Spectral tuning of Archaerhodopsin 3 in *Escherichia coli*

Owen Crowther

Exeter College

University of Oxford

A thesis submitted for the degree of

Masters by Research in Biochemistry

Trinity Term 2018

Abstract

Microbial rhodopsins are proteins that act as light driven proton pumps located in the plasma membrane of eukarya, bacteria and archaea. These proteins facilitate the movement of protons by light-induced isomerisation of a central retinal chromophore found in the transmembrane core of the opsin protein. Here we detail an extensive study of amino acid substitutions in the retinal binding pocket of archaerhodopsin 3 (AR3) and their effect on the absorbance spectrum of the protein for use in biotechnological applications.

In this thesis, protein engineering was used to generate a mutant library that produced hypsochromic and bathochromic shifts when purified from an *E. coli* model into a detergent environment. The largest shifts reported here (-25 and +7 nm compared to *wt*AR3) were achieved through strategic mutation of residues interacting with the retinal chromophore. The S151C mutant produced in this report perturbed the photocycle kinetics in numerous detergent systems tested, resulting in a delayed photo-intermediate decay and return to ground state.

The faster photocycle kinetics of AR3 compared to other rhodopsins makes it a promising target for both optogenetic and bio-sensory applications. For its use as a single molecule bio-sensor, a strategic cysteine mutation was incorporated into the intracellular EF loop of the protein to allow for its future immobilisation onto a conductive surface. With monomeric AR3 adsorbed onto a gold substrate, it can be assessed for its ability to generate a detectable proton current in response to activation by light of a wavelength specific to each mutant. For applications in optogenetics, distinctly shifted mutants can allow for targeted activation of different mutants in the same neural system, thus allowing for the assaying of more complex pathways. Red-shifted mutants are of particular interest in this field as they allow for non-invasive light penetration in the brain.

Acknowledgements

I would firstly like to express my gratitude and thanks to Professor Anthony Watts for his supervision and guidance throughout my Masters project and for offering me the opportunity to study at the University of Oxford.

My sincere thanks go to Javier Viñals Camallonga for helping to create this project. I am hugely grateful for the advice and assistance that you have given me throughout the year of my research. Your constant patience and willingness to train and help me with all aspects of the work has been an incredible experience. Javi, I will miss your questionable music taste, Spanish rants, constant smile and good mood, even on a weekend.

I must thank the other members of the lab, (the soon to be Dr) Juan Bada Juarez who has been another constant help with troubleshooting and advice in all aspects of the lab life and with the writing of this thesis. Your encyclopaedic knowledge of all things biophysics will be greatly missed by all in the lab. Henry Sawczyc 'Dad', you have been an integral part of my enjoyment over the last year, helping with both lab and college life has made my time here so much more enjoyable and helped me find my feet at the university. My thanks must also go to Dr Peter Judge who was more than willing to help whenever possible, especially with proof reading of this thesis. To the others who I have had the pleasure of working with: Sarah Frei; Dr Jason Yau; Steven Lavington and Constanze Cavalier, you have my thanks.

To all of my new-found friends in Oxford: in Biochemistry, Exeter College or elsewhere. Thank you.

Finally, and most importantly, I dedicate this thesis to my family. Ben, you have been a constant inspiration and rival throughout my academic life. I hope this thesis serves to prove that I am in fact the smarter sibling (at least on paper). Grandma, thank you for your everlasting love and support throughout all of my pursuits over the years. Mum and Dad, your unbelievable generosity in funding this Masters has allowed me to pursue my genuine passion. I cannot begin to thank you enough for your unfailing support and continuous encouragement throughout my year of study, through the process of researching and writing this thesis and generally throughout my entire life. None of this would have been possible without you.

Contents

List of Abbreviations

Å . . . . . Ångström

AAV . . . . . Adeno-associated Virus

AFM . . . . . Atomic Force Microscopy

AR3 . . . . . Archaerhodopsin 3

bR . . . . . Bacteriorhodopsin

C-AFM . . . . Conductive AFM

CD . . . . . Circular Dichroism

CHAPS . . . . 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate

CMC . . . . . Critical Micellar Concentration

CV . . . . . Column Volumes

DDM . . . . . *n*-Dodecyl-β-D-matloside

ddH_2O . . . . Double Distilled Water

E. coli . . . . . *Escherichia coli*

EtBr . . . . . Ethidium Bromide

GPCR . . . . . G Protein-Coupled Receptor

IMAC . . . . . Immobilised Metal Ion Affinity Chromatography

IPTG Isopropyl-β-D-1-thiogalactopyranoside

LB . . . . . Luria Broth

LCD . . . . . Liquid Crystal Display

LDS . . . . . Lithium Dodecyl Sulfate

MES . . . . . 2-(*N*-morpholino)ethanesulfonic acid

mMm . . . . . Metal-Molecule-Metal

MSP . . . . . Membrane Scaffold Protein

OD . . . . . Optical Density

OG . . . . . *n*-Octyl-β-D-glucopyranoside

ORF . . . . . Open Reading Frame

PBS . . . . . Phosphate-Buffered Saline

PCR . . . . . Polymerase Chain Reaction

PDB . . . . . Protein Data Bank

PEG . . . . . Poly(ethylene glycol)

PGP-Me . . . Phosphatidylglycerophosphate Methyl Ester

PM . . . . . Purple Membrane

RT . . . . . Room Temperature

SDS-PAGE . . Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

SEC . . . . . Size Exclusion Chromatography

SMA . . . . . Styrene Maleic Acid

TAE . . . . . Tris Base-Acetic Acid-EDTA

Tris . . . . . Tris(hydroxymethyl)aminomethane

WT . . . . . Wild Type

***wt*AR3** . . . . Wild Type Archaerhodopsin 3

$\lambda_{\mathbf{max}}$ . . . . . Wavelength of Maximal Absorbance

1 Introduction

1.1 Biological Membranes

All organisms are made up of single, or multiple cells encapsulated by a membrane[1, 2]. These membranes act as semi-permeable barriers between the intracellular and extracellular environments, controlling the movement of ions and compounds in and out as well as protecting the cell from the physical stresses of its environment[3]. The membrane also has a central role in cellular signalling and processing biological information, transport and cell cycle regulation[4].

Biological membranes are made up of a combination of lipids, proteins and carbohydrates. The composition varies depending on the cell type and function. In a typical cell, proteins contribute approximately 50% of the dry weight of the membrane, however this can vary between 20 and 80% depending on the specific cell type and its function[5].

Phospholipids spontaneously form a bilayer structure, due to the properties of the hydrophobic fatty acid tails and hydrophilic headgroups. The fluid mosaic model[6, 7] was developed to describe the bilayer structure, the dynamic and asymmetric interactions between the lipids and proteins and the roles that each play in the cell (Fig 1.1). Recently, the interactions between lipids and membrane proteins, historically thought of as transient and of limited functional significance[8, 9], have

emerged as an important consideration when characterising protein structure and function in the native membrane environment. Proteins may appear to function in a particular manner in the model system in which it is tested, but in its native environment can differ both structurally and functionally.

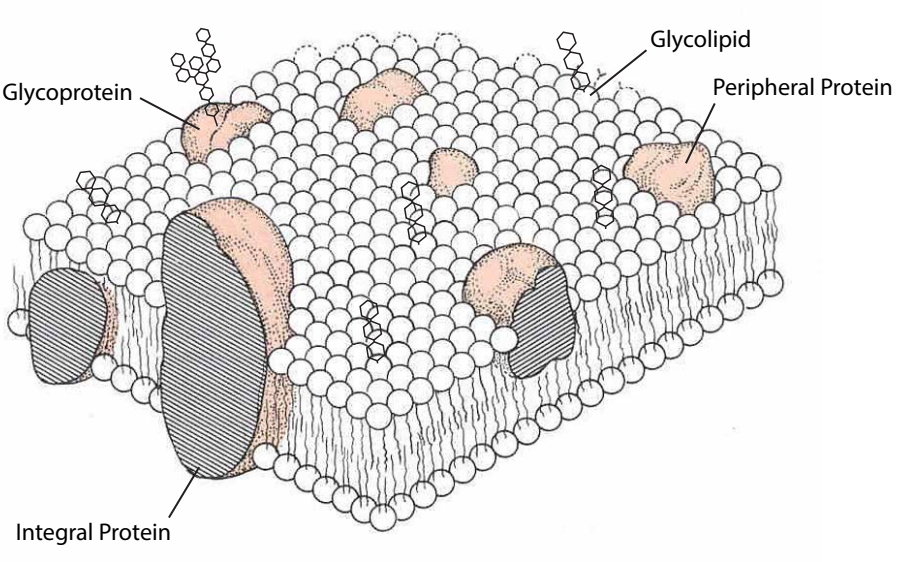


Figure 1.1: The fluid mosaic model of a membrane. The phospholipid bilayer is studded with proteins. Peripheral membrane proteins attach to one leaflet of the membrane. Integral proteins span the membrane with their hydrophobic core. The external face of the bilayer can accommodate glycolipid and glycoprotein modifications. Asymmetry between the bilayers arises from differing distribution of lipid species and proteins between leaflets. Adapted from [6].

