium acetate buffer.

Adduct B1

Input Mass	Matched Mass	Delta	Name	Formula	Ion
528.13	528.1759	.0459	LPS 16:4	C22H36NO9PK	[M+K]+
528.13	528.1390	.0090	PIP3 30:6	C39H66O22P4Na2	[M+2Na]2+
528.13	528.1208	.0092	PIP3 29:7;O	C38H62O23P4Na2	[M+2Na]2+

Adduct B2

Input Mass	Matched Mass	Delta	Name	Formula	Ion
1372.61	1372.6414	.0314	PIP3 54:10;O	C63H106O23P4	[M+NH4]+

Figure 25 - ***Results of lipid maps search using suspected masses of lipids adducts*** *suggesting presence of PIP3 in associated with both lipid adducts*

These results implied the retention of lipids using the C_6-C_2-50 polymer, although the identity of these lipids can't be confirmed from this data alone. The reader should be reminded that for the interpretation of the observed results the polymer is suspected of still being associated with the protein. This presents a model suggesting the attachment of lipids external to the polymer-protein nano disc. This means it is not clear if the lipids are associated with the protein or the polymer. Because of this the results were still somewhat unclear and not well validated as there was still the possibility of these series being contaminant proteins.

To determine if there were contaminant proteins present, It was decided to run the concentrated polymer isolated EbR on SDS-PAGE (Figure 26). This would provide clarity as to whether contaminant proteins are being observed within the spectra. It was also compared A8-35 preparations as a reference. It should be noted that Run2 of C_6-C_2-50 preparations was used for mass spectrometry. Given the absence of proteins in both

run1 of C_6-C_2-50 and C_8-C_0-50 preparations they were not used. This showed the known contaminant protein (OmpA) and the monomeric EbR band (~24kDa). This also showed other contaminant proteins not present within the previous purification gel at ~55 kDa.

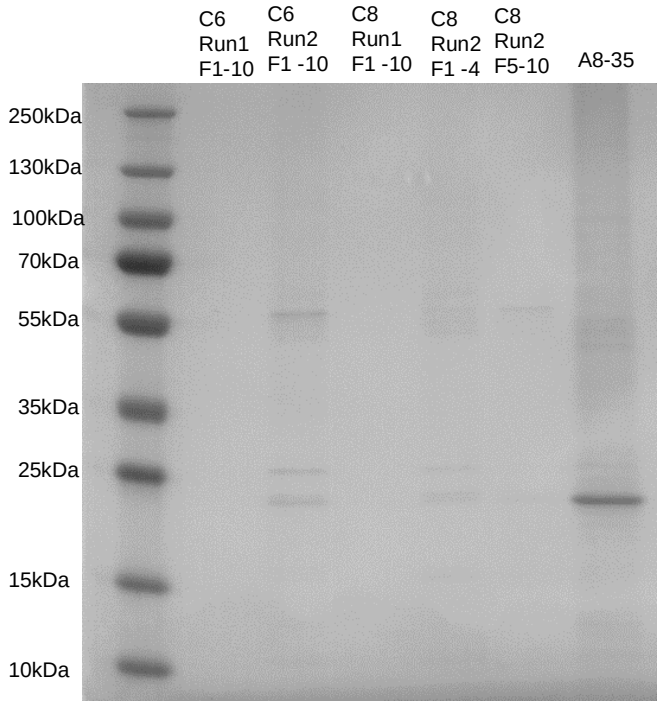

Figure 26 - ***SDS-PAGE of concentrated C6 and C8 EbR samples*** *showing known contaminant protein (OmpA) in both C_8-C_0-50 and C_6-C_2-50 preparations and ~24 kDa band suggestive of EbR with additional band at ~55 kDa suggesting another unidentified contaminant protein.*

This extra protein band was suspected of being another contaminant protein and Western blot was used to probe this (Figure 27A). This was also compared to concentrated samples of EbR isolated by DDM, SMA, A8-35 for screening of these samples also. The expected ~24 kDa his tagged protein band appeared in all samples. There also was the presence of a faint his tagged protein band just below 55 kDa in the A8-35 preparations, showing this was likely to be a dimer also shown by Bratanov et al (2015) (Figure 27B). This band was not observed for the C_6-C_2-50 and C_8-C_0-50 preparations.

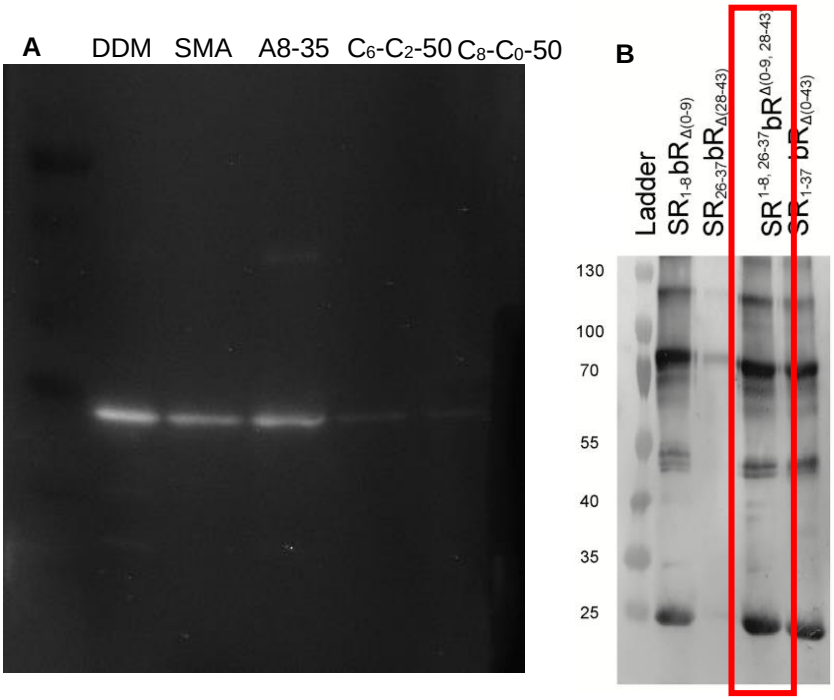

Figure 27 –**Western bot of EbR C_6-C_2-50 as compared with other reconstitution systems** A) Western blot of concentrated DDM, SMA, C_6-C_2-50 and A8-35 and C_8-C_0-50 C) Bratanov et al (2015) western blot of EbR showing presence of a dimer in DDM preparations for comparison B) SDS -PAGE of concentrated SMA and A8-35 EbR

The SDS-PAGE confirmed the purity of the sample, with small amounts of the already identified OmpA. The western blot further confirmed the presence of EbR monomer. The main 66.5 kDa protein series observed in nMS is therefore likely to be an EbR dimer. Given the variability in membrane protein associated ligands and the preservation of these in the gas phase it was thought that comparison with an alternative reconstitution system would prove to be beneficial. It was thought that the A8-35 preparations could provide a good reference as these preparations showed very limited contaminants and share a similarity between C_6-C_2-50 structure. This could facilitate the further characterisation of the C_6-C_2-50 spectrum highlighting patterns in oligomerisation states.

3.1.10 A8-35 EbR nMS

A8-35 has already been shown to solubilise efficiently directly from *E. coli* membranes during the solubilisation assay (Figure 15). It also showed a very pure sample of EbR with no contaminant proteins by SDS-PAGE (Figure 21). The sample was therefore characterised using native mass spectrometry (Figure 28) as a reference for the C_6-C_2-50 isolated samples. In particular, we know A8-35 has EbR dimer present (expected to be 58226.02 Da) seen in the western blot (Figure 27). A similar 68986.02 Da series (series C) was observed within the spectrum, a similar mass to the previously observed in C_6-C_2-50 preparations. The most prominent protein appeared to have a mass of 102,690.96 kDa z=+20, this could be suggestive of EbR trimer (87339.03 Da) especially with a similar mass also observed in the C_6-C_2-50 preparations. Other masses were also observed at 22.8kDa (series B) , 17.6 kDa (Series A) and 95 kDa (series D). These are at a low abundance compared to the main series E and so this is likely to be contaminant proteins.

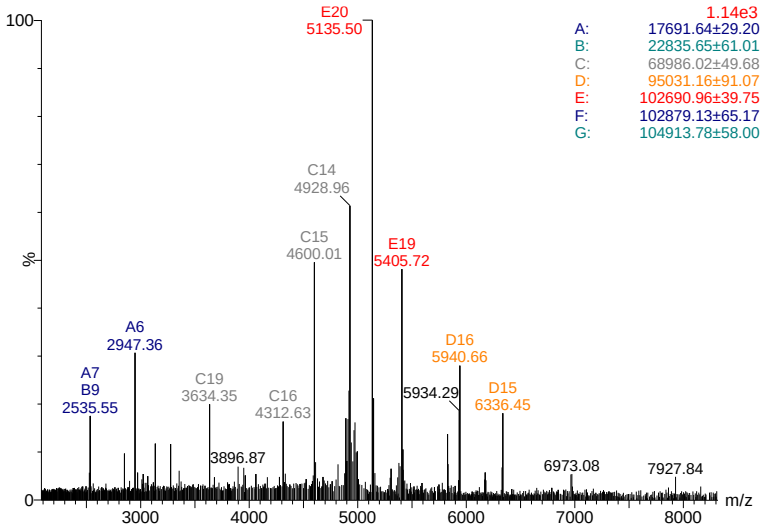

*Figure 28- **EbR A8-35 native mass spectrometry** showing 68.9 kDa and 102.8 kDa protein series seen in C_6- C_0-50 preparations among other molecules that cannot be identified.*

Comparing these A8-35 EbR results with the C_6-C_2-50 isolated EbR showed that there could be a reason for the observed masses seen. There were no unidentified contaminant proteins in either preparations observed using SDS-PAGE (Figure 26) especially as the extra 55 kDa protein band was suggestive of EbR dimer shown in the western blot (Figure 26). Furthermore, the protein masses seen In mass spectrometry from both A8-35 and C_6-C_2-50 preparations seen were similar. This started to suggest another explanation for the resulting masses. To understand the observed the masses they were investigated to determine the deviation in Mass

3.1.11 Polymer isolated EbR nMS mass differences

The masses seen in Figure 24 and Figure 28 are unexpected this is because dimeric EbR mass is 58226.02 Da and trimeric EbR is 87339.03 Da. Another possibility that was discussed the presence of OmpA oligomers, however all of the observed masses (66547.88 Da and 102849.08 Da) are too small to be dimeric OmpA (74401.52 Da) or trimeric OmpA (111602.28 Da). Furthermore, as mentioned previously no contaminant proteins were observed using SDS-PAGE (Figure 25). The presence of large lipid-EbR complex was also discussed but could not explain this either as these lipids would easily dissociate at the HCD cell voltage used. This implies that these protein series seen could be an oligomer of EbR and so these adducts and mass differences were collated together to understand the results observed (Table 5). The similarity in protein masses observed from both the A8-35 and C_6-C_2-50 implied that this was a polymer specific observation. This could imply that the polymer-EbR interaction was retained as opposed to dissociated, as previously assumed. To investigate this the mass deviation was used to calculate the theoretical number of either polymer present within these samples. From

the previous results it was assumed that the mass 66547.88 Da was the accurate mass due to the previously mentioned mass difference after the introduction of HCD cell voltage (suggesting dissociation of an associated molecule) and a similar mass (66639.13 Da) seen in Figure 23. This mass was used to calculate the additional mass from the expected EbR mass. It appeared that the suspected trimer complex had a deviation in mass of 15351.93 Da for A8-35 and 15510.85 Da for C_6-C_2-50. The suspected dimer showed an additional mass of 10760 for A8-35 and 8321.86 for C_6-C_2-50 preparations.

The mass of both polymers has been investigated, suggesting the average A8-35 monomer is 4300 Da (Zoonens & Popot, 2014) and C_6-C_2-50 is 4919 Da (Marconnet et al., 2020). The additional mass was therefore divided by the respective polymer mass to provide a theoretical number of A8-35 or C_6-C_2-50 monomer observed within the preparations.

Table 5 – Additional mass calculated based on expected mass during nMS of A8-35 and C_6-C_2-50 isolated EbR

Molecule	Expected mass	Observed mass	Additional mass	Theoretical polymer No.
A8-35				
EbR Dimer	58226.02	68986.02	10760	2.50
EbR Trimer	87339.03	102,690.96	15351.93	3.57
C_6-C_2-50				
EbR Dimer	58226.02	66547.88	8321.86	1.69
EbR Trimer	87339.03	102,849.88	15510.85	3.15

These data suggest that despite the increased activation both polymers could be retained within the mass spectrometer accounting for the large additional mass seen within each spectrum. This theory could be further validated using other reconstitution systems that could be removed or have a different mode of action and comparing the mass spectrometry results after dissociation of the reconstitution system.

To rule out the possibility of reconstitution system specific adducts and provide further understanding of the EbR complexes. These mass spectra were compared to the other reconstitution systems starting with DDM as it has a known mass of 510 kDa.

3.1.12 nMS of EbR after DDM reconstitution

After optimisation, EbR was purified using DDM and concentrated using a centrifugal filter for mass spectrometry. The previous SDS – PAGE showed EbR was the most abundant protein in the sample (Figure 17), showing its suitability for mass spectrometry. As mentioned previously a monomeric mass of 29113.01 Da and dimeric mass of 58226.02 Da is expected for EbR. After desalting twice the DDM sample was subjected to native mass spectrometry (Figure 29). This spectrum had to be acquired using a desolvation voltage at -125 V to release EbR from the DDM micelle. There was a very clear protein series (series L) displaying a mass of 215.1 kDa. This particular peak could not be identified as it did not fit with the calculated oligomers of EbR although it could be an EbR hexamer (174.6 kDa) with bound molecules of DDM, lipids or a mixture. The same could be said for another series seen with a mass of 161.8kDa, possibly showing a EbR pentamer (145.5 kDa) with additional ligands. Both these series were confidently assigned, However, the overlapping charge states and oligomers as the low m/z (series A and C) were particularly hard to assign, this put into question the validity of the masses seen within this area of the spectrum.

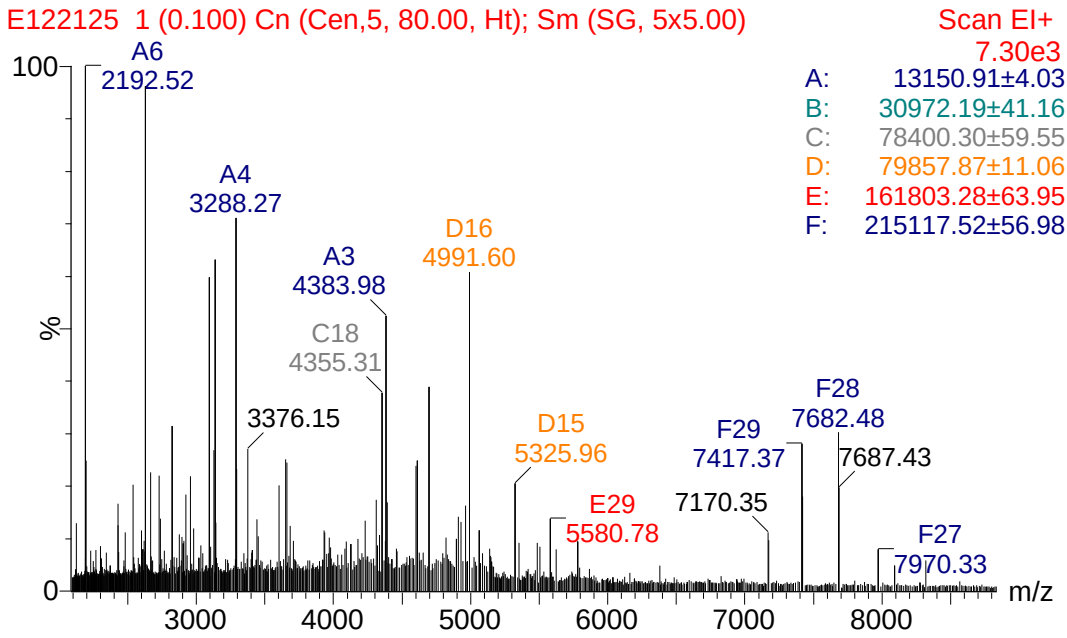

*Figure 29- **Native mass spectrometry of EbR DDM** showing a variety of series and clear protein series at 215.1 kDa.*

As per previous, spectrums the dissociation of associated molecules could clarify what is contained within the sample. The same spectrums were subjected to higher activation with HCD voltages 50 V, 100 V, 150 V with an aim to remove noncovalent interactions. The spectrums were processed and analysed together (Figure 30). It was hoped this dissociation might also dismantle larger unidentified complexes to further identify possible oligomers.

The same main protein series at 161.8 kDa and 215.1 kDa were observed. This gradual increase in voltage showed little to no difference in abundance of these proteins, suggesting they're likely to be contaminant proteins. The same was to be said for the peaks seen in the lower m/z, these peak assignments are still somewhat questionable due to the spacing between them.

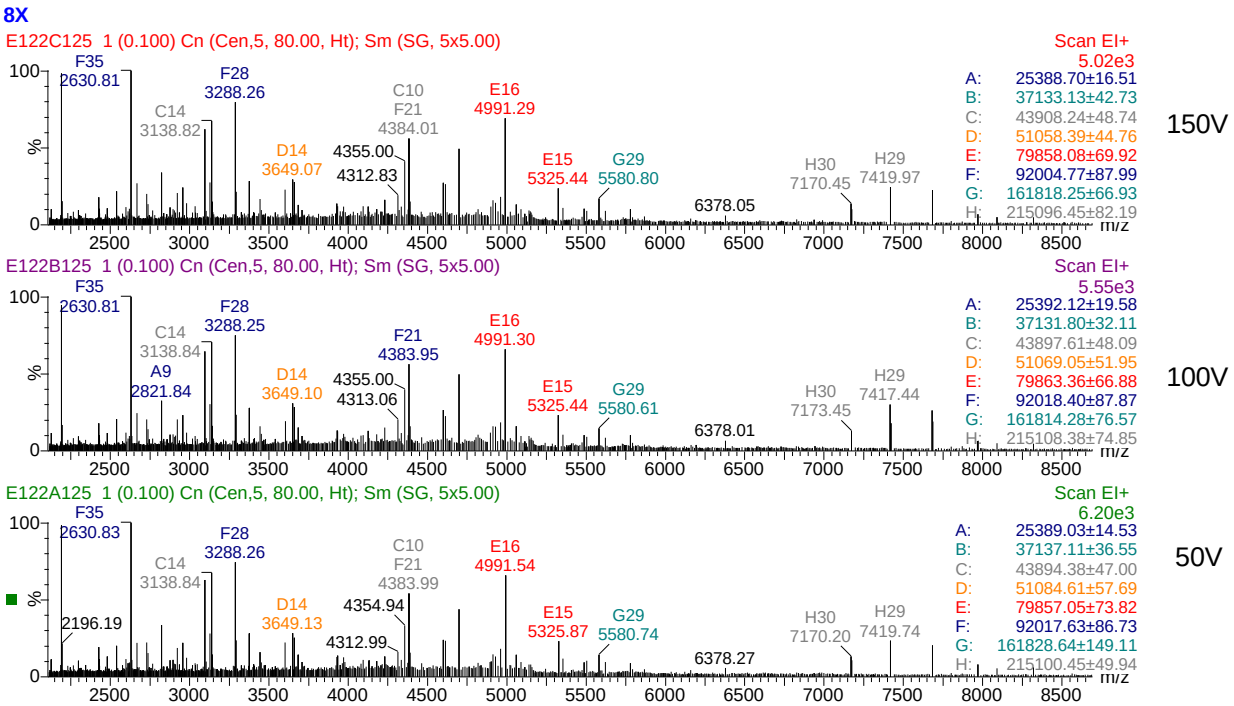

*Figure 30 - **EbR DDM dissociation experiment** with gradual increase of HCD cell activation (voltage) showing confidently assigned 215.1 kDa , 161.8 kDa molecules ; and low m/z peaks assignments that aren't clear*

There could be a variety of reasons for these results. There can be contaminant proteins that ionise well during ESI and show up rather abundantly. Another reason could be the presence of EbR PTMs or tightly associated molecules that are unknown.

This complexity of the samples implies that these preparations, despite showing high abundance of EbR in SDS-PAGE, need further characterisation. Some of the observed

peak assignments, particularly in the low m/z look like artifacts possibly due to the high desolvation voltage leading to unfolding of certain oligomers. Given the remit of this investigation, the next step was compare these results with preparations using a similar reconstitution system that is known to have no contaminants.

3.1.13 G1 OGD EbR nMS

EbR preparations solubilised in G1 OGD previously appeared to be a pure purification evaluated by SDS-PAGE (Figure 18). This allowed them to be validated for mass spectrometry analysis. The sample was prepared by desalting twice into 200 mM ammonium acetate, supplemented with 0.07 % G1 OGD detergent (2 x cac). To achieve displacement of protein from the G1 micelle for mass spectrometry a desolation voltage of -15 V was required, significantly lower than DDM (Figure 31). The main spectrum (Figure 31 bottom panel) showed a set of series with a variety of charge states. The most abundant peak series appeared to be series A with a mass of 27.1 kDa which is too small to be an EbR monomer. Other masses included 37.4 kDa and 49.2 kDa of which are too large to be an EbR monomer but too small to be an EbR dimer. There also appears to be associated molecules with each protein series (series A2, B2 and C2). Given the appearance of possible associated ligands it was wondered if these molecules were dissociatable, this would differentiate between a non-covalently bound molecules and post translational modification. The signal-to-noise ratio when using the HCD cell was too low to comprehend the spectrum. With the inability to acquire a spectrum using the HCD Cell, the in-source collision induced dissociation (ISCID) was used in an attempt to achieve a similar affect (Figure 31 top panel). This also displayed a similar issue creating an unreadable spectrum and so smaller increments of 1 V had to be used. this panel shows this method at 3 V, the maximum that could be used before an unclear spectrum was observed. It was unclear if the spectrums we showed dissociation of the additional molecules but it did however start to show series E with a mass of 164 kDa.

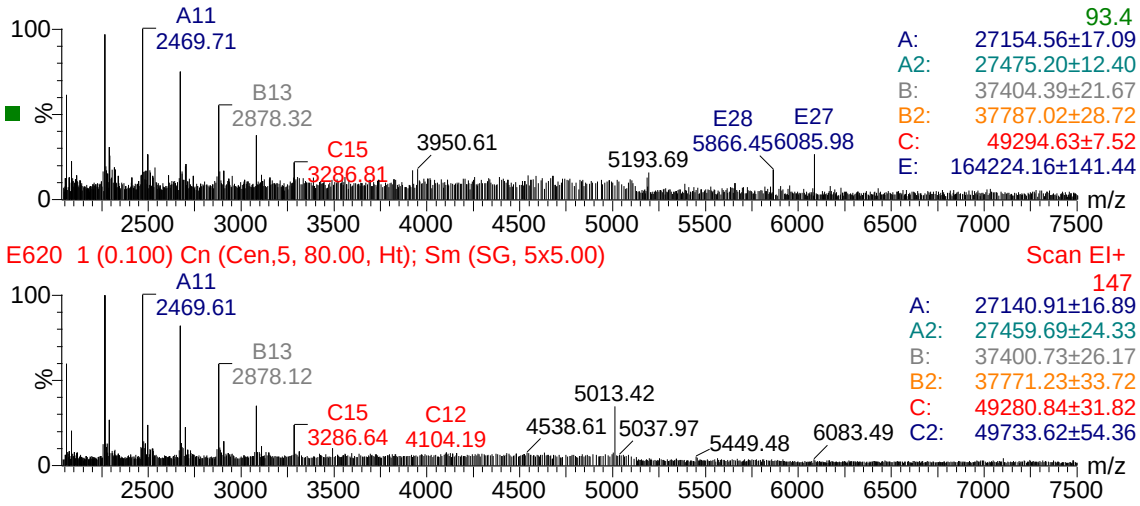

*Figure 31 - **G1 OGD native mass spectrometry** showing a variety of masses that can not be attributed to monomer EbR but also showing a pentameric EbR protein series after light activation using ISCID*

The main protein series masses (Figure 31 bottom panel) were the used to calculate the mass of the associated ligands (Table 7). This was done finding the difference between the main series mass and adduct mass. This found three different associated ligand masses with series A being association with an additional molecules of 305.13 Da, adduct B was associated a molecule of 382.63 Da and adduct C was associated with a molecules of 452.78 Da.

Table 6 – G1 OGD EbR nMS Adduct masses calculated from observed masses

Adduct	Observed protein ion mass	Observed associated adduct mass	Associated ligand mass
A	27140.91	27459.69	305.13
B	37400.73	37771.23	382.63
C	49280.84	49733.62	452.78

These results collectively show three different mass of associated molecules. It could not be verified if these were non-covalent modifications using mass spectrometry and so it is not clear if these are lipids. We can discern from these masses that they are not G1 OGD detergent associated with the main ions as G1 OGD has a mass of 408.3 Da,

measured by mass spectrometry (Urner et al., 2020). However, the variability in mass of the associated molecules could represent the variability of the *E. coli* membrane lipids composition.

The low protein masses show a few things about these samples. With there being little information on molecular interactions of EbR it is not clear how tightly it binds to its surrounding molecules. Furthermore, with the alterations required for the expression of EbR in *E. coli* it is not clear if any other unknown EbR PTMs occur that do insert into the membrane and therefore can be purified using IMAC. Despite this the Q-Exactive allowed characterisation of the samples and associated molecules. Molecules associated with these samples found in the MS do imply a retention of lipids or ligands and can be used to compare to HbR samples.

3.2 HbR

3.2.1 Enrichment and characterisation of Purple membrane (PM)

H. salinarium harbours our protein of interest in in dense patches of HbR known as purple membrane and has been used extensively for HbR production. After purchasing of freeze-dried *H. salinarium (ATCC)*, it was activated as described by the company's instructions. *H. salinarium* was grown in specific media described in the materials and methods. The purple membranes were separated from the red membrane of *H. salinarium* using sucrose gradient after cell disruption. As the name suggests, isolated purple membrane should appear as vibrant purple in the sucrose gradient and demonstrate the characteristic 560 nm absorbance peak. During enrichment, confirmatory characterisation was carried out to ensure these characteristics were as expected. Using the sucrose gradient yielded the characteristic purple band (Figure 32A). To further characterise the HbR sample it was analysed using SDS-PAGE. The observed size ~24 kDa was confirmed, with little to no contamination from other proteins (Figure 32B). Despite this not being the molecular weight of HbR, it is characteristic of many membrane proteins to show perturbed molecular weight using SDS-PAGE. Furthermore, upon investigating HbR within the literature it was found that Stauffer et al (2020) demonstrated the similar result using SDS-PAGE, validating our results.

To further confirm the identity of the protein, the absorbance of the purified sample was tested for absorbance at 560 nm, measuring three times to ensure accuracy. HbR is known to absorb specifically at this wavelength. Our analysis demonstrated the characteristic 560 nm peak in the absorbance spectrum which confirmed the presence of the protein of interest (Figure 32C).

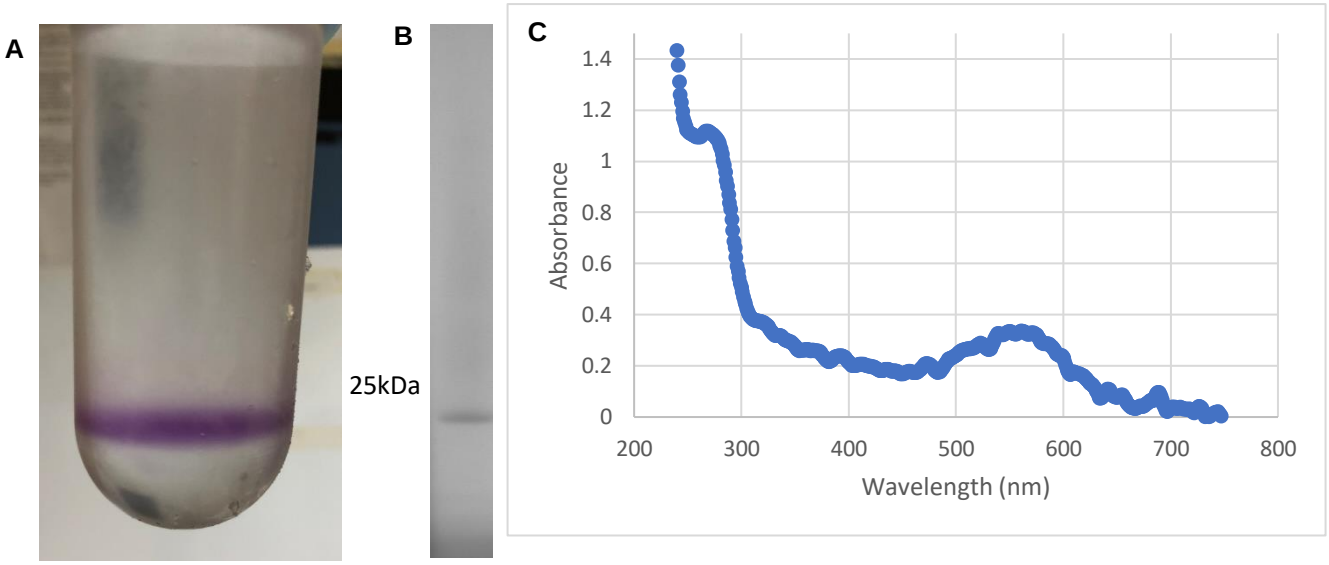

Figure 32 – ***HbR enrichment and SDS-PAGE****. A) PM sucrose gradient following* Shiu et al (2013) *B) coomasie stained SDS-PAGE of PM demonstrating the expected characteristics of isolated PM of 560 nm peak after three repeat measurements and showing to be pure*

The clear absorbance at 560 nm demonstrates the presence of the retinal chromophore of HbR and that this process shows no perturbations in secondary structure as consistent with the literature. These data confirmed the enrichment of the PM from *H. salinarium*.

3.2.2 HbR Solubilisation

After purification, PM was kept in DDH_2O and was pelleted before resuspending in the desired solubilisation buffer. To first facilitate the reconstitution of HbR a solubilisation assay was carried out at the same detergent and polymer concentrations used to solubilise EbR (Figure 15). This was run on SDS-PAGE to determine the amount of solubilised material (S) compared to total HbR present (T). The results show initially problems with blurred bands within the gel (not shown). It was determined to be due to an issue with acrylamide percentage being too low. Instead, the samples were then run on a prepared 15 % gel (Figure 33). This showed extremely faint protein bands mostly in the soluble fractions. This was suggestive of very minute solubilisation efficiency from the PM, even with the high concentrations of solubilisation systems used. Because of the lack of clarity of this gel, it was thought a more sensitive technique would be more accurate at determining solubilisation efficiency and spectrophotometric readings at 560nm were taken pre and post solubilisation.