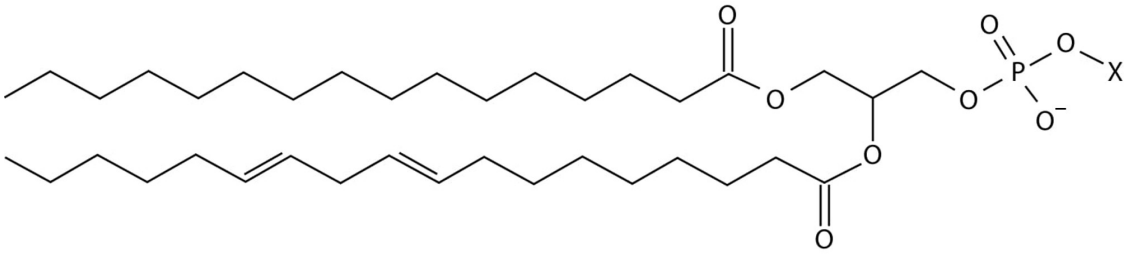
1.1.1 Membrane Lipids

Lipids are the most abundant component in the membrane by number[3]. The class of lipid is determined by their headgroup; most common are phospholipids and glycolipids. In a phospholipid, the head group is attached to a glycerol moiety which is then bound to two acyl chains from carbon *sn*-1 and *sn*-2. It is the acyl chains that form the hydrophobic tail of the lipid that are sandwiched in the non-polar centre of a bilayer. The acyl chains vary in length and degree of saturation.

Originally, lipids were thought of simply as the main structural component forming the barrier between the cell and its external environment, but have now been shown

to have important functional roles. For example the introduction of phospholipase A_2 at the human erythrocyte membrane, which hydrolyses 2-acyl lipids into a fatty acid and lysophospholipid[10], causes a reduction in the rate of glucose transport and Na^+/K^+-ATPase activity[11].

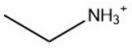
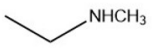
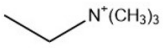
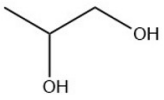
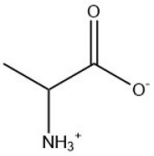
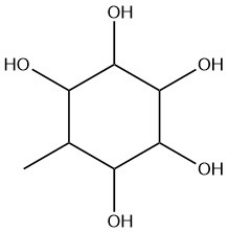
Bacteria contain many similar lipids to eukaryotic organisms, but also have a more restricted range of glycerophospholipids and cardiolipin[12, 13]. Figure 1.2 displays the general structure of a glycerophospholipid, the acyl chain length varies depending on lipid species, along with the head group (the most common of which are shown in Table 1.1). The most abundant phospholipid is phosphatidylethanolamine (PE), with approximately 58% of the total weight of the cell lipids in *E. coli*. Phosphatidylglycerol (PG) is also abundant, making up 15% of the membrane by weight[14]. Cardiolipin (Fig 1.3), an abundant lipid in the mitochondrial inner membrane, also accounts for 10% of the lipid composition by mass[15, 16]. One role of this lipid is to preserve the quaternary structure of cytochrome c oxidase in the membrane to allow for full enzymatic activity[17]. It also functions as an electron carrier between cytochrome c and cytochrome c oxidase as the final step in the electron transport chain[18].

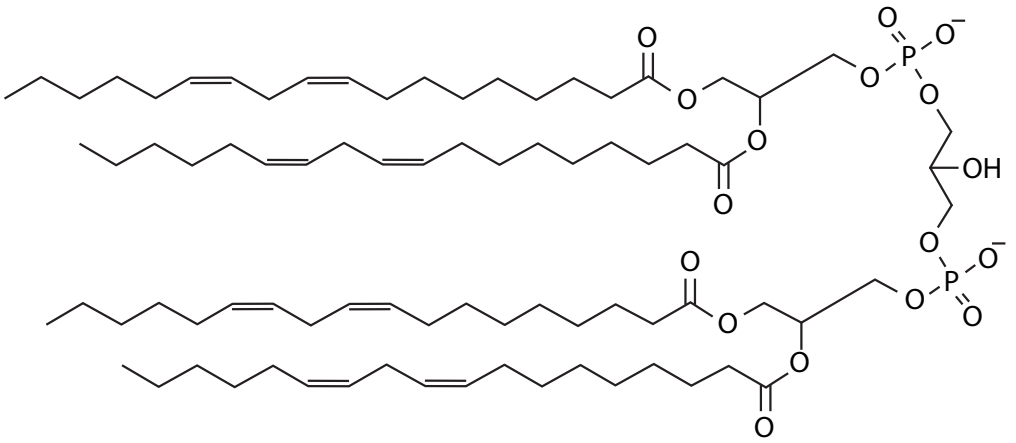


Glycerophospholipid

Figure 1.2: General structure of a phospholipid. The glycerol molecle is esterified to the two fatty acids and phosphate group. Acyl chain length and saturation vary depending on lipid species. 'X' denotes variable head group. Common phospholipid head groups shown in Table 1.1.

Table 1.1: Common head groups of phospholipids found in *Escherichia coli*. Head groups replace 'X' on Fig 1.2.

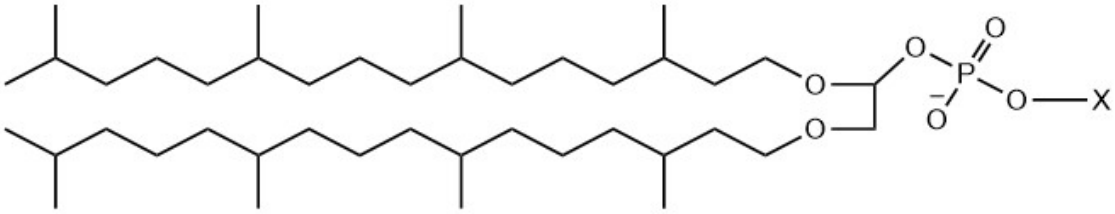
Phospholipid Class	Head Group
Phosphatidic Acid (PA)	
Phosphatidylethanolamine (PE)	
Phosphatidylmethylethanolamine (PME)	
Phosphatidylcholine (PC)	
Phosphatidylglycerol (PG)	
Phosphatidylserine (PS)	
Phosphatidylinositol (PI)	



Cardiolipin (CL)

Figure 1.3: The strucutre of cardiolipin, a common lipid found in bacteria, prodominantly at the poles and division septum of the cell[19]. Analogous lipids are present in small quantities in archaea.

The lipid content in archaea is distinct from that of the other domains of life. Whereas bacteria and eukaryotic cells have membrane lipids formed of acyl chains and charged head groups bound to a central glycerol by ester bonds, archaea form their lipids by ether bonding between the phytane chain and glycerol moiety[20]. The phytanyl groups are almost always fully saturated and made up of a 5-carbon repeating unit, binding to the *sn*-2,3 carbon positions of the glycerol backbone (instead of *sn*-1,2 glycerol carbons as in eukaryotes and bacteria). This unique lipid architecture is thought to enable these organisms to survive in more extreme environments, as ether bonds are more resistance to oxidation. Archaeol[21], a lipid with two 20 carbon chains ($A_{20,20}$), is ubiquitous in all archaea and is the only lipid phytanyl chain structure in extreme halophiles such as *Halobacterium salinarum* and *Halorubrum sodomense* (Fig 1.4). Some other halophilic archaea also contain a small proportion of $C_{20,25}$ and $C_{25,25}$ archaeols[22].

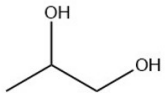
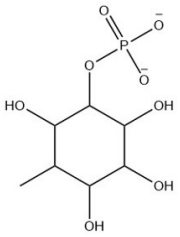
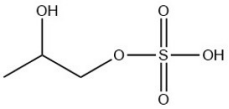
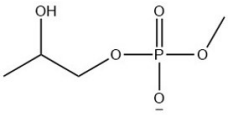


Archaeol

Figure 1.4: The general strucutre of an archaeabacterial phospholipid. Both C_{20} acyl chains are bound to the central glycerol by an ether bond. This is the only alkyl chain length in the membrane of the halophilic archaea *Halobacterium salinaurm* and *Halorubrum sodemense*. 'X' is replaced by the head group in specific archaeol species.

The monopolar archaeol lipids have been found to have different headgroups depending on the type of archaea[23]. Within their membrane, the halophilic archaea *H. salinarum* and *H. sodomense* contain an abundance of phospholipids and glycolipids, with a smaller portion of an analogue of cardiolipin (Fig 1.3)[23] which is ether-linked in the archaea[24]. These extreme halophiles also include sulphonated phospholipids and sulphonated glycolipids. Unlike in bacteria; ethanolamine and serine are absent in archaea[25], although inositol has been reported[26]. Common head groups of the archaeabacterial phospholipids are shown in Table 1.2. There have also been reports of glycophospholipid dimers found in the membrane[24]. These dimer lipids are made up of two units of archaeol bound together to span the membrane as a single, bipolar lipid, effectively forming a monolayer in the membrane. It is thought that these stronger, covalently bound lipids allow archaea to survive in their extreme environments. Other non-polar compounds present in the archaeal membrane include squalene (C_{30}) as the most abundant isoprenoid, with additional carotenoids (C_{40}) and bacterioruberin (C_{50})[27].

Table 1.2: Common head groups of phospholipids found in archaea. Head groups replace 'X' on Fig 1.4.

Phospholipid Class	Head Group
Phosphatidic Acid (PA)	
Phosphatidylglycerol (PG)	
Phosphatidylinositol 3-Phosphate (PI_3P)	
Phosphatidylglycerosulphate (PGS)	
Phosphatidylglycerophosphate Methyl ester (PGP-Me)	

1.1.2 Membrane Proteins

Membrane proteins constitute around 50% of the total dry weight of the membrane[5] and can be classified as either peripheral or integral. Peripheral proteins are found attached to one leaflet of the bilayer, without passing completely through the hydrophobic core of the membrane. These proteins mainly interact with the

membrane via electrostatic interactions with the polar head groups of the lipid species, with some hydrophobic interaction with the non-polar membrane core to increase stability. Integral membrane proteins are those which can only be removed from the membrane using detergents, and mostly span the hydrophobic interior of the membrane. The proteins present hydrophobic amino acid sidechains to the membrane interior which allows energetically favourable interactions with the lipids, whilst charged residues on the faces of the protein are able to interact with the headgroup moieties and the aqueous environment. Transmembrane proteins are classified as proteins that completely span the hydrophobic interior of the bilayer, either in a single pass or with multiple transmembrane segments.

Membrane proteins carry out a multitude of diverse functions in the cell including energy production, hormone signalling, neuronal transmission and adhesion[28]. It has been estimated that up to 40% of all genes in the human genome encode integral membrane proteins[29]. Because of their functional importance in almost all aspects of cellular life, they are an important target for drug development and disease treatment, accounting for 70% of all known drug targets[30]. With the continued development of drug design and screening assays, there has been a shift to targeted drug discovery, using protein structures to gain a deeper understanding of the specific interactions between drug ligand and target. However, with a total of 142,000 entries in the protein data bank (PDB), under 5,000 of which are for membrane proteins[31], there is still much information to be gathered for potential therapeutic targets.

1.2 Membrane Protein Reconstiution

Detergents allow for the study of membrane proteins by incorporation into a micelle of simple amphipathic molecules that mimics the hydrophobic interactions between the protein and its native membrane environment. One of the major problems with detergent reconstitution however is the inherent simplicity of the systems. As detailed in Section 1.1.1 above, the lipid membrane is highly complex, with many intricate interactions and a multitude of components. It is therefore important to

consider the differences in environment between the simple detergent and the native cellular environment of the protein of interest.

Selection of the most suitable detergent for membrane protein reconstitution can only be fully realised experimentally, although there are some general rules regarding class of detergent that should be considered. The most applicable detergent for stable membrane protein reconstiution may impede subsequent specific experimental procedures[32–34]. Futhermore, for more structural studies such as protein crystallography and mass spectrometry, the protein of interest benefits from the stripping of native membrane lipids, termed delipidation, often leaving it non-functional[35]. Removal of lipids can be achieved with organic solvent extraction, although it usually results in the loss of native protein structure.

Detergents are classed as either non-ionic, zwitterionic, anionic or cationic depending on their structure and physical properties. The properites of specific detergents varies depending on the strucutre and size of the polar headgroup as well as the length of the hydrocarbon chain to which it is attached. The critical micellar concentration (CMC) decreases by approximately one order of magnitude for every two methylene units added to the hydrocarbon chain [36]. The aggregation number of a detergent, that is the number of detergent monomers in a micelle, increases with increasing hydrocarbon chain length as more monomers are required to construct the micelle.

Typically, ionic detergents such as Sodium Dodecyl Sulfate (SDS) are too harsh for protein reconstitution and leave the protein denatured. Zwitterionic detergents are more successful at stabilising a membrane protein in a micelle, however the majority of purification protocols use non-ionic detergents for purification purposes. In general, detergents with a longer hydrocarbon chain and larger head group have milder properties, sometimes not fully stripping the native lipids from the protein. The non-ionic detergent Triton^{TM} X-100 is considered to be a mild detergent and rarely denatures the protein it is used with. Unlike other Poly(ethylene glycol) (PEG) based detergents, it has a relatively bulky branched alkyl group attached to its aromatic ring structure (Table 1.3). As a biproduct of the synthesis of Triton^{TM}

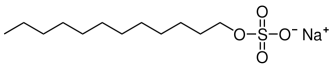
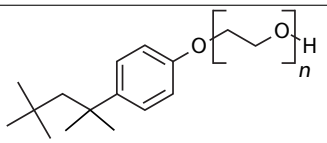
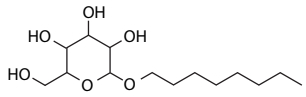
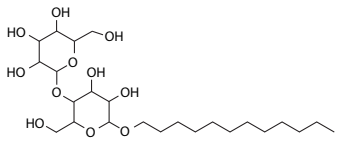
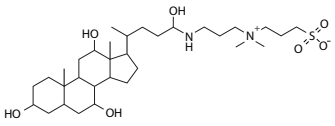
X-100, there are a number of peroxide contaminants that could interfere with the protein system[37].

Glycosidic non-ionic detergents are frequently used in cyrstallographic studies of membrane proteins[38]. The most common of this class of detergents are the alkyl β-D-glucosides and alkyl β-D-matlosides (Table 1.3), the chain lengths of which contain a total of between 8 and 14 carbons. The head groups of these detergents are based on carbohydrate ring structures, the glucosides are based on glucose, and the maltosides are based on maltose. This unique headgroup construction makes them particularly suitable for membrane protein purification as there is an aqueous-like microenvironment at the micelle interface, a feature that is not present in other non-ionic detergents[39].

The zwitterionic detergent 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate (CHAPS) (Table 1.3) is based on a cyclic head group but with sterol-based hydrophobic moieties. CHAPS has been shown to efficiently reconstitute membrane proteins without denaturing them.

The lipid-protein interactions present in the plasma membrane can be mimicked by the detergent but are unlikely to fully satisfy the interactions found in the lipid bilayer. The lateral pressures from the tight packing of lipids and proteins against the target protein that are experienced in the bilayer environment are lost in the simple detergent system[40]. Another major difficulty with detergent use compared to native membrane preparations is the inability, thus far, to develop asymmetric detergent environments for the protein. The lipid and protein content of each of the bilayers of the cell membrane is distinct and variable, allowing for precise interactions between the protein of interest and its environment. This must be considered when assaying protein activity, as some of the sophisticated interactions are not present in detergent which could affect the protein folding and leave the protein non-functional *in vitro*.

Table 1.3: Common detergents used for membrane protein reconstitution. Critical micellar concentrations (CMCs) are given in[41].

Detergent	Structure	CMC (mM)
Ionic		
Sodium dodecyl sulfate (SDS)		8.2
Non-ionic		
Triton^{TM} X-100		0.3
n-Octyl-β-D-glucopyranoside (OG)		25
n-Dodecyl-β-D-matloside (DDM)		0.15
Zwitterionic		
3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate (CHAPS)		6.5

1.3 Retinal-containing proteins

Eukaryotic rhodopsin proteins are light-sensitive 7-TM G protein-coupled receptors (GPCRs) that are made up of two components, the opsin protein and a prosthetic group, retinal. The retinal molecule is bound via a Schiff Base linkage to a lysine residue on helix G of the protein. Rhodopsin is the pigment in the rod cells of the retina and is responsible for the detection of photons.

Microbial rhodopsins can be found in the membranes of bacterial, archaeal and eukaryotic cells, although they occur more rarely in complex multicellular organisms[42]. These proteins are not GPCRs, but instead use the absorption of a photon to facilitate ion translocation across the membrane or to allow chemotaxis[43].

Because of their ability to generate a proton gradient when activated by a specific

wavelength of light, microbial retinal-containing proteins have a number of applications for imaging and electrical sensors. In the field of optogenetics, microbial rhodopsins can be used to track the propagation of an action potential in neurons and enhance or block a signal by depolarising or hyperpolarising the membrane potential. Wavelength-specific bio-sensors can be created using microbial rhodopsins to absorb light and generate a signal due to their ion translocation which could be converted into an electrical readout.

1.3.1 Bacteriorhodopsin

Bacteriorhodopsin (bR) is a seven transmembrane protein made up of 248 amino acids with a total molecular weight of 27 kDa[44]. It functions as a light driven proton pump found in the purple membrane (PM) of the halophilic archaea *Halobacterium salinarum*. These cells grow in high salt environments (4 M NaCl) where they utilise sensory rhodopsins (SRI and SRII) and their respective phosphorylation signalling cascades to cause attractant and repellent phototaxis, respectively[45].
bR is the only protein in the purple membrane (which can cover up to 80% of the total membrane surface[46]) and is arranged in a 2D-trimeric lattice that makes it suitable for diffraction analysis; it was the first membrane protein to have its structure solved[47], now with high resolution 3D X-ray crystallographic structures available for a number of the different protein intermediates (Fig 1.5A). In patches of the purple membrane, the transmembrane proton gradient generated by bR can then be used by ATP synthase to generate ATP from ADP and P_i.

The purple membrane is coloured due to the retinal chromophore found in the core of bR (Fig 1.5B), which is bound via a Schiff Base linkage to Lys216 on helix G. In the ground state, the retinal is in its all-*trans* form. Upon light-induced isomerisation to the 13-*cis* isomer, the protein undergoes a series of conformational changed, and the K216 residue is deprotonated. This sequential transfer of a proton between amino acid sidechains is the mechanism behind the pumping of protons to the outside of the cell. The different isomers of retinal absorb at differing

wavelengths. Dark-adapted bR contains an equimolar ratio of 13-*cis* (λ_{max}= 558 nm) and all-*trans* retinal (λ_{max}= 568 nm). Upon illumination, light-adapted bR, absorbing maximally at 568 nm[48], undergoes the photoreaction cycle resulting in proton translocation.