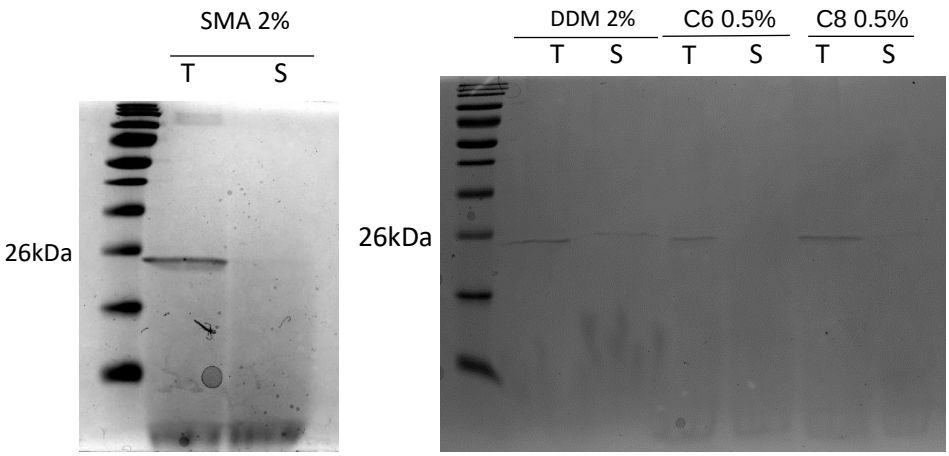

Figure 33- ***SDS-PAGE comparing total material and solubilised HbR*** *with faint protein bands in DDM, SMA and C_8-C_0-50 solubilised preparations*

The solubilisations were carried out as before. The baseline absorbance of HbR, was measured at 560 nm using a nanodrop spectrophotometer. The absorbance readings were taken pre solubilisation (total) to demonstrate total material presence. The A560 measurement was repeated after centrifugation which removed insoluble material. This would represent the total HbR solubilised (soluble). To present this, the A560 of the total material was compared to A560 of solubilised material after triplicate measurements of each samplel (Figure 34). This demonstrated the ability of each agent to solubilise HbR from the PM and also allowed the identification of interference within the spectra. Much interference was seen in this graph suggesting solubilised HbR might be overestimated. This was mostly present within DDM solubilised sample, presumably due to scattering of light by DDM micelles within the solution. Apart from DDM, all absorbances demonstrated a micromolar concentration in each sample explaining the minimal protein seen in the SDS-PAGE gel. Despite amphipol (A8-35, C_6-C_2-50 and C_8-C_0-50) preparations demonstrating minimal yield, the wavelength spectrum did not show a clear 560nm peak seen in other preparations (not shown). SMA showed a minimal yield also, however it appeared that the spectrum was clear enough to investigate this further and optimise the solubilisation.

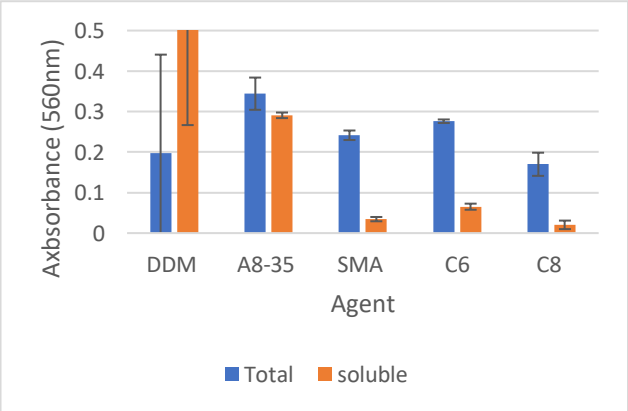

Figure 34- ***Solubilisation of HbR directly from the PM*** *showing micromolar yields HbR from solubilisation and interference of DDM with spectrophotometric reading all readings were carried out in triplicates to ensure accuracy (n=9). Further characterisation of these samples is required*

The solubilisation efficiency of these agents were thought not to be accurate from this graph and so had to be optimised/investigated individually to properly understand the efficiency seen. This would then allow optimisation of solubilisation environment and buffer composition for each reconstitution method.

3.2.3 G1 OGD: solubilisation directly from the PM

The G1 OGD was prepared much the same as the previous reconstitutions. G1 OGD was added at to the purple membranes at 1 % and incubated overnight at 4 °C. The pre and post solubilisation samples were run on SDS-PAGE on a precast gel (Figure 35A). This showed that the solubilised HbR (S) band had a similar thickness to the total material (T), implying that most if not all HbR had been solubilised. There also appeared to be larger oligomers in the solubilised HbR with faint protein bands in the soluble fraction at ~55 kDa. To further quantify the solubilisation efficiency of G1 OGD from the PM the solubilisation was quantified using absorbance comparing total material to solubilised material present after centrifugation (Figure 35B). A280 was used due to alterations in baseline absorbance explained later in these results. This solubilisation showed a high yield of ~ 59 % compared to that of other reconstitution systems.

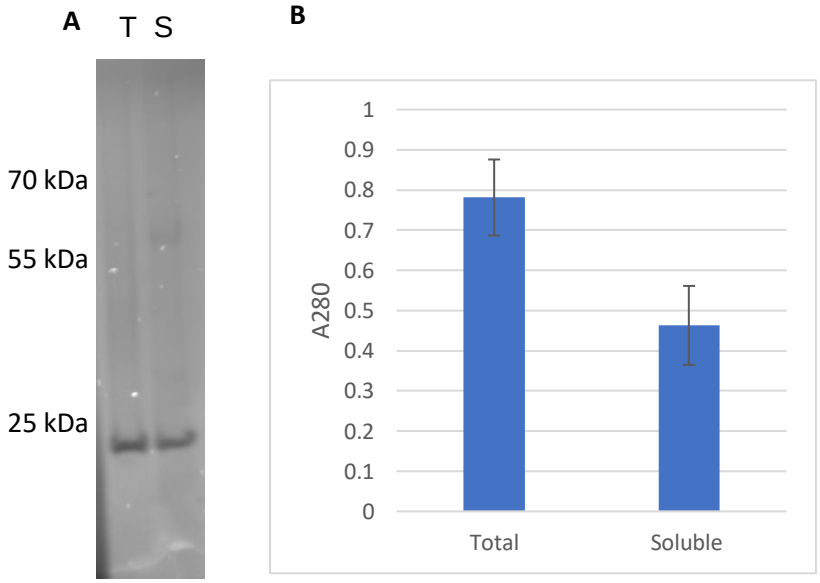

Figure 35- ***Solubilisation of HbR using G1 at 1 %****.A) Showing possible higher oligomerisation states present and similar concentration to total material B) spectrophotometric quantification of G1 HbR showing total material present against solubilised material with ~59 % efficiency*

These results could not inform a size of either protein bands present due to the alterations in MP migration through SDS gels. However, comparison of Figure 34A with Figure 30B shows a similar size and purity to that of the stock PM sample within the G1 HbR preparations. The high solubilisation yield of HbR G1 OGD mean that this modular detergent was more efficient to solubilise native HbR than EbR. Probably due to the

nature of lipids surrounding the protein. The sample was validated for further for nMS characterisation. And so was prepared in 200 mM Ammonium acetate.

3.2.4 nMS of G1 HbR

During nMs HbR-G1 showed an abundant 8+ charge state suggesting that the this was near native fold during nMS (Figure 36). The modular detergent allowed well resolved spectra, providing abundant peaks suggestive of monomeric HbR at a mass of 27.08 kDa, this would imply the release of HbR from the modular detergent micelle. This release of HbR from the micelle was able to be carried out at desolvation voltage of -20 V. Adducts were also detected with an additional mass around 906.9 Da, suggestive of PGP-Me. To ensure that the adducts showed were not due to presence of covalent modifications, the HCD cell was used to dissociate the adduct from the suspected lipid-protein complex. The suspected lipid adducts dissociated upon application of HCD cell voltage at 200 V confirming the presence of HbR-PGP-Me complex (Figure 34).

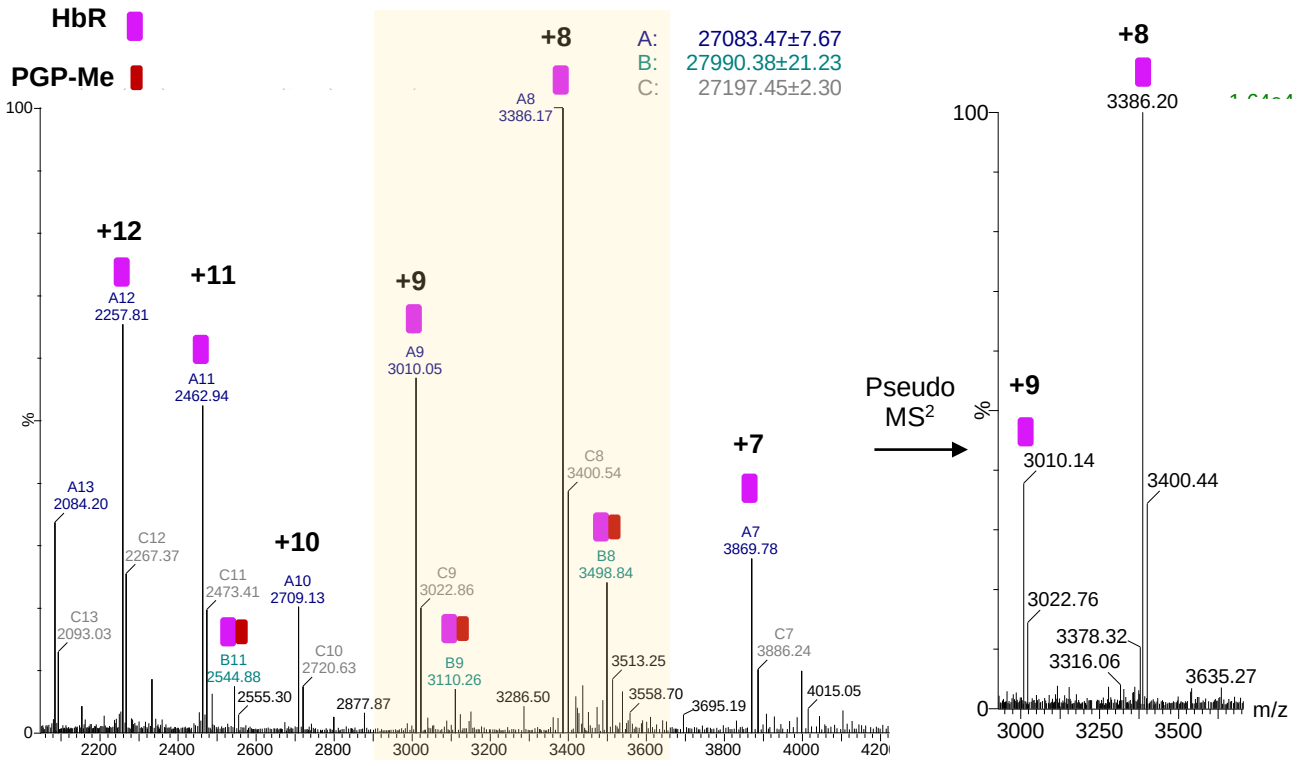

*Figure 36– **HbR-G1 mass spectrum with suspected lipid and dissociated lipid adducts** using the HCD cell for an activation experiment, confirming the presence of a lipid with mass.*

These masses are extremely similar to some of the cited masses and so have been compared to the cited masses to highlight variations in mass (Table 7). The sequence mass of bacteriorhodopsin (28000 Da) was also compared to the cited protein mass

(27000 Da)(Campuzano et al, 2016). PGP-Me structural mass is 901.2 but has been previously cited as 900.65 using lipid mass spectrometry (Angelini et al., 2010).

Table 7 - Mass of molecules found in HbR G1 OGD preparations compared to the literature

Molecule	Expected mass (Da)	Observed mass (Da)	Cited Mass (Da)
Bacteriorhodopsin	28000 (sequence derived)	27083	27000
PGP-Me	901.2	906.9	900.65

From this it is clear that the sequence mass of bacteriorhodopsin, also called known as bacterioopsin, is larger than the cited protein mass. This is due to maturation process of HbR before insertion into the purple membrane, where the retinal group is included into the correctly folded HbR (Hoi et al., 2021). This data confirms the presence of HbR within the G1 OGD micelles and the presence of HbR-PGP-Me complex in these samples.

3.2.5 SMA: solubilisation directly from the PM

SMA appeared to show minimal solubilisation in both the SDS-PAGE and spectrophotometry. This meant that further optimisation of the solubilisation directly from the PM could be possible. Its known that there are a variety of factors that could affect SMAs ability to solubilise HbR. The first two factors to be tested were incubation time and temperature (Figure 37). Previous solubilisations were carried out at 25 °C for 2 hours, so it was decided to compare this to 4 °C and 37 °C and taking samples at 1 hr, 2 hrs, 3 hrs, and 16 hrs. Solubilisation efficiency was investigated using A560. The A560 pre solubilisation was compared for the time periods mentioned (after centrifugation). The values were used to calculate a percentage solubilisation. This revealed that there appeared to be little difference in time related solubilisation percentage between 1 ,2 and 3 hours when incubated at room temperature. It also illustrated that at 4°C, there was a significant jump in solubilisation percentage from 0 % to ~35 % between hour 1 and 2 which was also higher than 3 hours (17%) and 16 hour (0%) incubation. Given that the 2 hours at 4 °C incubation also showed a higher yield than the other incubation temperatures and times, this appeared to be the optimal environment for SMA

solubilisation directly from the PM. These were implemented as the conditions for HbR solubilisation.

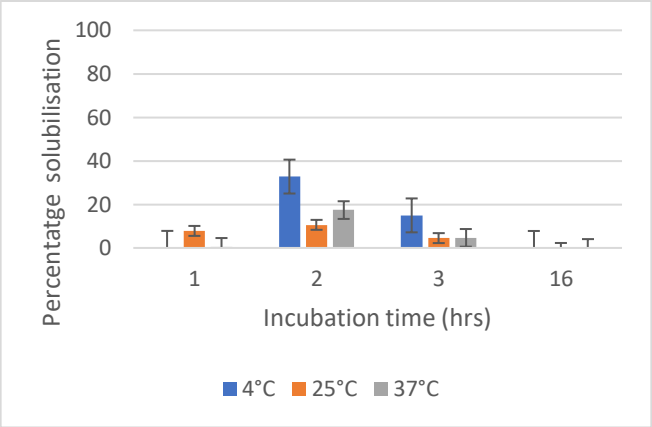

Figure 37- ***Temperature dependent solubilisation of HbR using SMA****. Reconstitution of HbR from the PM using SMA and varying temperatures and times Showing solubilisation for 2 hours at 4 °C was optimal for SMA solubilisation.*

Given these results further optimisation of HbR solubilisation was tested to evaluate if the if a higher solubilised HbR yield was possible, as it remained between 20 % and 30 % solubilisation. Another key factor was the concentration of PM used for solubilisation. The previous solubilisation were executed at PM concentration of 1mg/ml and it was not clear if higher concentrations could yield a higher solubilisation efficiency. Therefore, the dependency of SMA on the PM concentration was investigated (Figure 38). PM concentrations 0.5, 1, 2, 3, 4 and 5 mg/ml were solubilised for 2 hours at 4°C. This demonstrated that a PM concentration of 2 mg/ml would be optimal.