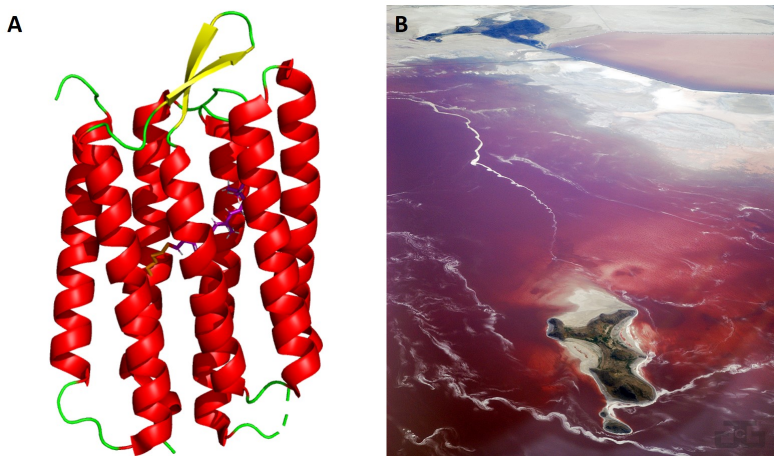


Figure 1.5: A. The crystal structure of a single bacteriorhodopsin monomer. The protein is coloured by secondary structure, the retinal chromophore is shown in *purple* and the Schiff Base linkage to residue K216 is shown in *orange*.Generated from PDB ID: 1M0L. **B.** Gunnison Island in the Great Salt Lake, Utah, USA. The purple colour is due to the abundance of the halophilic archaea *Halobacterium salinarum* and the bacteriorhodopsin present in the purple membrane (Photo Credit: Flickr[49]).

Bacteriorhodopsin is expressed in the purple membrane as a homo-trimer. The monomers make extensive interactions with the other two proteins in the trimer as well as with lipids such phosphatidylglycerophosphate Methyl ester (PGP-Me) and squalene. bR has been reported to mediate the production of the pigment bacterioruberin. This pigment molecule shares a precursor with retinal and was shown to accumulate with a bR deletion mutant of *Halobacterium salinaurm* (Δbop)[50].

Light-induced isomerisation of the ground state all-*trans* retinal results in photo-induced isomerisation to the 13-*cis* retinal. This is followed by proton transfer from the Schiff base to an acceptor residue (Asp85) on the extracellular side. The

Schiff base then switches its orientation to face a donor residue (Asp96) on the intracellular side of the proton channel within the protein. Asp96 accepts a proton from the cytoplasm and transfers it to the Schiff base, whilst the retinal isomerises back to the all-*trans* state. The final step of the photocycle returns bR back to the ground state, with the release of a proton from the proton-release complex (Gln194 and Gln204) into the extracellular environment. The intermediate states all have overlapping spectra between 550 and 630 nm, but the M state (in which the Schiff Base is deprotonated) has a characteristic absorption maximum at around 410 nm (Fig 1.6)[51].

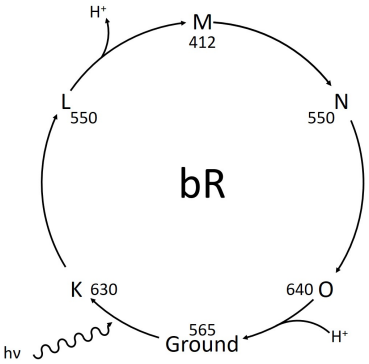


Figure 1.6: The photocycle of bacteriorhodopsin with absorbance maxima of each photo-intermediate. Adapted from[52].

1.3.2 Archaerhodopsin 3

Archaerhodopsin 3 (AR3), a homologue of bR, but much less well characterised, is also a 7-TM protein, acting as a proton pump in a different halophilic archaea, *Halorubrum sodomense* where the protein is located in a specific membrane, much the same as bR, called the claret membrane. AR3 is made of 252 amino acids, contributing to a molecular weight of 27.2 kDa. The sequence identity between AR3 and bR is 59%, although within the retinal binding pocket it is much more highly conserved. There is also considerable sequence homology between archaerhodopsins AR1, 2 and 3. Fig 1.7 shows a sequence alignment between bR

and the archaerhodopsins. All archaerhodopsins contain a retinal chromophore at its centre, but as a result of slight differences in amino acid composition in the retinal pocket, the delipidated protein absorbs light at a wavelength which is shifted realtive to bR (λ_{max}= 550 nm).

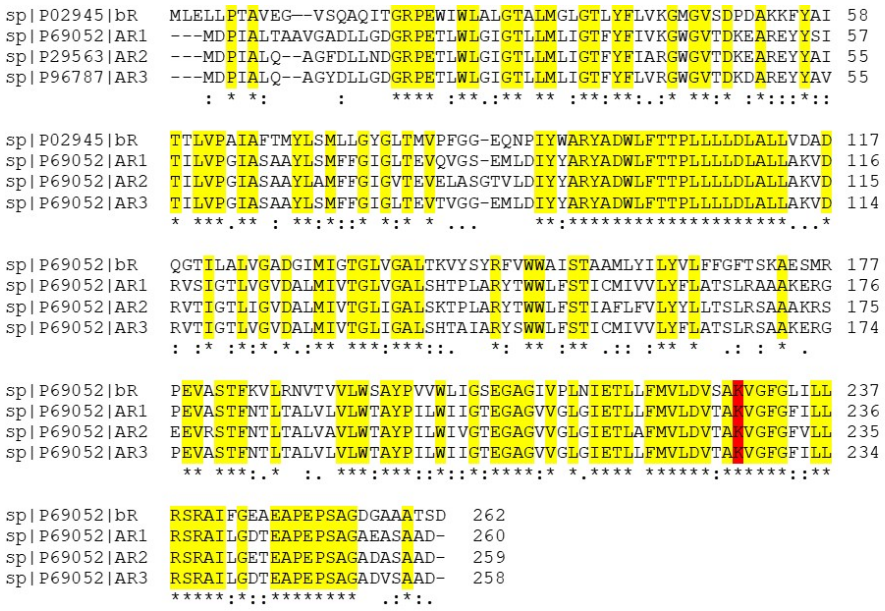


Figure 1.7: Sequence alignment between bacteriorhodopsin (bR) and archaerhodopsins AR1, AR2 and AR3 using the UniProt sequence alignment tool. Conserved residues are shown in *yellow*. Schiff Base Lysine shown in *red*.

Similarly to bR, AR3 is thought to exist as a homo-trimer in the claret membrane of *Halorubrum sodomense*. However, the pigment bacterioruberin is thought to be present at the interfaces between AR3 monomers. This has been shown in AR2 to bind to hydrophobic residues between Helix A and E of the two subunits and at a location called the omega loop (Fig 1.8)[53]. Comparing monomeric and trimeric AR2 crystal structures, it has been found that with the addition of bacterioruberin, the proton release channel is in a more closed conformation, becoming stabilised in the trimeric form[53].

As observed in bR, the photocycle of AR3 progresses through a number of intermediates that correspond with the movement of protons from the intracellular to the

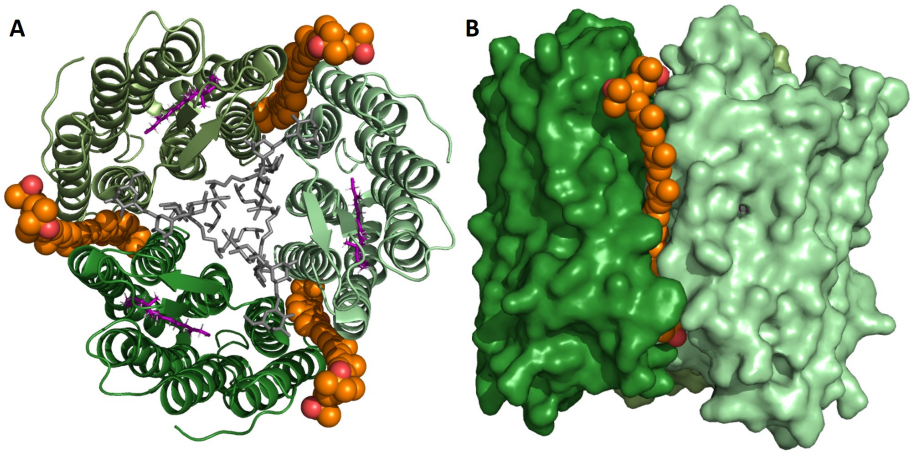


Figure 1.8: The trimeric organisation of archaerhodopsin 2 (AR2). Individual subunits are shown in *green*. The retinal is shown in *purple*. Lipids co-crystallised with protein are shown in *grey*. Bacterioruberin is shown in *orange*. **A.** A top view of each subunit with bacterioruberin clearly binding at interface between momoners. **B.** Side view of monomer A and B, bacterioruberin is nestled in the interface and makes extensive hydrophobic interactions with the protein residues. PDB ID: 2Z55. Apadpted from[53].

extracellular environment via two 'half channels' separated by the Schiff Base lysine. Each photo-intermediate corresponds to a distinct protonation state of the Schiff Base (K226). The M state signifies the deprotonation of the Schiff Base, the decay of this intermediate corresponds to the reprotonation at the N state. A fluorescent photo-intermeidate, the Q state, has recently been discovered[54]. This allows the protein to release a fluorescent signal if it is further excited with an appropriate wavelength of light. Kinetic measurements of the AR3 photocycle suggest a much faster recovery time than bR, allowing a faster generation of a proton gradient across the membrane (Fig 1.9)[54, 55].

For many applications, such as optogenetics and molecular bio-sensing, proteins that absorb light at distinct wavelengths whilst still maintaining their native functionality are needed. Archaerhodopsin 3 has an absorbance maximum at a different wavelength in its native environment compared to bR and other microbial rhodopsin proteins. However, due to the kinetic differences between bR and AR3, it is thought that AR3 may be more suited to a number of applications in optogenetics and as a single molecule bio-sensor. To generate distinct mutants of AR3 which can

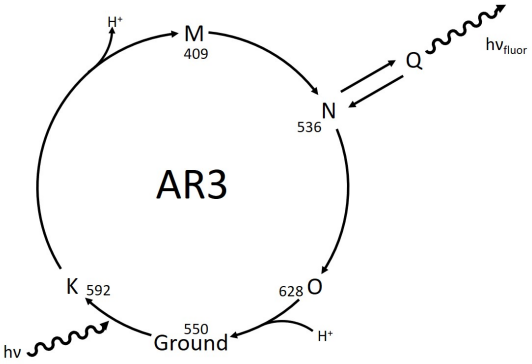


Figure 1.9: The photocycle of archerhodopsin 3 with absorbance maxima of each photo-intermediate. Adapted from[54, 55].

absorb light at different wavelengths, the protein environment around the retinal can be mutated. A number of studies have been published exploring the effect of both directed evolution[56, 57] to create a library of rationally designed mutants, as well as with direct mutagenesis targeting residues located within the retinal environment to alter the polarity of the retinal pocket changing the interactions between the chromophore and protein[58–60] (Fig 1.10).

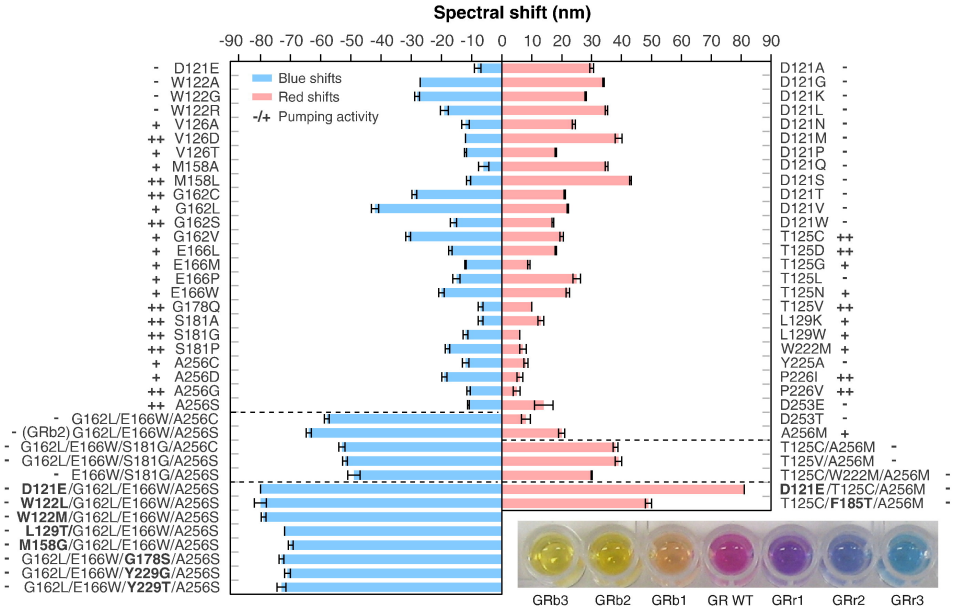


Figure 1.10: Directed evolution of *Gloeobacter violaceus* rhodopsin (GR) to produce spectrally shifted mutants. Shifts reported with respect to λ_{max} of WT GR (538 nm) with standard error of 3 repeats. Adapted from[57].

Depending on the location of the mutation, whether it is in close proximity to the β-ionone ring of the retinal or towards the Schiff Base, there is expected to be an opposing effect on the tuning of absorbance maxima. By introducing a more hydrophobic residue within 5 Å of the Schiff Base, a bathochromic shift of the absorbance maximum is seen. This tunes the protein to absorb light at a longer wavelength, closer to the near-infrared wavelengths of the absorbance spectrum. Introduction of a more polar group near the β-ionone ring would also have this effect. An increase in hydrophobicity in the portion of the retinal pocket encompassing the β-ionone ring has the effect of tuning the absorbance maxima of AR3 towards the shorter wavelengths of light (Fig 1.11)[61]. Generating multiple mutations in a single protein can allow for combinative shifting of the wavelength depending on the individual effects of the added mutations. By using hypsochromic-shifting and bathochromic-shifting mutations it is possible to generate AR3 proteins with a wide range of distinct absorption maxima for applications in biotechnology.

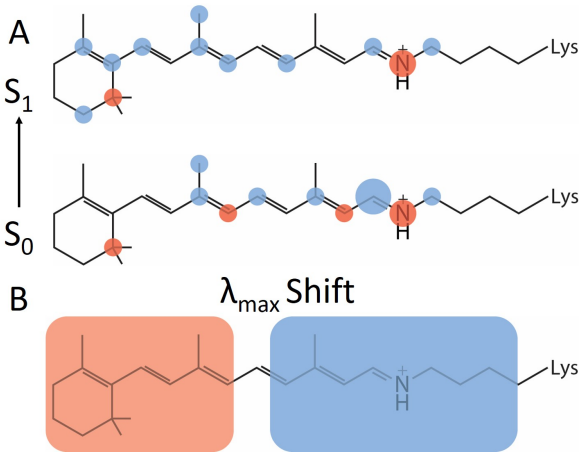


Figure 1.11: Electrostatic effects of mutations on retinal carbons. **A.** The ground state (S_0) and excited state (S_1) of the protonated Schiff Base retinal. The change in charge distriubtion along the chromophore after excitation. **B.** Schematic representation of the effect of decreasing the hydrophobiciy in areas of the retinal pocket. Increasing polarity near the Schiff base causes blue-shifted spectra, near the β-ionone ring causes red-shifted spectra. An inversion of the displayed mutations will have an opposing effect. Adapted from[61].

1.4 Molecular Bioelectronics

The field of bio-sensor design aims to exploit biological molecules as efficient, compact and precise sensors for both biological systems and physical properties[62]. The limiting factor when developing silicon-based transistors is their continued miniaturisation[63]. The miniaturisation of silicon based devices occurs at the insulator, a gate made of silicon oxide, which is the most important aspect to consider when developing a new appliance[64]. The insulator can only be scaled down to a point at which it no longer functions as an insulator in the device. It has been predicted that these silicon-based semiconductors will reach the molecular scale by 2030[65, 66], as predicted by Moore's Law[67] beyond which no miniaturisation is possible. It is therefore necessary to find an alternative approach to conventional silicon devices that can be used at even smaller scales.

Protein engineering can be used to develop molecular-scale signal processors, in contrast to the large-scale manipulation currently used to generate integrated circuits. The precision achievable with genetic manipulation cannot yet be met by current methods. Liquid crystal displays (LCDs) are possibly the best known application of molecular bioelectronics currently available. Bioelectronics is a distinct field within molecular electronics that uses both native and modified biological materials in electronic devices. This offers a high level of control due to the ability to specifically alter pre-existing biological molecules to be more appropriate for their required application. The eventual goal of this research is to develop nanoscale processors from a few (or single) biological molecules which could be detected and be able to transfer an electrical signal.

Both cytochrome c and azurin[68] have been attached to gold surfaces to study their electron transfer mechanisms *in vitro*. These proteins contain abundant surface cysteine residues, which can be used to form gold-thiol bonds to the conductive surface, which, in turn, allow the protein to self-assemble into monolayers. Other immobilisation methods can be used, such as electrostatic and hydrophobic interactions, however chemisorption with a covalent bond between protein and gold is strong[69] and allows for a homogenous orientation of immobilised protein[70].

Bacteriorhodopsin is the smallest naturally occuring photosynthetic system, comprising a single protein unit that allows for the generation of a proton gradient upon activation by light. bR has previously been used as a molecular-scale photoswitch once bound to a gold surface[71–73]. Archaerhodopsin 3 has a faster photocycle meaning it can recover after activation more rapidly, and is therefore a prime candidate for molecular signal processing. This assumption is the basis for much of the work detailed in this thesis.

One of the crucial properties of both bR and AR3 is their ability to be immobilised onto a conductive surface as was performed with cytochrome c and azurin. In these earlier examples, the abundance of surface cysteine residues found in the protein's native conformation allowed their adherence onto the surface by a gold-thiol bond with the cysteine side chain. AR3 contains only a single natural cysteine molecule in its mature sequence, which is buried within the 5^{th} transmembrane helix (helix E). It is therefore necessary to introduce an additional cysteine residue onto an acesible loop of the protein to allow its immobilisation. This has already been accomplished with bR, where work was carried out to locate the optimal position for immobilisation, whilst not perturbing the protein function[71–73]. A position on the intracellular EF loop was chosen for cysteine mutation. This specific location was preferred due to the conformational changes that the EF loop undergoes during the photocycle. It has been reported that, during the N state of the photocycle in bR, there is a conformational change at the intracellular interface between helices E, F and G to allow for better Schiff Base reprotonation from Asp96[74]. In bR the cysteine mutant M163C was chosen due to its optimal position for immobilisation, offering strong gold-thiol bonding, whilst also being located in the sensitive EF loop to allow for optimal photocycle progression[75].

1.5 Optogenetics

The study of neuronal circuits requires an ability to observe, initiate and control nerve impulses inside the organism of interest. Working with invertebrate model organisms with simple neuronal circuitry, single neurons that could be individually

investigated were identified, leading to the understanding of behavioural patterns in these organisms. However, vertebrate brains are several orders of magnitude more complex, with the rat brain containing upwards of hundreds of millions of neurons and the human brain hundreds of billions of neural cells, of many different types. Each neuron type has its own complex and refined expression motif, function, and localisation creating a heterologous, associative, network. Classical techniques do not allow for precise assessment of individual neuron types, only regions of the brain.