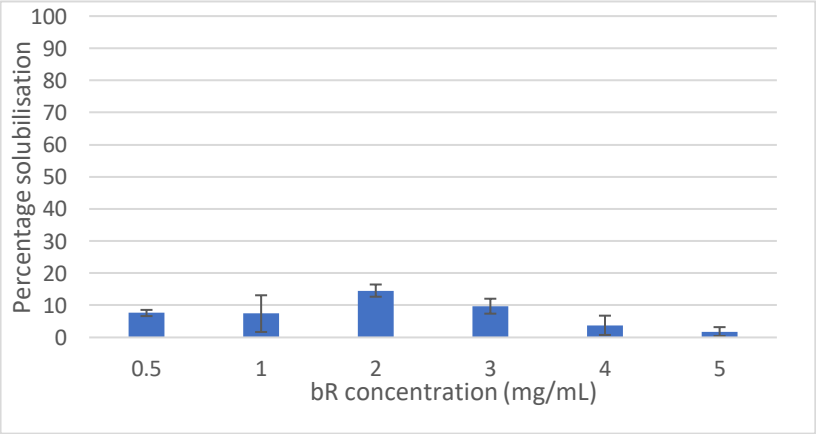

Figure 38- ***PM concentration dependant solubilisation of HbR*** *using SMA with 2 mg/ml providing an optimal polymer to protein ratio for most efficient reconstitution (measurements were taken in triplicate for accuracy)*

Despite these positive results, the yield of SMA-HbR was still considerably low. Given that the PM is grown in the halophilic environment at 4 M NaCl, the lipids associated with HbR could be more rigid in lower salt concentrations. This could imply that increasing NaCl factor could improve solubilisation efficiency. Solubilisation buffers were then prepared with NaCl concentrations 250 mM, 500 mM, 750 mM, 1 M, 2 M, 3 M and 4 M. The PM samples were pelleted and resuspended in the respective buffer at 2 mg/ml and

incubated at 4 °C for 2 hours. The samples were also quantified using A560 comparing the total material (T) to the solubilised material (S) of the varying NaCl concentrations (Figure 39). It was unclear from this investigating of the effect of varying NaCl concentration on the efficiency. However, it appeared to suggest 4 M NaCl would be optimum for solubilisation from the purple membranes. Given the still low efficiency of solubilisation, it was thought that using a higher volume for solubilisation would balance out the low solubilisation efficiency and yield enough solubilised protein for characterisation.

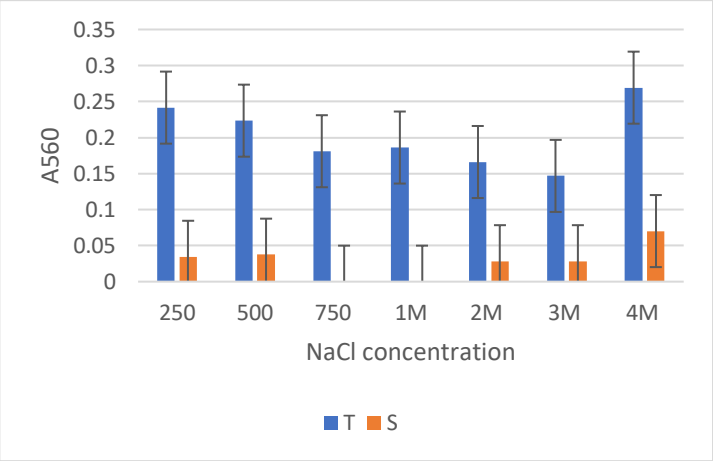

Figure 39- ***NaCl Dependant solubilisation of HbR using SMA.*** *4 M appeared to show increased efficiency for solubilisation compared to other concentrations and was trialled in larger volumes of PM.*

Because of this, a larger volume solubilisation was then carried out at 4 M NaCl with the intention of concentrating the sample post solubilisation. Similarly, the sample was prepared in 4 M NaCl and incubated at 4 °C for 2 hours and quantified using A560 comparing the total material and material left post centrifugation, the soluble material (Figure 40). Surprisingly this showed a significant improvement over previous solubilisation demonstrating ~73 % solubilisation.

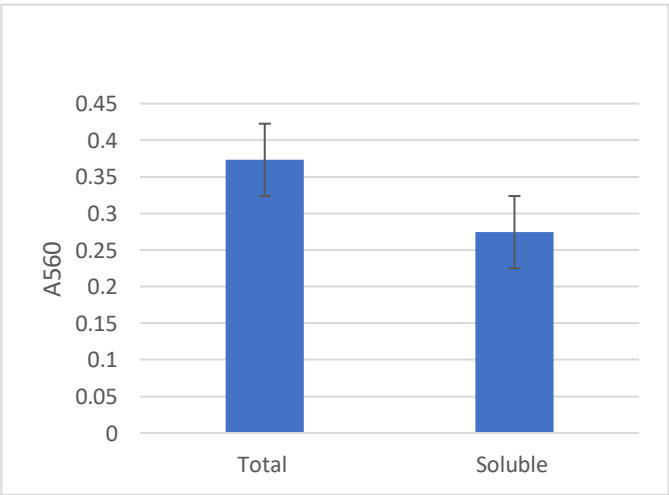

Figure 40– ***SMA Solubilisation efficiency of HbR directly from the PM at 4 M NaCl concentration.*** *Yield of solubilised HbR from the purple membrane after solubilisation in 4 M NaCl showing an ~73 % efficiency.*

3.2.6 nMS of SMA HbR

After SMA reconstitution, the HbR-SMA sample was then buffer exchanged into ammonium acetate as done with previous samples and applied to native MS (Figure 41). Previously, nMS of HbR reconstituted DMPC liposomes achieved a release of monomeric bR from the SMALP in the mass spectrometer at a in source trapping voltage (desolvation voltage) of -80 V (Hoi et al., 2021). This voltage was used for this sample also and appeared to show an abundant 69 kDa complex. This could however suggest lipids paired with a HbR dimer (54.16 kDa).

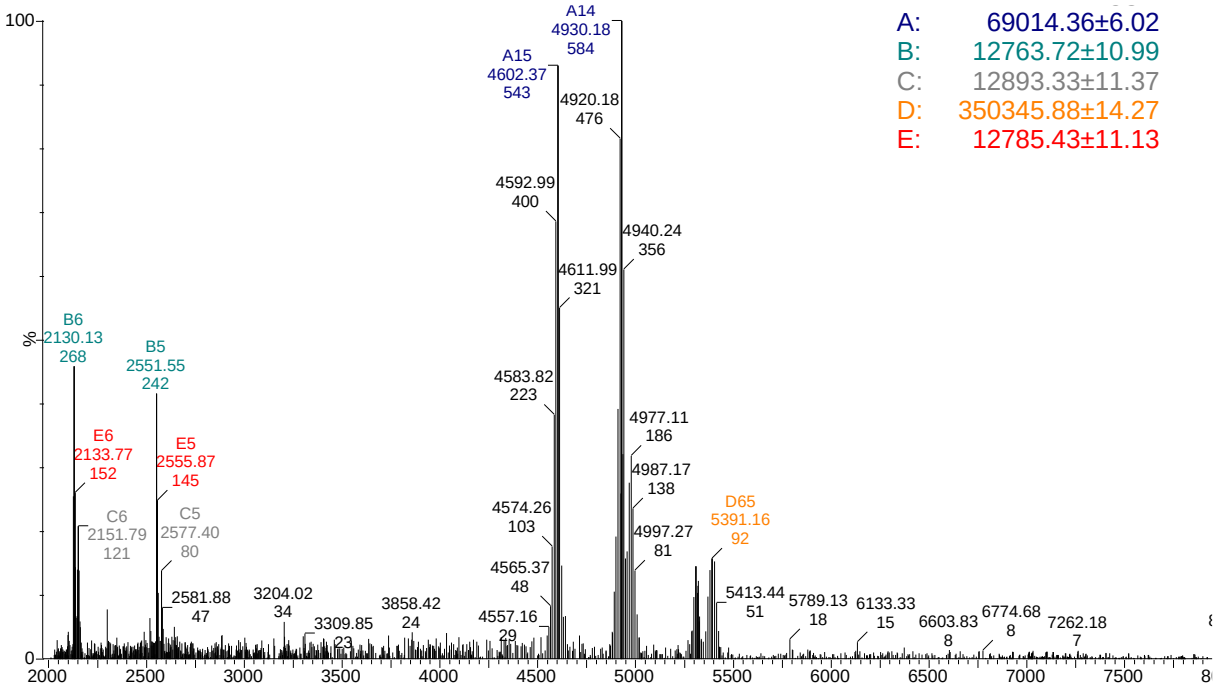

Figure 41 - ***nMS of SMA HbR at -80 V desolvation.***

However, increasing the activation of this sample did not yield clearer spectrums or changes in mass. Suggesting that HbR was still contained within the SMALP (not shown). This would be further explained with the smaller series seen at 12.7 kDa, this could be a SMALP containing lipids. It was thought that chemical dissociation of the SMA from the protein, as opposed to the electromagnetic forces from the Q-Exactive, would provide a better understanding of the contents of the SMALP.

Due to the HbR appearing to be contained within the SMALP MS lipid adducts could not be seen within the spectrum. This meant the application of 7.9 V for the injection flatapole was not enough to eject the HbR from the SMALP. It was thought chemical dissociation of the SMALP would facilitate characterisation of the internal environment of the

HbRSMALP. SMALPs are known to be sensitive to divalent cations, particularly magnesium, causing dissociation of SMALP at micromolar concentrations. It was thought that the use of divalent cations would be beneficial for this application. The low concentrations wouldn't cause the presence of salt adducts within the spectrum. The dissociation was therefore achieved by the addition of magnesium acetate (MgAc) and ammonium acetate buffer for a final concentration of 10 µM MgAc, just before mass spectrometry (Figure 42 top panel) the main spectrum was compared to the raw spectrum during analysis to ensure processing parameters captures all associated peaks (Figure 42 bottom panel). It was hoped that the HbR would retain enough stability and lipids to acquire a mass spectrum. Dissociation of the SMA from the HbR-SMA complex was found to be somewhat sensitive to certain parameters during mass spectrometry. However, setting the desolvation voltage as -80 V was more than sufficient at dismantling the complex seen. This showed an interesting pattern with it suggesting the loss of many molecules from the SMALPs indicated by the products ions being lower on the m/z scale than the parent ions. However, this made this spectrum hard to interpret as there was a variety of molecules present most likely to be polymer and other associated adducts. It appeared there could possibly be a proteins series with protein-lipid complexes at m/z 3000-4000 with a mass of 29.7 kDa (series C), although it was not clear form this spectrum due to the variety of molecules present. To facilitate the characterisation of this sample during chemical dissociation the inclusion of precursor ion selection and HCD dissociation experiment would allow further analysis of each molecule present.

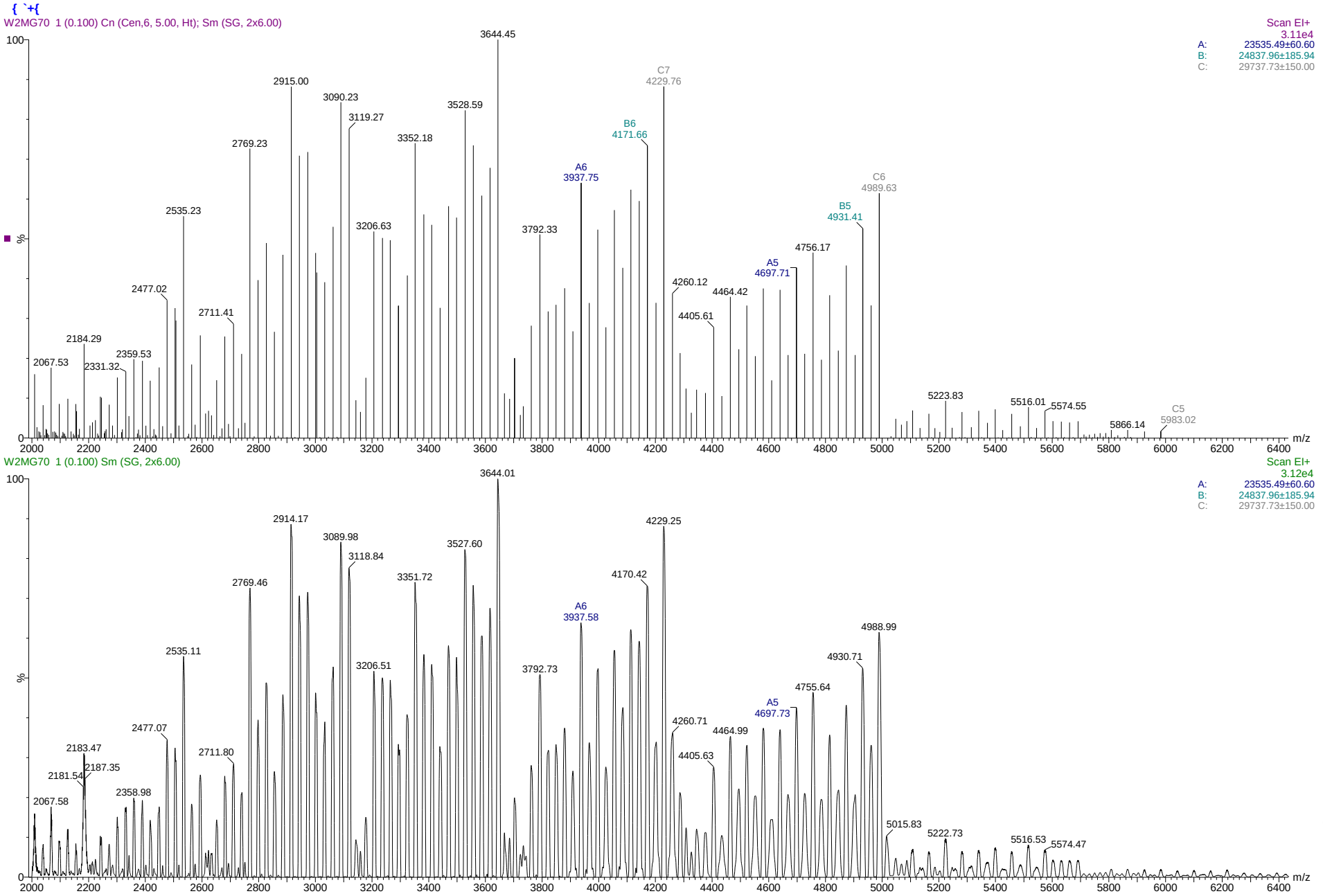

Figure 42 - ***Mass spectrometry of HbR-SMALP complex after SMA dissociation*** *mass spectra were acquired after addition to ammounium acetate, Magnesium acetate solution to chemically weaken the SMALP complex*

It was clear that there were a variety of protein series and associated ligands possibly overlapping in the previous spectrum. Given the normal m/z of HbR sits at around 3000 it was decided using the quadropole at an isolation window of 3000-3700 m/z with an aim to harbour monomeric HbR seen in other spectrums in Hoi et al (2021) and Figure 34. To further characterise the proteins present an activation experiment was carried out by increasing the HCD cell voltage from 50 V – 150 V (Figure 43). To facilitate peak assignment in the masslynx software, the main peak series were focused on.