Optogenetics has allowed for the integration of new techniques into specific nerve cells for imaging in more complex systems. Using viral systems to incorporate the required protein for the specific target cell, new light-sensitive proteins can be expressed in the model organism and used for neural stimulation or silencing, allowing the propagation of action potentials to be accurately followed throughout the complex network. There are some drawbacks to this approach, chiefly the size of the protein which may be inserted by the virus is limited by the length of the nucleic acid which can be packed into the virion. Commonly used lentivirus and adeno-associated virus (AAV) vectors are limited to 10 kb and 5 kb constructs, respectively[76]. This reduces the promoter size for the protein gene and thus reduces the integration of the protein into the nerve cells. The use of transgenic animals can allow for the specific and complete expression of optogenetic modulators with extreme precision. Individual cell types can be targeted with different proteins, allowing for more detailed visualisation of neural networks in complex organisms.

Bacteriorhodopsin has previously been used as a neural silencer, due to its ability to pump protons across the membrane of a neuron resulting in hyperpolarisation and the inhibition of the action potential. Compared to other light-activated pumps, such as Cl^- pumps, proton pumps have a few advantages, most important of which is their faster recovery time[77–79]. AR3 is therefore believed to be even more advantageous than bR, due to its faster photocycle and higher intensity[77]. There have also been published studies investigating the photocurrent differences between bR and AR3 leading to the discovery that AR3 has a photocurrent approximately 10 fold

higher than bR, meaning it causes a much more intense signal upon activation (Fig 1.12)[77]. There is now a push to develop mutant proteins with shifted absorbance wavelengths for medical applications. Red-shifting AR3 mutants will allow their use with non-invasive light penetration in the brain. Blue-shifted mutants are required to allow multiple proteins (with non-overlapping absorption spectra) to be used in the same system to probe multiple neuronal pathways. Using a combination of tuned mutants, it is possible to use their respective maximal wavelengths to precisely and selectively activate one protein over another. Because of its kinetic advantages over other proton pumps, AR3 is now routinely used in the field of optogenetics, however, the development of photostable mutants with distinct absorption spectra has not yet been fully realised.

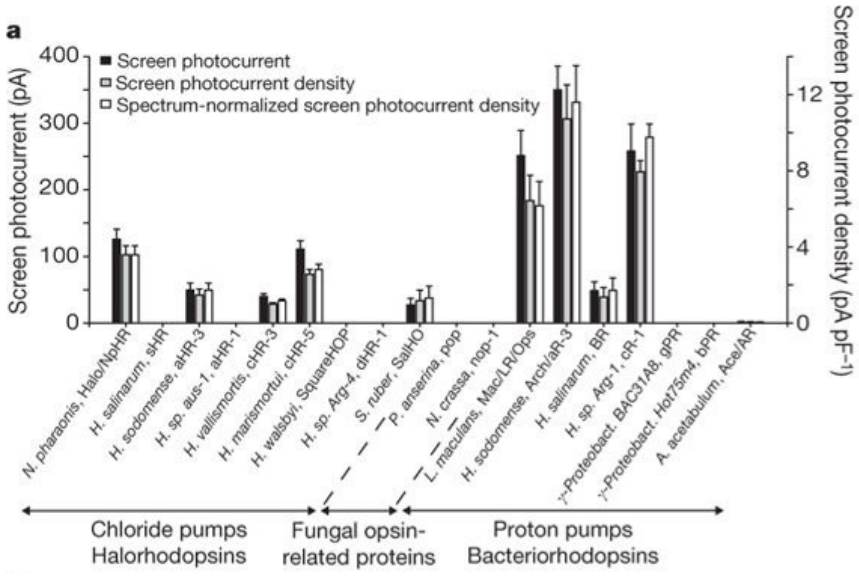


Figure 1.12: The photocurrent of common proteins used for optogenetic experiments. It can be seen that archaerhodopsin 3 has a photocurrent of around 350 pA, compared to bR which has only approximately 50 pA. The same magnitude of difference is seen with the screen photocurrent density (pA PF^{-1}). Adapted from[77].

1.6 Thesis aims

1. The first aim of the work in this thesis is to generate a library of AR3 mutants with tuned absorbance spectra by mutating key amino acids around the retinal chromophore. These protein constructs will then transformed into *E. coli* strain BL21 for expression.

2. Following successful expression in the bacterial model, the proteins will be grown at a large scale and purified into n-Octyl-β-D-glucopyranoside (OG) micelles. The absorbance spectrum of each mutant will recorded to assess the production of both hypsochromic- and bathochromic-shifted absorbances compared to *wt*AR3.

3. For the development of AR3 as a molecular bio-sensor, immobilisation of the protein onto a conductive surface must be achieved. To allow for covalent adsorption, two locations were chosen for cysteine mutation, Arg173 and Pro175. Both of these positions are on the EF loop, allowing for favourable and homogenous binding of the protein onto gold.

4. In the last section of this work AR3 was assessed for its suitability as a bio-sensor. Atomic force microscopy and visible CD will characterise the oligomeric state of AR3 in detergent. Solid-phase UV-Vis spectroscopy proves the adsorption of the cysteine mutants onto a conductive surface for downstream biotechnological applications.

2
Materials and Methods

2.1 Materials

Template plasmid WT Arch-3 in pET28b [addgene ID: 58487][54] was purchased from addgene[80]. The Plasmid contains a KanR kanamycin resistance gene and an AR3 open reading frame (ORF) with an additional fused poly-His tag (His_6). Detergent stocks (n-Octyl-β-D-glucopyranoside; n-Dodecyl-β-D-matloside and 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate) were purchased from Glycon Biochemicals (Germany). Triton^{TM} X-100 was purchased from Sigma (UK). All detergents were used at concentrations above their respective critical micellar concentrations (CMCs). Reagents, unless otherwise stated, were used without further modification or deviation from package instructions.

2.2 Polymerase Chain Reaction

Plasmid pET28b with an added gene for *wt*AR3 was used for site-directed mutagenesis. Mutant primers were designed with reference to the Thermo Fisher primer design and tools[81]. PCR was performed with reagents from the Agilent Technologies QuikChange II Site-Directed Mutagenesis Kit (Agilent, USA). The PCR programme for competent *E. coli* cell incorporation is shown in Table 2.1. Post-reaction, the samples were incubated with 1 µl DpnI (37 ℃, 1 hr) to digest the parental strand.

Table 2.1: PCR programme for site-directed mutageneis of the AR3 gene.

Cycles	Temperature (℃)	Time
1x	95	1 min
15x	95	30 s
	55	1 min
	68	10 min
1x	68	10 min
1x	4	-

2.3 Transformation

Transformations were performed in *E. coli* strain BL21 to allow for increased protein expression. Competent cell transformation protocol was adapted from PGC Scientifics Corporation[82]. In brief, BL21 cells were thawed on ice before addition of 5 μl PCR DNA. Tubes were incubated (on ice, 30 min) before heat shock (42 ℃, 45 s). Tubes were replaced on ice (5 min) and 500 μl Luria Broth (LB) was added and tubes were incubated in a shaking incubator (37 ℃, 180 rpm, 60 min). Cultures were spread on agar plates supplemented with 100 mg/ml kanamycin. Plates were incubated at 37 ℃ overnight.

Successfully grown colonies were picked and added to 5 ml LB with an additional 100 mg/ml kanamycin and were grown with shaking (37 ℃, 180 rpm, 4 hrs). Glycerol stocks were made with 200 μl 90% glycerol and 800 μl liquid culture for storage at -80 ℃.

Minipreps were carried out on liquid cultures as described in the Qiagen QIAprep Spin Miniprep Kit (Qiagen, Netherlands). A 4 ml sample was initially centrifuged, DNA was eluted with 30 μl elution buffer to increase DNA extraction concentration. DNA concentration was determined by Nanodrop analysis, blanked against elution buffer. Samples were sent to Source Bioscience[83] for sequencing to confirm uptake of plasmid and presence of mutation.

2.4 Preparation of AR3 expressed in *Escherichia coli*

AR3 glycerol stock was grown in 20 ml LB with 20 μl kanamycin stock at a concentration of 100 mg/ml. Pre-cultures were grown (37 ℃, 180 rpm, overnight). LB media flasks with 200 μl additional trace metals (50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$ and $ZnSO_4$, 2mM each $CoCl_2$, $CuCl_2$, $NiCl_2$, Na_2MoO_4, Na_2SeO_3 and H_3BO_3) induced with pre-cultures and were grown until OD_{600}=0.5. Addition of 0.5 mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG) and 5μM all-*trans* retinal (Sigma, UK) were used to induce protein production. Flasks were incubated for a further 4 hrs before harvesting. Cell cultures were harvested by centrifugation (8,000 *g*, 4 ℃, 20 min). Pellets were frozen in liquid Nitrogen (LN_2) and stored at -80 ℃ until required.

2.5 Detergent purification of AR3

Purification protocols were adapted from[84]. Cell pellets were resuspended in 75 ml sonication buffer (150 mM NaH_2PO_4, 20 mM NaCl, 5 mM $MgCl_2$, pH 7) per litre of culture. Lysozyme (0.2 g), 60 μl of 1.5 mM Aprotinin and 400 μl of 300 mM DNase, 2 mM Leupeptin, 1.5 mM Pepstatin A and 1 M $MgCl_2$ were added to the solution and stirred (4 ℃, 2 hrs). The solution was sonicated with a tip sonicator with a 10 s on/off pulse, 5 minutes, at an amplitude of 50% and repeated three times for the 75 ml sample. Sonicated samples were homogenised in a glass/Teflon tissue homogeniser 5 to 10 times. The homogenised sample was centrifuged (70,000 g, 4 ℃, 1 hr) and resuspended in solubilisation buffer containing detergent (30 mM Na_2HPO_4, 20 mM NaH_2PO_4, 300 mM NaCl) and either 1.5% (*w/v*) OG, 0.3 mM Triton™ X-100, 1% (*w/v*) DDM or 0.75% (*w/v*) CHAPS at pH 7 at 4℃ overnight. The detergent-solubilised protein was centrifuged (70,000 *g*, 4 ℃, 1 hr) and the supernatant purified on an immobilised metal ion affinity chromatography (IMAC) column (GE Healthcare, UK).

For the DDM detergent samples, the solution was additionally diluted to a final

detergent percentage of 0.2% (*w/v*) in solubilisation buffer without detergent. The column was washed with five column volumes (CV) of ddH_2O before loading the protein. All samples were filtered through a 0.45 μm cellulose nitrate filter (Whatman, UK) and the solution was then loaded onto the HisTrap column at a rate of 1 ml/minute. The column was washed with 10-15 CV buffer (50 mM NaH_2PO_4, 100 mM NaCl, 50 mM imidazole, pH 7) with added detergent at previous ratios (0.2% (*w/v*) DDM used in chromatography buffers) until the OD_{280} became stable. Elution was performed with 5-10 CV elution buffer (50 mM NaH_2PO_4, 100 mM NaCl, 300 mM Imidazole, pH 7). Coloured fractions were pooled for UV-Vis spectroscopy. Purified fractions were stored at 4 ℃ for further experiments.

2.6 Lipodisq purification of AR3

Detergent-free purification of AR3 from *E. coli* membranes was performed using styrene maleic acid (SMA) polymer (3:1 ratio)[85]. Frozen *E. coli* membranes were resuspended in phosphate buffer (50 mM NaH_2PO_4, pH 7.4) to a concentration of 40 mg/ml. The solution was homogenised and SMA (5% (*w/v*), pH 8) was added and incubated at room temperature (RT) with gentle agitation for 2 hrs. The preparation was then centrifuged (100,000 *g*, 10 ℃, 45 minutes) and purified by chromatographic methods.

First, a HiTrap column was used to remove excess free SMA. The fractions from Size Exclusion Chromatography (SEC) were subsequently buffer exchanged into 20 mM NaH_2PO_4, 200 mM NaCl, pH 8 and were loaded onto a HisTrap column (GE Healthcare) and recirculated throught the column at a rate of 1.5 ml/min for 2 hours. To elute the sample, increasing concentrations of imidazole were used until the coloured fraction eluted.

2.7 Biochemical Characterisation

2.7.1 Agarose Electrophoresis

For visualisation of PCR and DNA samples, a 1% agarose gel was prepared by melting 0.5g agarose in 50 ml of TAE buffer (1 mM EDTA-Na_2,40mM Tris, 20mM acetic acid, pH 8.0). 3.5 μl ethidium bromide (EtBr) was added and the gel was left to set for 30 minutes. Samples were prepared by addition of 3 μl running buffer (3 ml glycerol, 25 mg bromophenol blue, 7 ml ddH_2O) to 20 μl sample. The gel container was filled with TAE buffer, samples were loaded and run at 110 V until the dye front had migrated to the end of the gel. Bands were visualised under UV light.

2.7.2 SDS Gel Electrophoresis

NuPAGE^{TM} 12% Bis-Tris gels used with MES SDS PAGE buffer (Novex, UK). Gel samples were prepared by combining 15 μl protein sample with 5 μl lithium dodecyl sulphate (LDS). Gels were run at 180 V, 400 mA to completion.

2.7.3 Native PAGE

Native PAGE was performed using a light blue cathode buffer (1 ml 20x cathode additive per 200 ml 1x NativePAGE running buffer) with 4-16% Bis-Tris gels (150 V, RT, 100 min).

Coomasie Blue Staining

InstantBlue^{TM} (Expedeon, USA) protein stain was used for staining gels with Coomasie blue dye. The gel was stained (RT, at least 15 minutes) with gentle agitation.

Silver Staining

Silver stain was used as a more sensitive method for protein staining. After electrophoresis, the gel was placed in fixative enhancer solution (100 ml methanol, 20 ml acetic acid, 20 ml fixative enhancer concentrate, 60 ml ddH_2O; 20 min) with gentle agitation. After fixing, the gel was rinsed (200 ml ddH_2O, 10 min) before

discarding and repeating with a second washing step (200 ml ddH_2O). Staining was carried out (17.5 ml ddH_2O, 2.5 ml each of silver complex solution, reduction moderator solution and image development reagent (Bio-Rad)). Just before use, development accelerator solution (25 ml), brought to room temperature, was added and the solution was mixed. Staining was allowed to continue until the desired band intensity was seen. The reaction was stopped by the additon of 5% acetic acid solution for 15 minutes.

2.7.4 Spectroscopy

UV-Vis spectroscopy was carried out on a Jasco V-630 UV-Vis spectrophotometer fitted with microcell attachment EMC-709. All analyses were corrected against respective solution buffers. AR3 spectra were recorded in the range 350-700 nm, at a scan speed of 100 nm/min using a 1 cm quartz cuvette.

Solid phase UV-Vis spectroscopy was carried out using a film holder attachment (FLH-740). Gold on glass substrate nanochips (NanoSPR, USA) were washed with ethanol, water and phosphate buffered saline (PBS) (30 mM Na_2HPO_4, 20 mM NaH_2PO_4, pH 7.4).The protein sample was prepared to 25 µM and left agitating with the gold chip (overnight, 4℃). The adsorbed gold chip was washed with PBS and its abosrbance spectrum collected between 400 and 700 nm, a baseline measurement was taken against a blank chip.

2.7.5 Circular Dichroism

Purified samples were buffer exchanged with Cl^- free PBS to remove imidazole and loaded onto Jasco J-815 CD Spectrometer. Spectra were measured at RT between 190 and 260 nm using a 1 mm cuvette. Protein secondary structure was predicted using the Jasco spectra manager software and calculated from the Dichroweb website[86] using the CDSSTR programme[87] with the membrane protein SMP180 set as the reference[88]. Visible CD spectra were measured between 400 and 700 nm at RT.

2.8 Atomic Force Microscopy

AFM was used to image AR3 after solubilisation into a detergent micelle. The protein was depositied on a mica substrate to estimate the surface coverage of both *wt*AR3 and a cysteine mutant, P175C. AFM imaging was performed with the Veeco Dimension 3100 system (Veeco, USA) and the data was analysed with the Nanoscope analysis programme (Bruker, USA).

Both samples were imaged after purification into an OG micelle (Section 2.5), utilising a mica substrate. The freshly cleaved substrate was incubated with the protein solution (5 μM in phosphate buffer) for 10 min, rinsed in phosphate buffer to remove non-specific binding, and dried under nitrogen. Imaging was performed in tapping mode in air using non-contact AFM probes (Tap300, Budget Sensors).

3

Results - Spectral Tuning of Archaerhodopsin 3

The amino acids surrounding the retinal chromophore in AR3 are well characterised by referencing the sequence alignment between AR3 and other microbial rhodopsins. With the recently elucidated crystal structure of AR3 at 1.3 Å (PDB ID: 6GUX), the location and interactions of these residues in the core of the protein can be further investigated. By referring to the crystal structure (Fig 3.1), the key residues that form the retinal binding pocket can be identified. With this knowledge, and previous studies in AR3 and homologous proteins, a number of residues could be identified that, when mutated, result in a shift of the absorbance spectra in AR3. As is represented in Fig 3.1, the residues surrounding the retinal chromophore form close contacts with each other as well as water molecules found within the chromophore pocket. Therefore, mutating these residues would likely affect the interactions between sidechains and between the amino acid and retinal.

3.1 PCR and Transformations

To allow for AR3 mutagenesis, four mutations in the retinal binding pocket were selected based on previous studies of AR3 or bR. Primers were designed (Table 3.1) for each mutation and used to introduce it into the template plasmid. Site-directed

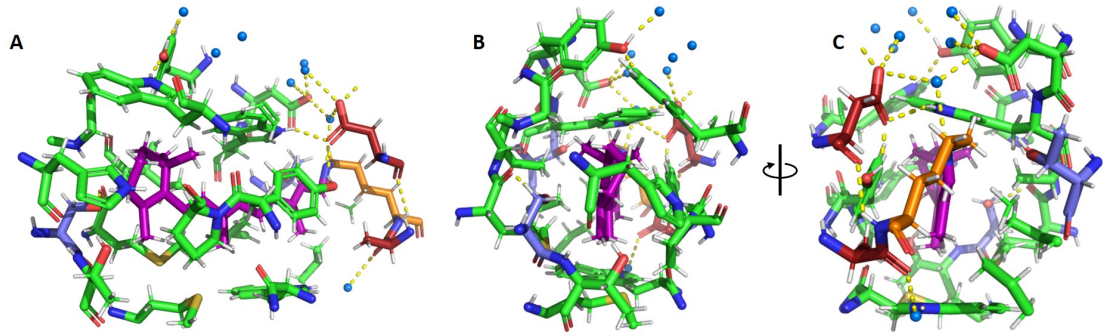


Figure 3.1: Molecular representation of the retinal pocket in AR3. Retinal shown in *purple*, K226 Schiff Base shown in *orange*. Water molecules shown as *blue spheres*. Polar contacts between side chains shown as *yellow lines*. **A.** All residues within 5 Å of AR3 are shown in *green*. The residues mutated in this report are represented as either *blue* or *red* depending on their expected hypsochromic or bathochromic shift, respectively. **B.** The retinal pocket oriented looking down the retinal from the β ionone ring. **C.** The same representation rotated 180° to show the pocket from the Schiff Base. Generated from PDB ID: 6GUX.

mutagenesis PCR was used to generate all AR3 mutants. Each mutant was carefully chosen due to its close proximity to the retinal chromophore.