The main spectrum without any dissociation capabilities applied (Figure 43A), appeared to have two main protein series with masses of 100.6 kDa (series B) and 77.9 kDa (Series A). Given that SMA is known to retain lipids, the assumed retention of lipids could explain the size of the molecules seen in this spectrum. Therefore, further dissociation is required to understand the protein within the SMA. When 50 V HCD cell is applied (Figure 43B), there was an increase in protein series B with a mass of 83.3 kDa, this and protein series C (93.3 kDa). Trimeric HbR has a mass of 81.2 kDa as determined using previous spectrum (Figure 34), so to further characterise if this was trimeric EbR, increasing the HCD cell voltage should be able to dismantle the trimer. Further increasing the HCD cell voltage (Figure 43C) started to show an increase in protein series with a mass of 55.3 kDa. With the mass of dimeric HbR being 54.1 kDa, it was clear that it was possible this comes with lipids from the trimer hence the increase in mass. Furthermore, at 150 V (Figure 43D) there was significant increase in abundance of peaks with masses of 27.9 kDa. This would signifiy the monomer HbR with PGP-Me seen previously within the G1 OGD HbR preparations (Figure 36).

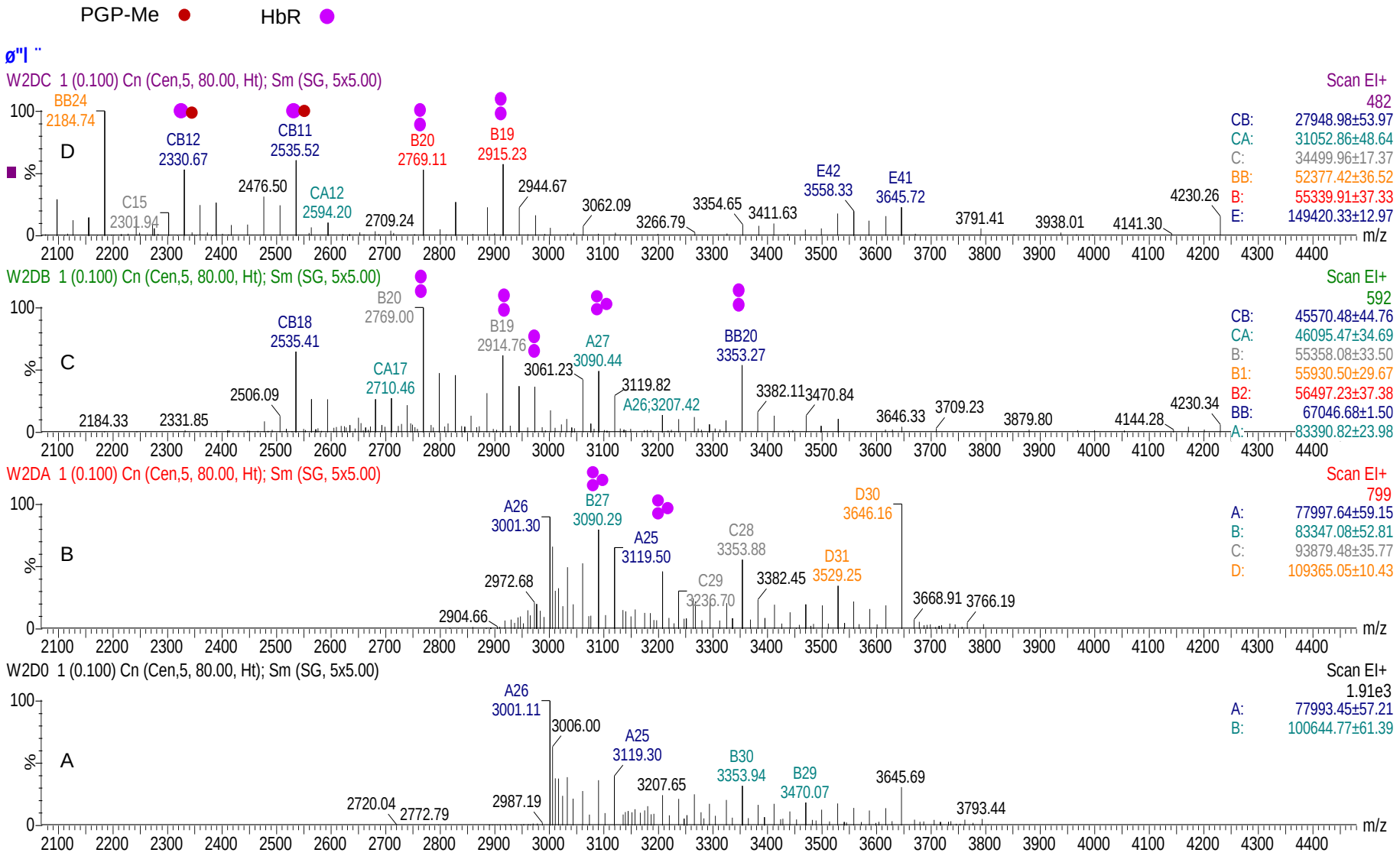

Figure 43 - ***Characterisation of SMA HbR after chemical dissociation*** *showing the dismantling of HbR trimer from trimer to monomer and their associated lipids*

These results not only demonstrate the complexity of the SMALP but also demonstrate another workflow that could be of benefit. By chemically dissociating the SMA from the HbR-SMALP, it allowed gradual dismantling of the HbR trimer to probe its associated molecules and masses.

3.2.7 Spectrophotometry of solubilised HbR

After successful solubilisations with SMA and G1 OGD of HbR, solubilised sample were spectrophometrically characteriseed to determine if there were any structural changes (Figure 44). The centre of the 560 nm peak were indicated by the blue lines. The left line is the centre of the G1 OGD HbR line showing a lower absorbance peak than both the native membranes and SMA solubilised membranes. This means that as compared with the native membrane's absorbance spectrum, it was determined that the OGD G1 cause alterations in HbR secondary structure demonstrated by the shift in absorbance from the expected 560 nm. However, SMA appeared to show no alteration in absorbance and was therfore selected for further structural studies.

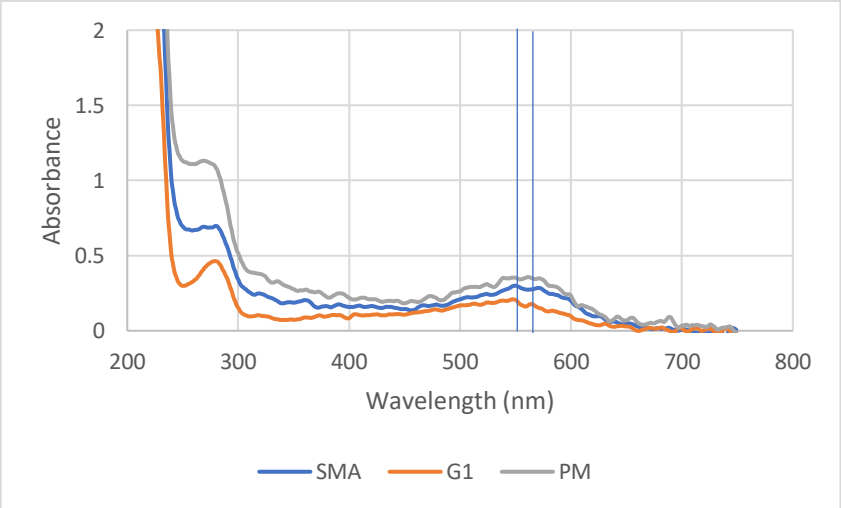

Figure 44 - ***Absorbance spectrum of HbR solubilised by G1.*** *Lines represent the centre of each peak and shows an alteration of HbR baseline absorbance of 560 nm to 547 nm as compared with the PM and SMA preparation suggesting alterations in secondary structure caused by G1.*

Given these results, mass photometry was employed to further paint a picture as to the oligomerisation state of both the G1 and SMA HbR preparations.

3.2.8 Mass photometry

Mass photometry was employed to determine the oligomerisation states of HbR within the sample associated with the reconstitution systems. It was applied to HbR SMA and HbR -G1 OGD, this is specifically due to their purity shown in SDS-PAGE. Possible higher oligomerisation state present were also observed in HbR -G1 OGD preparations

using SDS-PAGE (Figure 35A). Also given SMA flexibility to the 4M salt it was suggestive that this would be a stable construct when mixed with the PBS which is used as the buffer for calibration and so is also required for an accurate mass measurement.

3.2.8.1 SMA HbR

Initial mass photometry of HbR in SMA appeared to be primarily dimeric HbR (Figure 45). However, the landing event counts were low. To be statistically significant the counts preferentially would be better at ~1000. To improve the landing event count it was decided a high concentration should be used for mass photometry.

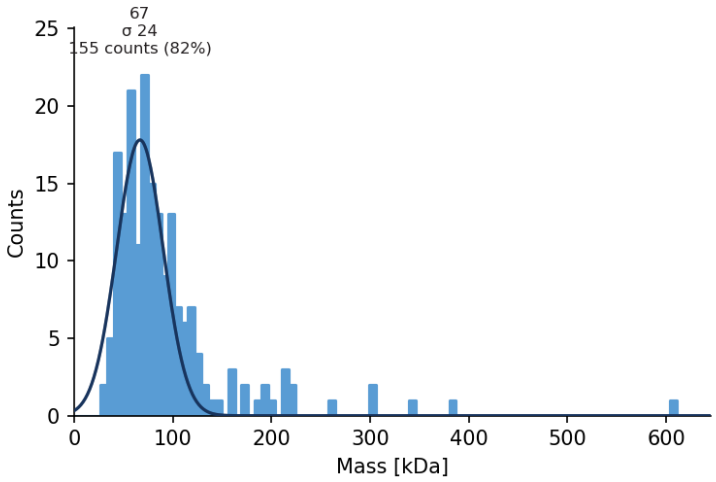

*Figure 45- **SMA HbR mass photometry** suggestive of dimer with associated molecules*

In repeating the experiment with a higher concentration of HbR-SMA, the mass detected from using mass photometry increased 122 kDa and 113 kDa (Figure 46). These distributions also did not show a typical gaussian style distribution. This does appear to be due to an abundant and specific mass that is present. This however does also put into question the data we're seeing and the interpretation of this mass from the software used.

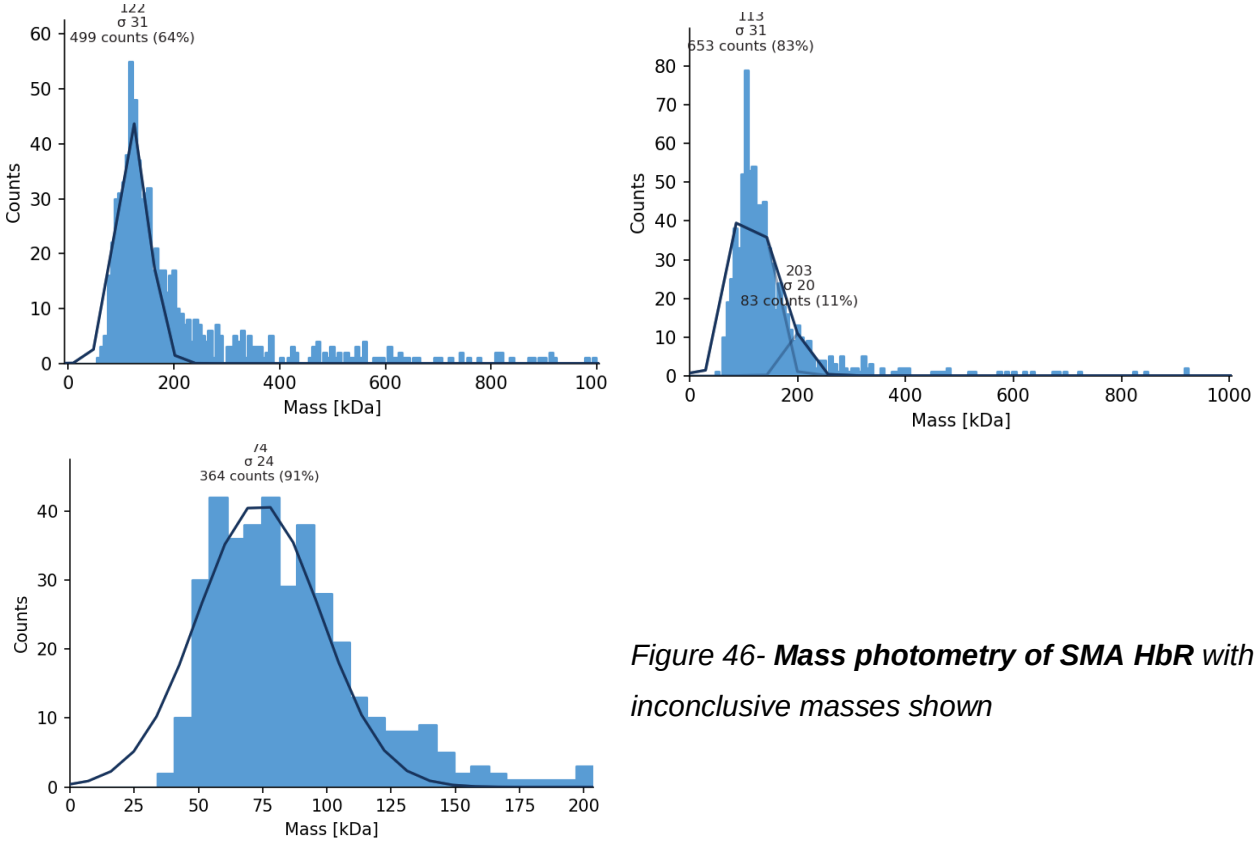

*Figure 46- **Mass photometry of SMA HbR** with inconclusive masses shown*

It was decided a comparison reconstitution could provide some answers to this mass variability seen.

3.2.8.2 G1 OGD HbR

Mass photometry of HbR in G1 OGD was carried as with SMA. This show showed a mass that is somewhat unexpected at around 350-360 kDa (Figure 47). This was thought to be attributed to the G1 being sensitive to the buffer composition as the measurements are carried out in PBS. Because of this, the measurements were then repeated in the solubilisation buffer used for G1 OGD.

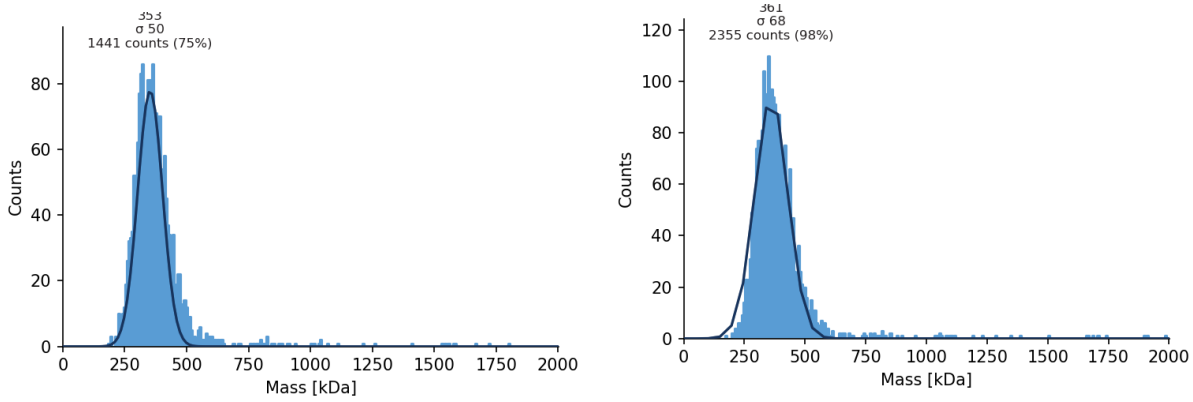

*Figure 47 - **Mass photometry of SMA HbR with increased sample concentration***

A distribution was apparent when using the G1 OGD buffer, showing a pattern corresponding to a trimeric and pentameric HbR (Figure 48). However, similarly to SMA HbR the landing counts were low. A similar approach to SMA HbR was taken to increase the significance of the reading.

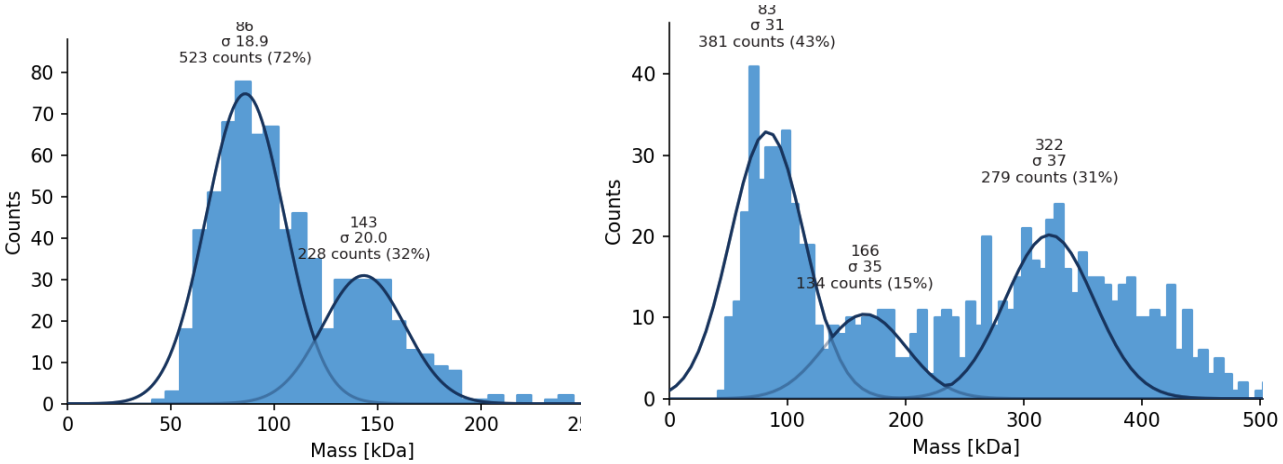

Figure 48- ***Mass photometry of HbR in G1 OGD solubilisation buffer***

The same pattern was observed using this dilution, an increased count came with an increased mass (Figure 49). This reading however, presented the same pattern of distribution as the lower concentrations.

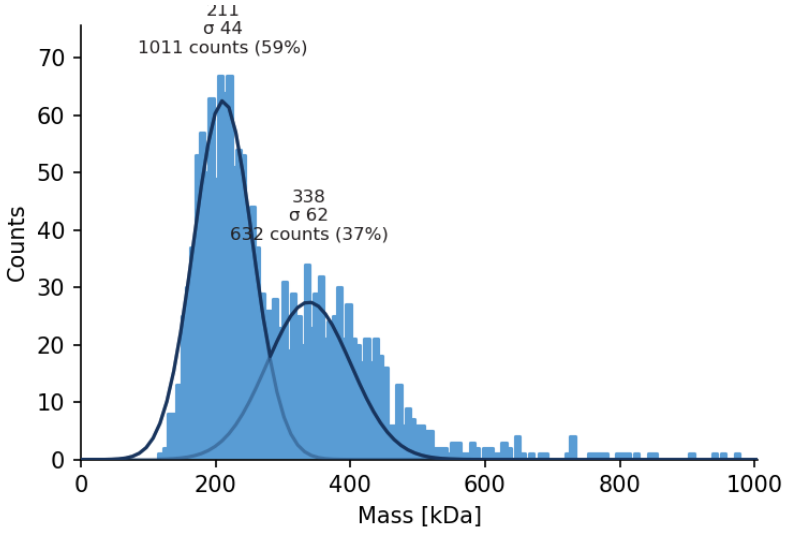

Figure 49- ***Mass photometry of G1 HbR increased sample concentration***

Because of the lack of reproducibility in mass and statistical significance, no numerical information can be taken from these figures. However, this does demonstrate some problems with the application of mass photometry to investigate membrane protein sample heterogeneity.

Chapter 4: Discussion

4.1 Protein expression and purification

4.1.1 EbR is stably expressed

Molecular biology investigation demonstrated that the recombinant EbR DNA insert within the given plasmid was as expected. The insert sequence was further confirmed using DNA sequencing. This data collectively demonstrated that the plasmid to be used contained the correct DNA sequence for EbR expression. The later expression and purification from membranes showed a His-tagged protein of consistent size by SDS-PAGE and western blot. This investigation therefore demonstrated that EbR was stably expressed in *E. coli* mimicking the results of Bratanov et al (2015). Furthermore, nMS characterisation of EbR was also feasible despite minimal biophysical investigations of EbR seen within the literature. Additionally, this investigation achieved this characterisation using a variety of membrane protein stabilisation methodologies. This ability to characterise EbR in using this workflow demonstrated that EbR is stably expressed.

4.1.2 EbR has low expression levels

Despite the stable expression of EbR, it was difficult to gauge the expression using the typical workflow using SDS-PAGE. The levels of expression were very low and could not be detected on Coomassie stained acrylamide gel. However, it was clearly detectable using western blot with anti-His antibody. This data implies that the expression levels of EbR were significantly lower than that of a typical heterologous expression of proteins with most protein expression being visible using SDS-PAGE. Bratanov et al (2015) did demonstrate that the expression of EbR was higher than that of other methodologies that attempted to express wild-type bacteriorhodopsin in *E. coli.* Future EbR studies should take this into account when designing their methodology.

4.1.3 EbR Possibly has unknown PTMs

This investigation found a variety of EbR masses using mass spectrometry using different reconstitution systems to isolate EbR. Of particular interest were results from G1 OGD isolated EbR where a variety of unidentifiable masses were observed. More specifically being too low to be monomeric EbR and some being too low to be EbR dimer but too high to be monomers. SDS-PAGE also did not reveal any contaminant proteins. This suggest that there must be other reasons for the masses seen. As G1 OGD was not identified within he spectrum also, this could suggest the presence of unknown PTMs that can be isolated due to the presence of a His-tag, allowing the purification by IMAC.

This is further confirmed by the fact no other oligomerisation states of G1 OGD HbR were observed implying that higher oligomers are not retained after isolation by G1 OGD. Furthermore, it is not known how expressing this modified EbR affects the insertion of EbR into the *E. coli* membrane, this is an important consideration as HbR (from which EbR is derived) has a maturation process that folds HbR to include the characteristic retinal group and ensures correct folding. Despite this being minimal information, there is a premise for further investigation preferably using denatured mass spectrometry investigations to identify possible EbR PTMs.

4.2 Novel amphipols (C_6-C_2 -50 and C_8-C_0-50)

4.2.1 Cyclic groups of novel amphipols reduce solubilisation efficiency

The amphipols (A8-35, C_6-C_2 -50 and C_8-C_0-50) appeared to preferentially solubilised EbR compared to that of HbR. This was demonstrated by the increase solubilisation efficiency from *E. coli* membranes (Figure 15) as compared with the resistance to solubilisation HbR from the PM (Figure 34). This further suggests that the environment from which these amphipols reconstitute a target MP is highly dependent on the nature of the membrane around the target MP. Increasing the solubilisation efficiency of HbR from the PM was tough to address given that these polymers precipitate beyond 500 mM NaCl, a variation in NaCl concentration could not be tested to investigate if this could improve solubilisation. This was desirable to test as HbR resides in high NaCl concentrations (~4 M) and it appeared to be a major factor in solubilising directly from the PM using SMA.

Initial solubilisation of EbR using the cyclic amphipols demonstrated a reduced efficiency of solubilisation compared their parent amphipols (A8-35). The novel amphipol structure included cyclic groups of different sizes. This is structurally a key difference between A8-35, C_6-C_2 -50 and C_8-C_0-50. In particular, A8-35 appeared to solubilise efficiently directly from the *E. coli* membranes all EbR present. Lipids were however identified within the C_6-C_2 -50 preparations within mass spectrometry. This does imply that the novel amphipols do retain some lipids in the gas-phase. As to whether this is more than their counterpart A8-35 is still unclear from this investigation as there were no clearly identifiable lipids within the spectrum.

Purification of EbR using both the novel amphipols showed a strong protein contaminant within the SDS-PAGE at ~35 kDa. This was then digested and analysed using LC-MS and identified as OmpA. This reduced efficiency of solubilisation resulting in a decreased concentration of solubilised EbR during IMAC. This reduced solubilisation efficiency

could be the reason for the contaminant protein observed. This is important as the reduced concentration of solubilised EbR also reduces competition for binding to the cobalt resin. This means usually that proteins with multiple histidine residues will also bind and can be eluted in fractions.

4.2.2 A8-35 and C_6-C_2-50 preserves native confirmation in the gas phase

This investigation set out to determine how well amphipols preserve native conformation and whether they are amenable to mass spectrometry. Given the use of A8-35 for mass spectrometry, the new amphipols were likely to also be amenable to mass spectrometry. Clear high resolution mass spectra were acquired using the A8-35 and the novel amphipols with low charge states (+16 and +20) suggesting that these amphipols are good contender for native mass spectrometry. Given the resistance to HbR to reconstitution by the novel amphipols, these amphipols demonstrated to primarily be suited to *E. coll* based expression and purification. It could be suggested that a mixture of both novel amphipols could provide a flexible purification methodology for other membranes.

The mass spectrometry of C_6-C_2-50 EbR preparations proved difficult due to the complexity of the samples and unexpected masses. One thing that was clear was the protein peaks that were present demonstrated to come from many associated ligands suspected to be lipids, implying that the environment within the novel amphipols are also fairly complex. Although, these did not require chemical dissociation like SMA. Lipids have already been found in purifications of AcrB using this amphipol also (Higgins et al., 2021). This data confirms the presence of associated molecules to these amphipols rather than excess lipids form the purification process. Use of these amphipols will also be facilitated by identifying the molecules retained by amphipols. The molecule retention can be understood by characterising a variety of novel amphipol isolated membrane protein complexes to determine lipid specific interactions as well as protein specific interactions. Particularly with a workflow similar to this investigation, using techniques such as native MS paired with Denatured MS based workflows. This is further evidence however, that these molecules could retain native confirmations and interactions with proteins of interest.

A key observation was that despite the observed higher oligomers (dimers and trimers) all demonstrated to have a low charge state for their oligomerisation state, with C_6-C_2-50 dimer observing a 15+ charge state and 19+ charge state of a suspected EbR trimer after HCD activation. Given the suspected oligomeric state of EbR, this would imply the

compact, folded state reducing solvent accessible surface of EbR during electrospray ionisation. Collectively this information shows that this C_6-C_2-50 and A8-35, could preserve the native protein fold.

4.2.3 C_6-C_2-50 Polymer-protein interactions and Protein-lipid interactions

During this investigation it was found that C_6-C_2-50 isolated EbR preparations came with large additional masses when investigated using native mass spectrometry. Initially it was assumed that the C_6-C_2-50 was being dissociated; additionally, it was clear that there were no unidentified contaminant proteins present in the C_6-C_2-50 preparations shown on SDS-PAGE (Figure 26). Activation experiments using the HCD cell at 190 V also yielded minimal change in mass of the observed molecules suggesting there were no large lipid-protein complexes (Figure 24). Similar masses were also observed when EbR was isolated using the parent polymer A8-35 which was known to be contain the EbR dimer as it was shown in subsequent western blots (Figure 27). It became clear that the polymer could actually be tightly associated with the protein revealing some patterns in additional masses observed (Table 5) especially as C_6-C_2-50 and A8-35 are of similar mass. There is some merit in this observation as initially membrane proteins were isolated using detergents before the additional use of amphipols. The inclusion of the detergents could reduce the strength of polymer-protein interaction making it easier to dissociate the polymer during mass spectrometry, explaining its repeated success thus far. In this investigation, the polymer was used to solubilise directly from *E. coli* membranes. Additionally, there is also minimal information on the mode of action of solubilising directly from membranes using these polymers and associated retained molecules (If Any). This data implies the strong association of C_6-C_2-50 to EbR and should be cause for further investigation of solubilisation directly from membranes using amphipols.

This investigation also demonstrated the ability to dissociate some suspected lipid molecules from the observed oligomers in native mass spectrometry (Figure 24). This thought was interesting as this would imply that the lipids are attatched on the outside of the polymer due to the suspected retention of polymer after this dissociation. It is not clear if this sort of retention would be still due to the lipid-protein interactions sites available despite the presence of the polymer or direct association of the lipids with the polymer. These polymers were designed with the cyclic groups to replicate the SMA mode of actions via the direct interaction of lipids. This is a further cause of investigation as it appears that the C_6-C_2-50 environment, like SMA, is also somewhat complex and

so it's mode of action and association to surrounding lipids should be a cause for future investigation.

4.2.4 Spectrophotometric limitation of C_6-C_2-50, C_8-C_0-50

Due to interference of HbR C_6-C_2-50, C_8-C_0-50 and DDM with spectrophotometric reading. A clear absorbance spectrum could not be acquired for HbR from these preparations. This does mean that an understanding of how these solubilising agents affect the photocycle of HbR could not be concluded upon. This does highlight that the use of these novel amphipols could affect spectrophotometric quantification methods of proteins after solubilisation. It should be noted that this was not observed before solubilisation. Furthermore, C_6-C_2-50 or C_8-C_0-50 does not cause significant interference on their own. The significantly low HbR concentrations suggests this is not to do with the protein content or protein characteristics after solubilisation. Therefore, suggesting the possible light scattering after solubilisation is more specific to the polymer. Furthermore, interference post solubilisation was not observed with other polymers such as SMA. Further confirming this hypothesis. This is important given our current minimal understanding of the amphipol mode of action.

4.3 SMA

4.3.1 SMA self-assembly follows the current proposed model

We have shown that HbR can be solubilised with a ~73 % solubilisation efficiency directly from its native purple membrane. The solubilisation of HbR using SMA was investigated by altering buffer composition. More specifically we decided to increase NaCl composition in increments from 250 mM to 4 M to reflect the high salt concentration in the natural environment of *H. salinarium*. This resistance to solubilisation directly from the purple membrane by both the novel amphipols and SMA has also been observed by Marconnet et al (2020). It should be noted that this investigation and Hoi et al (2021) used sonication and additional DMPC liposomes to increase solubilisation efficiency. This is likely due to the density of HbR in the PM and the rigidity of the lipids. However, these data suggest information on the SMA mode of action during solubilisation. Since SMAs introduction into MP investigations there has been studies into its mode of action during solubilisation. In particular, given the proposed direct interaction of SMA with lipids there are a variety of factors that govern the reconstitution efficiency of SMA. Key factors have been demonstrated to be lipid packing and electrostatic interactions (Scheidelaar et al., 2015). SMA is a negatively charged copolymer implying that insertion can be inhibited by anionic lipids of the purple membrane, also demonstrated by Scheidelaar et

al (2015). Therefore, to overcome this repulsion of the SMA by the charge of lipids, the electrostatic interactions have to be weakened to facilitate SMA insertion into the bilayer. This was also demonstrated to be provided by the addition of NaCl. It could therefore be proposed that due to the high density of HbR in the PM and few lipids present, a significant weakening of the electrostatic interactions of the PM are required for the insertion of SMA into the PM for HbR reconstitution, this is therefore provided by the extremely high salt (4 M) used. The contrast of using lower NaCl concentration (250 mM) for EbR reconstitution further confirms this theory. Given that the *E. coli* membrane is primarily Zwitterionic PE and is not as densely packed suggesting there is less requirement for the disruption of these electrostatic interactions between SMA and the MP annular lipid environment for MP reconstitution. It should be noted that 250 mM NaCl was not the optimum concentration for EbR but required minimal alteration to acquire pure SMA-EbR fractions. These findings follow the current proposed mode of action of SMA and therefore provide additional experimental data to support it.

4.3.2 SMA Preserves annular lipid interactions

This investigation set out to highlight methodologies to facilitate the use of SMA in structural biology due to its limitations during biophysical characterisation. Initially this investigation used typical workflow for mass spectrometry of HbR SMA samples. This proved to be difficult as the protein would not eject from the SMA. It became apparent that alterations to the workflow by chemically dissociating the SMA using a divalent cation solution would not only facilitate the characterisation of the protein but of the environment of the SMALP. Despite SMALPs being used for a variety of proteins there is little information characterising the internal environment of MP-SMALPs. During this workflow, SMA solubilised HbR did retain stability enough to characterise the sample with the use of quadrupole of the Q-Exactive UHMR.

It is clear from these results that the environment within the HbR-SMALP is complex and dissociating the polymer chemically provided insight into this complexity. As already mentioned, like many membrane proteins, HbR relies heavily on its local lipid interactions for structure and function. Despite the difficulty in acquiring nMS spectrum of HbR from SMALPS, it demonstrated no perturbations in the HbR baseline absorbance. Furthermore, it is already been shown by Cui et al (2015) that HbR has a reliance on the phosphatidylglycerophosphate methyl ester, 2,3-di-O-phytanyl-sn-glycero-1-phospho-3'-sn-glycerol-1'-methyl phosphate (PGP-me) headgroup for the photocycle. Additionally, Inada et al (2019) has found the trimerization and photocycle is dependent

on STGA-1 presence. Both these findings demonstrate the requirement of all annular lipids associated with HbR for the photocycle. Furthermore, implying the importance of the finding that SMA does not cause alterations in baseline absorbance and therefore structure of the proteins. This is presumably because of retention of these annular lipids. Further mass spectrometry analysis also attests to this hypothesis after chemical dissociation of the SMA polymer. The SMA solubilised HbR molecules started at 100kDa, which is too large to be a HbR trimer and too small to be a HbR tetramer. These were then systematically dismantled to show monomer with suspected lipid still associated. Collectively this data demonstrates that SMA at the very least preserves annular lipids but the SMALP environment is complex and requires further characterisation. The retention of the annular lipids and observed trimer demonstrated the requirement of this complex SMALP environment to retain the characteristic trimeric HbR and its photocycle activity.

4.3.3 SMA is tightly associated with HbR

This investigation demonstrated that HbR SMALPs can not be dissociated easily using the desolvation voltage of the Q-Exactive UHMR. These samples required chemical dissociation using divalent cations to investigate the solubilised protein. In contrast, Hoi et al (2020) ejected HbR from SMALP-DMPC lipodisqs with a desolvation voltage at -80 V could suggest that interactions between DMPC and SMA are weaker than those of PM lipids and SMA. This does imply the tight association of SMA with HbR after reconstitution from its native purple membrane.

4.4 G1 OGD modular detergent

4.4.1 G1 is a versatile reconstitution system

The idea of testing the modular detergents within this project was to determine their flexibility within a structural biology setting and to determine the possible effects of using this reconstitution system on target membrane proteins. The use of modular detergent also provided a reference when it was unclear to what was seen in other spectra. G1 OGD demonstrated the ability to solubilise HbR from both the PM and *E. coli* membranes with minimal optimisation of the purification process by Dr Leonhard Urner. This is an important characteristic when considering flexible purification systems. As demonstrated here certain reconstitution system require considerable optimisation before moving to sample characterisation. This reduced need for optimisation using the modular detergents could be beneficial in investigating a wide variety of protein in different membrane environments. The reader should be reminded however that this is one of few

investigations attempting to understanding this novel reconstitution system and the disadvantages would become more apparent with further use and investigations.

4.4.2 G1 OGD interactions are weaker

This reconstitution system required very little desolvation voltage (-30 V) to liberate HbR and EbR (-15 V) from the G1 OGD micelle (Figure 31 and Figure 36, respectviely). This is as compared with DDM, requiring -125 V to release EbR form the DDM micelle (figure 30). This implies that the modular detergent -protein interactions and G1 OGD -G1 OGD interaction are weaker than with the classical DDM detergent. This could also be a reason for the preservation of protein-lipid interactions seen within the G1 OGD HbR preparations.

These findings suggest G1 OGD could be a strong contender for a versatile system for membrane protein solubilisation and characterisation. However, further work should aim to understand how well this system preserves protein-protein interactions and lipid-protein interactions. As of currently it appears G1 OGD has only been used for mass spectrometry investigations and so it is not clear how well this applies to other characterisation methods such as CryoEM. This knowledge will come with further investigations of the limitations of G1 OGD.

4.4.3 G1 OGD doesn't retain MP oligomerisation

The absorbance spectrum of HbR reconstituted in SMA and G1 (Figure 44) suggested alterations in the secondary structure caused by G1 detergent as compared to SMA and the PM. Despite this there was clear retention of native lipids as demonstrated by nMS. It could be suggested that the modular detergent is replacing too many lipids and causing this alteration in structure. This is particularly important as trimeric HbR was not observed, which could be due to the facts that native lipids are particularly important in HbR trimer formation. In addition, smaller oligomeric sizes were observed for EbR preparations. It is not known however which interactions regulate EbR trimer and dimer formation but presumably it is a mix of lipid-protein and protein-protein interactions. It should be noted that the amount of lipids that are required to prevent structural alterations could be protein specific. This data demonstrates the importance of the annular lipid environment for MP structure and function and implies that modular detergents do not preserve all lipid-protein and protein-protein interactions leading to the absence of higher oligomeric states. Despite this it could be used as a tool to investigate monomeric membrane protein interactions such a lipid specific interactions.

4.4.4 G1 OGD is sensitive to buffer composition

Despite the lack of quantitation of mass photometry, there was a clear difference in distribution between using the PBS and using the G1 OGD solubilisation buffer. This was not observed with SMA suggesting that G1 OGD is significantly more sensitive to buffer composition than SMA, particularly as no other oligomeric states were observed in nMS of G1 OGD preparations. It is not clear from this data as to what part of the buffer it is sensitive to but a suggestion might be the salt content given the low 100 mM NaCl within the solubilisation buffer. Given the desired application of this reconstitution system to a variety of proteins and methodologies it would be desirable to understand the limitations of flexibility of these molecules.

4.5 Mass spectrometry for MP analysis

4.5.1 The Q-Exactive is ideal for probing protein-lipid interactions

This project aimed to use mass spectrometry to investigate lipid interactions. More specifically with the use of high resolution orbitrap mass spectrometry via the Q-Exactive UHMR. The application of this instrument during this investigation particularly proved more than capable of characterising membrane proteins. Having the flexibility in providing tools to remove undesired molecules from proteins in the gas phase. With capability of MS^3 investigations, this can be achieved using the Insource CID and HCD in conjunction with ion selection of the quadrupole to facilitate product ion identification. The sensitivity and high resolution allowed well resolved lipid peaks within the spectra and extremely narrow protein peaks.

The desolvation voltage helped remove reconstitution system from the protein, this then allowed the declustering and capture of the environment of the micelles and lipodisqs. In such a case as SMA HbR requiring chemical dissociation, it still allowed selection via the quadropole and gentle conditions to retain native fold and show protein peaks within the spectrum among other associated ions. This challenged the Q-Exactive UHMR in being both gentle for chemically dissociated SMA preparations and harsh for stronger interactions while still maintaining high resolution. This investigation proved the full capability of the UHMR in characterising complex biological mixtures.

4.6 Mass photometry of MPs

Mass photometry was used with the aim of elucidating the heterogeneity of oligmerisation states associated with each reconstitution systems. It presented to be a reliable and simple technique for mass analysis of membrane proteins in Olerinyova et

al (2021). Despite this, the presented investigation highlighted areas of this techniques that require further development.

4.6.1 Membrane protein standards should be used for mass photometry

In being recently used for membrane protein investigations, little is known how well mass photometry lends itself to the typical workflow of structural biology. One of the key problems associated with this technique was the lack of reproducibility associated with reading of samples. Increasing the sample concentration also demonstrated an increase in observed mass. It is not incorrect to perceive that this is likely to do with the complexity of mixture used within this investigation, containing glycerol, salt, lipids and excess polymers or even micelles.

It should be noted that calibrations are also carried out with soluble proteins in PBS buffer. It is therefore becoming increasingly clear that to facilitate the use of mass photometry for membrane protein oligomerisation analysis, membrane protein standards should be used while accounting for heterogeneity of the desired lipid particles. This would be done using membrane protein standards containing the same reconstitution system as your sample with the same lipids. Furthermore, the application of standard protocol for the use of a mass photometer would greatly benefit frequent users. To this point, it should be noted that a recent methodology paper was published detailing how best to prepare glass slides and prepare a variety of samples for mass photometry by Wu & Piszczek (2021). It is currently unclear even after alterations in calibrations and protocols whether more accurate readings could be obtained.

Much work has been carried out attempting to understand how the presence of detergents varies the readings from a mass photometer. It has been demonstrated that below Critical micelle concentration (CMC), detergents will create the least amount of effect on readings. With the use of detergents in samples for membrane proteins being significantly higher than CMC it could be desirable to avoid the use of these detergents for mass photometry or factor these into the calibrations as mentioned before. As of currently, there is minimal information on the use of SMA and other polymers in mass photometry. Given this effect of detergents on the readings, it could be preferential to investigate how polymers affect these readings. This could determine their application in this setting and if they’re preferential to detergents.

4.6.2 Mass photometry isn’t accurate at low molecular weights

Another factor that needs to be taken into account for this investigation is size of proteins. This investigation used small protein at 27 kDa monomeric mass which was not expected

to be detected. It was expected that higher order oligomers would be detected. being expected such as dimer (54 kDa) and trimer (81 kDa) are just above the 40 kDa detection limit of mass photometry. Also given the use of SMALPs, incorporation of lipodisq would further increase the mass. It could be that the accuracy is hindered this close to the detection cut off. Finally, the difference between oligomers is only 27 kDa, suggesting that using mass photometry to read minute differences of oligomer is not preferential. Increasing the size of the target protein will create a larger difference in mass, and therefore a larger mass difference between oligomeric distributions. This could explain as to why Olerinyova et al (2021) showed statistically significant readings even with such complex mixtures. This shows that as of currently mass photometry will likely be most applicable to larger membrane proteins.

Chapter 5: Conclusion

This investigation set out to compare current and novel membrane protein reconstitution systems. Particularly, we aim to understand the limitations of applying them to current workflows of biophysical characterisation of membrane proteins in their native conformations, especially around mass spectrometry and mass photometry.

Despite causing alterations in secondary structure, G1 demonstrated to preserve minimal protein-lipid interactions. From this investigation G1 could be considered as a potential universal reconstitution system for a variety of membrane proteins. The main theme of investigations of this protein should focus on finding limitations to sufficiently understand G1 OGD and how well the associated lipid environment and protein-protein interactions are retained.

Despite the difficulties of using SMA for biophysical characterisation, in solution dissociation of SMA polymer from the SMALP could provide a good workflow for characterisation studies using this polymer, particularly using nMS. The preservation of annular lipid interactions still makes this an attractive reconstitution system and effort should be made to alter workflows to accommodate for it pitfalls and understand the heterogeneity of the SMALP.

The novel amphipols used in this study provide a reliable solubilisation molecule to retain native conformations. Purification protocols still need to be optimised to limit contaminants. Further investigations should aim to understand better how these novel amphipols work in various membrane and protein combinations. Aiming to understand

their mode of action and how best to apply them (concentrations, temperature pHs…etc) as they might be optimised for other membrane types.

The Q-Exactive UHMR has proved it to be a reliable and powerful mass spectrometer with much control over the forces experienced by the sample. Altering the injection flatapole and potential gradient was optimal for this size membrane protein and should be adopted by investigations using similar sizes proteins.

All of the reconstitution systems used were amenable to native mass spectrometry although it was clear some more than others. A8-35 and G1 OGD required the lowest desolvation voltages to liberate the protein from the solubilisation system, suggesting they're the most amenable. A8-35 appears to retain more oligomerisation states than G1 OGD. Although, the associated lipids were detected in all G1 OGD preparations from both EbR and HbR which was not the case for A8-35. Therefore, the use of either methodology would therefore be dependant on the investigation at hand.

Mass photometry is a promising technique with high sensitivity in its infancy. Despite this it is not suitable for investigating small membrane proteins and their molecular interactions but could be useful for larger proteins and oligomers. Further work should also aim to understand how the reconstitution systems affect the readings using mass photometry such as amphipols, SMALPs.

References

Angelini, R., Babudri, F., Lobasso, S. & Corcelli, A. (2010) MALDI-TOF/MS Analysis of Archaebacterial Lipids in Lyophilized Membranes Dry-Mixed with 9-Aminoacridine. *Journal of Lipid Research*, 51 (9), pp. 2818–2825.

Bender, J. & Schmidt, C. (2019) Mass Spectrometry of Membrane Protein Complexes. *Biological Chemistry*, 400 (7), pp. 813–829.

Bondar, A. N., Elstner, M., Suhai, S., Smith, J. C. & Fischer, S. (2004) Mechanism of Primary Proton Transfer in Bacteriorhodopsin. *Structure*, 12 (7), pp. 1281–1288.

Braimen, M. & Matheis, R. (1982) Resonance Raman Spectra of Bacteriorhodopsin ' s Primary Photoproduct : Evidence for a Distorted 13-Cis Retinal Chromophore Author (s): Mark Braiman and Richard Mathies Published by : National Academy

of Sciences Stable URL : Https://Www.Jstor.Org/Stabl. 79 (2), pp. 403–407.

Bratanov, D., Balandin, T., Round, E., Shevchenko, V., Gushchin, I., Polovinkin, V., Borshchevskiy, V. & Gordeliy, V. (2015) An Approach to Heterologous Expression of Membrane Proteins. The Case of Bacteriorhodopsin. *PLoS ONE*, 10 (6), pp. 1–16.

Calabrese, A. N., Watkinson, T. G., Henderson, P. J. F., Radford, S. E. & Ashcroft, A. E. (2015) Amphipols Outperform Dodecylmaltoside Micelles in Stabilizing Membrane Protein Structure in the Gas Phase. *Analytical Chemistry*, 87 (2), pp. 1118–1126.

Campuzano, I. D. G., Li, H., Bagal, D., Lippens, J. L., Svitel, J., Kurzeja, R. J. M., Xu, H., Schnier, P. D. & Loo, J. A. (2016) Native MS Analysis of Bacteriorhodopsin and an Empty Nanodisc by Orthogonal Acceleration Time-of-Flight, Orbitrap and Ion Cyclotron Resonance. *Physiology & behavior*, 176 (1), pp. 139–148.

Catalano, C., Ben-Hail, D., Qiu, W., Blount, P., Georges, A. des & Guo, Y. (2021) Cryo-Em Structure of Mechanosensitive Channel Ynai Using Sma2000: Challenges and Opportunities. *Membranes*, 11 (11).

Chorev, D. S., Baker, L. A., Wu, D., Beilsten-Edmands, V., Rouse, S. L., Zeev-Ben-Mordehai, T., Jiko, C., Samsudin, F., Gerle, C., Khalid, S., Stewart, A. G., Matthews, S. J., Grünewald, K. & Robinson, C. V. (2018) Protein Assemblies Ejected Directly from Native Membranes Yield Complexes for Mass Spectrometry. *Science*, 362 (6416), pp. 829–834.

Chorev, D. S., Tang, H., Rouse, S. L., Bolla, J. R., Kügelgen, A. von, Baker, L. A., Wu, D., Gault, J., Grünewald, K., Bharat, T. A. M., Matthews, S. J. & Robinson, C. V. (2020) The Use of Sonicated Lipid Vesicles for Mass Spectrometry of Membrane Protein Complexes. *Nature Protocols*, 15 (5), pp. 1690–1706.

Cong, X., Liu, Y., Liu, W., Liang, X., Russell, D. H. & Laganowsky, A. (2016) Determining Membrane Protein-Lipid Binding Thermodynamics Using Native Mass Spectrometry. *Journal of the American Chemical Society*, 138 (13), pp. 4346–4349.

Cui, J., Kawatake, S., Umegawa, Y., Lethu, S., Yamagami, M., Matsuoka, S., Sato, F., Matsumori, N. & Murata, M. (2015) Stereoselective Synthesis of the Head Group of Archaeal Phospholipid PGP-Me to Investigate Bacteriorhodopsin–Lipid

Interactions. *Organic & Biomolecular Chemistry*, 13 (41), pp. 10279–10284.

Eliuk, S. & Makarov, A. (2015) Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Review of Analytical Chemistry*, 8, pp. 61–80.

Essen, L. O., Siegert, R., Lehmann, W. D. & Oesterhelt, D. (1998) Lipid Patches in Membrane Protein Oligomers: Crystal Structure of the Bacteriorhodopsin-Lipid Complex. *Proceedings of the National Academy of Sciences of the United States of America*, 95 (20), pp. 11673–11678.

Gohon, Y., Dahmane, T., Ruigrok, R. W. H., Schuck, P., Charvolin, D., Rappaport, F., Timmins, P., Engelman, D. M., Tribet, C., Popot, J. L. & Ebel, C. (2008) Bacteriorhodopsin/Amphipol Complexes: Structural and Functional Properties. *Biophysical Journal*, 94 (9), pp. 3523–3537.

Gupta, K., Li, J., Liko, I., Gault, J., Bechara, C., Wu, D., Hopper, J. T. S., Giles, K., Benesch, J. L. P. & Robinson, C. V. (2018) Identifying Key Membrane Protein Lipid Interactions Using Mass Spectrometry. *Nature Protocols*, 13 (5), pp. 1106–1120.

Hasegawa, N., Jonotsuka, H., Miki, K. & Takeda, K. (2018) X-Ray Structure Analysis of Bacteriorhodopsin at 1.3 Å Resolution. *Scientific Reports*, 8 (1), pp. 1–8.

Hellwig, N., Peetz, O., Ahdash, Z., Tascón, I., Booth, P. J., Mikusevic, V., Diskowski, M., Politis, A., Hellmich, Y., Hänelt, I., Reading, E. & Morgner, N. (2018) Native Mass Spectrometry Goes More Native: Investigation of Membrane Protein Complexes Directly from SMALPs. *Chemical Communications*, 54 (97), pp. 13702–13705.

Hesketh, S. J., Klebl, D. P., Higgins, A. J., Thomsen, M., Pickles, I. B., Sobott, F., Sivaprasadarao, A., Postis, V. L. G. & Muench, S. P. (2020) Styrene Maleic-Acid Lipid Particles (SMALPs) into Detergent or Amphipols: An Exchange Protocol for Membrane Protein Characterisation. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1862 (5), p. 183192.

Higgins, A. J., Flynn, A. J., Marconnet, A., Musgrove, L. J., Postis, V. L. G., Lippiat, J. D., Chung, C., Ceska, T., Zoonens, M., Sobott, F. & Muench, S. P. (2021) For CryoEM Analysis. pp. 1–9.

Hoi, K. K., Bada Juarez, J. F., Judge, P. J., Yen, H.-Y., Wu, D., Vinals, J., Taylor, G. F.,

Watts, A. & Robinson, C. V (2021) Detergent-Free Lipodisq Nanoparticles Facilitate High-Resolution Mass Spectrometry of Folded Integral Membrane Proteins. *Nano letters*.

Inada, M., Kinoshita, M. & Matsumori, N. (2019) Archaeal Glycolipid S-TGA-1 Is Crucial for Trimer Formation and Photocycle Activity of Bacteriorhodopsin. *ACS Chemical Biology*, 15 (1), pp. 197–204.

Jawurek, M., Dröden, J., Peter, B., Glaubitz, C. & Hauser, K. (2018) Lipid-Induced Dynamics of Photoreceptors Monitored by Time-Resolved Step-Scan FTIR Spectroscopy. *Chemical Physics*, 512, pp. 53–61.

Jeganathan, C., Thamaraiselvi, K. & Sabari Girisun, T. C. (2019) Improved Production of Bacteriorhodopsin from Halobacterium Salinarum through Direct Amino Acid Supplement in the Basal Medium. *Extremophiles*, 23 (1), pp. 133–139.

Kühlbrandt, W. (2000) Bacteriorhodopsin — the Movie. 406 (August).

Lee, S. C., Knowles, T. J., Postis, V. L. G., Jamshad, M., Parslow, R. A., Lin, Y. P., Goldman, A., Sridhar, P., Overduin, M., Muench, S. P. & Dafforn, T. R. (2016) A Method for Detergent-Free Isolation of Membrane Proteins in Their Local Lipid Environment. *Nature Protocols*, 11 (7), pp. 1149–1162.

Li, Y., Struwe, W. B. & Kukura, P. (2020) Single Molecule Mass Photometry of Nucleic Acids. *Nucleic Acids Research*, 48 (17), p. E97.

Marconnet, A., Michon, B., Bon, C. Le, Giusti, F., Tribet, C. & Zoonens, M. (2020) Solubilization and Stabilization of Membrane Proteins by Cycloalkane-Modified Amphiphilic Polymers. *Biomacromolecules*, 21 (8), pp. 3459–3467.

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A.-N., Standfuss, J., Neutze, R. & Iwata, S. (2016) A Three-Dimensional Movie of Structural Changes in Bacteriorhodopsin. *Science*, 354 (6319), pp. 1552–1557.

Nekrasova, O. V, Wulfson, A. N., Tikhonov, R. V & Yakimov, S. A. (2010) A New Hybrid Protein for Production of Recombinant Bacteriorhodopsin in Escherichia Coli A New Hybrid Protein for Production of Recombinant Bacteriorhodopsin in Escherichia Coli. *Journal of Biotechnology*, 147 (3–4), pp. 145–150.

Olerinyova, A., Sonn-Segev, A., Gault, J., Eichmann, C., Schimpf, J., Kopf, A. H., Rudden, L. S. P., Ashkinadze, D., Bomba, R., Frey, L., Greenwald, J., Degiacomi, M. T., Steinhilper, R., Killian, J. A., Friedrich, T., Riek, R., Struwe, W. B. & Kukura, P. (2021) Mass Photometry of Membrane Proteins. *Chem*, 7 (1), pp. 224–236.

Orwick-Rydmark, M., Lovett, J. E., Graziadei, A., Lindholm, L., Hicks, M. R. & Watts, A. (2012) Detergent-Free Incorporation of a Seven-Transmembrane Receptor Protein into Nanosized Bilayer Lipodisq Particles for Functional and Biophysical Studies. *Nano Letters*, 12 (9), pp. 4687–4692.

Parmar, M., Rawson, S., Scarff, C. A., Goldman, A., Dafforn, T. R., Muench, S. P. & Postis, V. L. G. (2018) Using a SMALP Platform to Determine a Sub-Nm Single Particle Cryo-EM Membrane Protein Structure. *Biochimica et Biophysica Acta - Biomembranes*, 1860 (2), pp. 378–383.

Pimlott, D. J. D. & Konermann, L. (2021) Using Covalent Modifications to Distinguish Protein Electrospray Mechanisms: Charged Residue Model (CRM) vs. Chain Ejection Model (CEM). *International Journal of Mass Spectrometry*, 469, p. 116678.

Raab, S. A., El-Baba, T. J., Laganowsky, A., Russell, D. H., Valentine, S. J. & Clemmer, D. E. (2021) Protons Are Fast and Smart; Proteins Are Slow and Dumb: On the Relationship of Electrospray Ionization Charge States and Conformations. *Journal of the American Society for Mass Spectrometry*.

Ratkeviciute, G., Cooper, B. F. & Knowles, T. J. (2021) Methods for the Solubilisation of Membrane Proteins: The Micelle-Aneous World of Membrane Protein Solubilisation. *Biochemical Society Transactions*, 49 (4), pp. 1763–1777.

Reichow, S. L. & Gonen, T. (2009) Lipid-Protein Interactions Probed by Electron Crystallography. *Current Opinion in Structural Biology*, 19 (5), pp. 560–565.

Renner, C., Kessler, B. & Oesterhelt, D. (2005) Lipid Composition of Integral Purple Membrane by 1H and 31P NMR. *Journal of Lipid Research*, 46 (8), pp. 1755–1764.

Rohner, T. C., Lion, N. & Girault, H. H. (2004) Electrochemical and Theoretical Aspects of Electrospray Ionisation. *Physical Chemistry Chemical Physics*, 6 (12), pp. 3056–3068.

Rose, R. J., Damoc, E., Denisov, E., Makarov, A. & Heck, A. J. R. (2012) High-Sensitivity Orbitrap Mass Analysis of Intact Macromolecular Assemblies. *Nature Methods*, 9 (11), pp. 1084–1086.

Scheidelaar, S., Koorengevel, M. C., Pardo, J. D., Meeldijk, J. D., Breukink, E. & Killian, J. A. (2015) Molecular Model for the Solubilization of Membranes into Nanodisks by Styrene Maleic Acid Copolymers. *Biophysical Journal*, 108 (2), pp. 279–290.

Shiu, P. J., Chen, H. M. & Lee, C. K. (2014) One-Step Purification of Delipidated Bacteriorhodopsin by Aqueous-Three-Phase System from Purple Membrane of Halobacterium. *Food and Bioproducts Processing*, 92 (2), pp. 113–119.

Shiu, P. J., Ju, Y. H., Chen, H. M. & Lee, C. K. (2013) Facile Isolation of Purple Membrane from Halobacterium Salinarum via Aqueous-Two-Phase System. *Protein Expression and Purification*, 89 (2), pp. 219–224.

Sohlenkamp, C. & Geiger, O. (2016) Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiology Reviews*, 40 (1), pp. 133–159.

Soltermann, F., Foley, E. D. B., Pagnoni, V., Galpin, M., Benesch, J. L. P., Kukura, P. & Struwe, W. B. (2020) Quantifying Protein – Protein Interactions by Molecular Counting with Mass Photometry. pp. 10774–10779.

Sonn-segev, A., Belacic, K., Bodrug, T., Young, G., Vanderlinden, R. T., Schulman, B. A., Schimpf, J., Friedrich, T., Dip, P. V., Schwartz, T. U., Bauer, B., Peters, J., Struwe, W. B., Benesch, J. L. P., Brown, N. G., Haselbach, D. & Kukura, P. (2020) Machines by Mass Photometry. *Nature Communications*, pp. 1–10.

Stauffer, M., Hirschi, S., Ucurum, Z., Harder, D., Schlesinger, R. & Fotiadis, D. (2020) Engineering and Production of the Light-Driven Proton Pump Bacteriorhodopsin in 2d Crystals for Basic Research and Applied Technologies. *Methods and Protocols*, 3 (3), pp. 1–22.

Stroud, Z., Hall, S. C. L. & Dafforn, T. R. (2018) Purification of Membrane Proteins Free from Conventional Detergents: SMA, New Polymers, New Opportunities and New

Insights. *Methods*, 147 (2018), pp. 106–117.

Tahara, S., Kuramochi, H., Takeuchi, S. & Tahara, T. (2019) Protein Dynamics Preceding Photoisomerization of the Retinal Chromophore in Bacteriorhodopsin Revealed by Deep-UV Femtosecond Stimulated Raman Spectroscopy. *Journal of Physical Chemistry Letters*, 10 (18), pp. 5422–5427.

Tamara, S., Boer, M. A. Den & Heck, A. J. R. (2021) High-Resolution Native Mass Spectrometry. *Chemical Reviews*.

Tribet, C., Audebert, R. & Popot, J. L. (1996) Amphipols: Polymers That Keep Membrane Proteins Soluble in Aqueous Solutions. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (26), pp. 15047–15050.

Urner, L. H., Goltsche, K., Selent, M., Liko, I., Schweder, M. P., Robinson, C. V., Pagel, K. & Haag, R. (2021) Dendritic Oligoglycerol Regioisomer Mixtures and Their Utility for Membrane Protein Research. *Chemistry - A European Journal*, 27 (7), pp. 2537–2542.

Urner, L. H., Liko, I., Yen, H. Y., Hoi, K. K., Bolla, J. R., Gault, J., Almeida, F. G., Schweder, M. P., Shutin, D., Ehrmann, S., Haag, R., Robinson, C. V. & Pagel, K. (2020) Modular Detergents Tailor the Purification and Structural Analysis of Membrane Proteins Including G-Protein Coupled Receptors. *Nature Communications*, 11 (1).

Wickstrand, C., Dods, R., Royant, A. & Neutze, R. (2015) Bacteriorhodopsin: Would the Real Structural Intermediates Please Stand Up? *Biochimica et Biophysica Acta - General Subjects*, 1850 (3), pp. 536–553.

Wu, D. & Piszczek, G. (2021) Standard Protocol for Mass Photometry Experiments. *European Biophysics Journal*, 50 (3–4), pp. 403–409.

Yamagami, M., Tsuchikawa, H., Cui, J., Umegawa, Y., Miyazaki, Y., Seo, S., Shinoda, W. & Murata, M. (2019) Average Conformation of Branched Chain Lipid PGP-Me That Accounts for the Thermal Stability and High-Salinity Resistance of Archaeal Membranes. *Biochemistry*, 58 (37), pp. 3869–3879.

Zheng, H., Liu, W., Anderson, L. Y. & Jiang, Q.-X. (2011) Lipid-Dependent Gating of a Voltage-Gated Potassium Channel. *Nature Communications*, 2 (1).

Zoonens, M. & Popot, J. L. (2014) Amphipols for Each Season. *Journal of Membrane*

Biology, 247 (9–10), pp. 759–796.

Zubarev, R. A. & Makarov, A. (2013) Orbitrap Mass Spectrometry. *Analytical Chemistry*, 85 (11), pp. 5288–5296.