T99C has been shown in AR3 to induce a blue-shift when used as a single mutation, but in conjunction with other hypsochromic mutants can switch to cause a fluorescent red shift[56]. The mutation S141C in bR has been demonstrated to cause a large blue shift, as much as 80 nm, but has also been shown, by inducing a pH change in solution in a light-dependent manner, to reduce proton pumping ability to 20% compared to the wild type protein[60]. The corresponding mutation in AR3, S151C, was chosen for its propensity to cause a large shift in the absorbance maxima. D212E, when studied in bR, caused a red shift of around 30 nm and absorption maximum became pH sensitive[89]. D222E, the corresponding AR3 mutant, was therefore selected. A225M has been shown to only cause a small bathochromic shift in absorbance, but this effect is amplified in combination with other mutants in double- or triple-mutated AR3[56].

For future immobilisation onto a conductive surface, two locations on the EF loop were selected as appropriate anchor points for cysteine point mutation. The two locations in AR3, R173 and P175, both remove a non-essential amino acid so as

to not perturb the proton pumping activity of AR3. Previously, the equivalent location between helix E and F was successfully used for immobilisation of bR[71–73].

Table 3.1: All primer sequences used for AR3 mutagenesis. Codons in *bold* signify point mutations. T99C and S151C are hypsochromic mutants, D222E and A225M are bathochromic.

Spectral Tuning Mutants	
T99C Forward	**T99C Reverse**
TGGCTGTTT**TGC**ACCCCACTT	AAGTGGGGT**GCA**AAACAGCCA
S151C Forward	**S151C Reverse**
CAGTTGGTGGTTGTTC**TGT**ACAATTTGC	GCAAATTGT**ACA**GAACAACCACCAACTG
D222E Forward	**D222E Reverse**
ATGGTGTTGGA**GGT**GACTGCC	GGCAGTC**ACC**TCCAACACCAT
A225M Forward	**A225M Reverse**
GGACGTGACT**ATG**AAGGTCGGCTTT	AAAGCCGACCTT**CAT**AGTCACGTCC
Cysteine Immobilisation Mutants	
R173C Forward	**R173C Reverse**
CAAAGGAG**TGC**GGCCCCGA	TCGGGGCC**GCA**CTCCTTTG
P175C Forward	**P175C Reverse**
CGGGGC**TGC**GAGGTGGCA	TGCCACCTC**GCA**GCCCCG

The single mutants were created using the wild type AR3 plasmid template. Once transformed and sequenced, the miniprepped mutant DNA was used as the template for production of subsequent double mutants. Table 3.2 shows all the mutants that were successfully transformed and sequenced, Fig 3.2 shows an agarose gel of succesfully transformed PCR mixtures. Not all of these mutants could be succesfully purified in OG or DDM detergents.

Table 3.2: All mutants successfully sequenced. Not all mutants could be succesfully purified in OG or DDM detergents. Mutants purified in OG are listed in Table 3.3, those in DDM are shown in Table 3.4.

Single Mutants			
T99C	S151C	D222E	A225M
R173C	P175C		
Double Mutants			
T99C-S151C	T99C-A225M		
T99C-R173C	T99C-P175C		
S151C-A225M			
S151C-P175C			
Triple Mutants			
T99C-S151C-P175C	T99C-A225M-P175C		
Quadruple Mutant			
T99C-S151C-A225M-P175C			

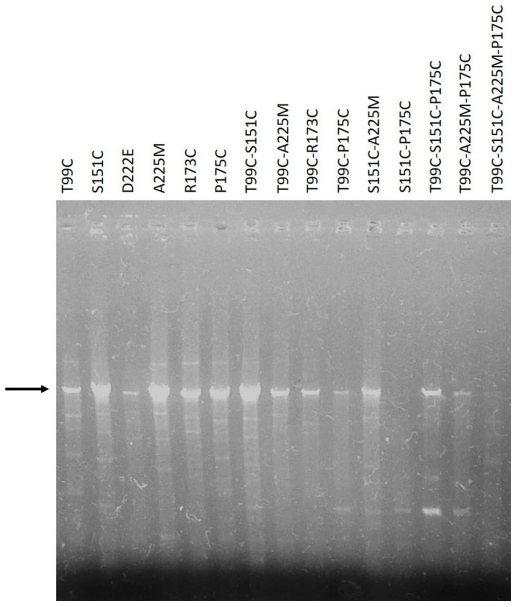


Figure 3.2: 1% agarose gel of sucessful PCR products. The single intense band is the PCR product, *arrow*, other bands seen in the wells may be due to non-specific amplified DNA.

3.2 OG Purification

The purification of membrane proteins requires their solubilisation from the bilayer of the host organism into a detergent micelle of known composition. Initially, in keeping with previous work in the Watts lab and other work on rhodopsin proteins which used n-Octyl-β-D-glucopyranoside (OG) as a detergent for solubilisiing rhodopsin from the host expression system, AR3 was purified from *E. coli* using OG as the detergent of choice. This allowed for successful purification of the *wt*AR3 and several mutants (Fig 3.3). The spectral shift obtained by mutation could then be assessed and rationalised against similar studies carried out on AR3 homologues. Using this approach, a total of four single mutants and two double mutants could be successfully purified and their respective absoprtion spectra recorded.

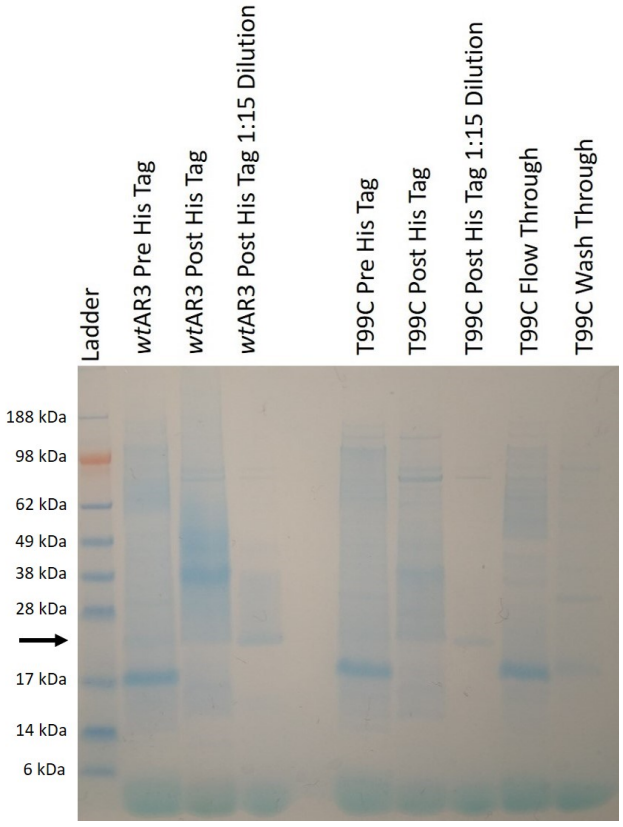


Figure 3.3: Coomasie blue stained SDS PAGE of *wt*AR3 and T99C purified sample at different points during the purification process. Samples diluted in ddH_2O. Monomeric AR3 can be seen in the band highlighted by the *arrow* at approximately 27 kDa.

The ability for the AR3 mutants to alter the retinal pocket in the hydrophobic core of the protein was assessed by UV-Vis spectroscopy. The absorbance spectra were measured, and the peaks compared to quantify the tuning introduced from the respective mutants. As observed in Fig 3.4 and Table 3.3 the wild type AR3 absorbs maximally at 554 nm. Mutating a specific threonine to a cysteine (T99C), which is in close proximity to the retinal Schiff base, results in a 14 nm blue shift. Increasing the hydrophobicity of the retinal pocket by introducing a methionine group for an alanine (A225M) shifts the spectra, slightly, towards a longer wavelength. This bathochromic shift of only 2 nm is consistent with previously published work investigating fluorescent mutants of AR3 from *E. coli* in DDM[56]. The cysteine mutants selected for immobilisation, R173C and P175C do not exhibit a large shift in their absorbance maxima. This is due to the relative distances of the side chains from the retinal chromophore. Arg173 and Pro175 are over 30 Å from the retinal, on the intracellular EF loop and thus any effect seen is not due to direct alteration of the retinal pocket environment but due to small structural changes in the whole protein caused by the introduction of a non-native amino acid near to the end of one of the transmembrane helices.

The double mutants exhibited here, T99C-R173C (TR) and T99C-P175C (TP), both combine one retinal pocket mutant and an accessible cysteine for future immobilisation. The combination of both single mutants does not add up constructively. Instead, there is a slightly diminished effect of the T99C mutation that is distinct from the effect of the single R173C and P175C mutations when compared to *wt*AR3.

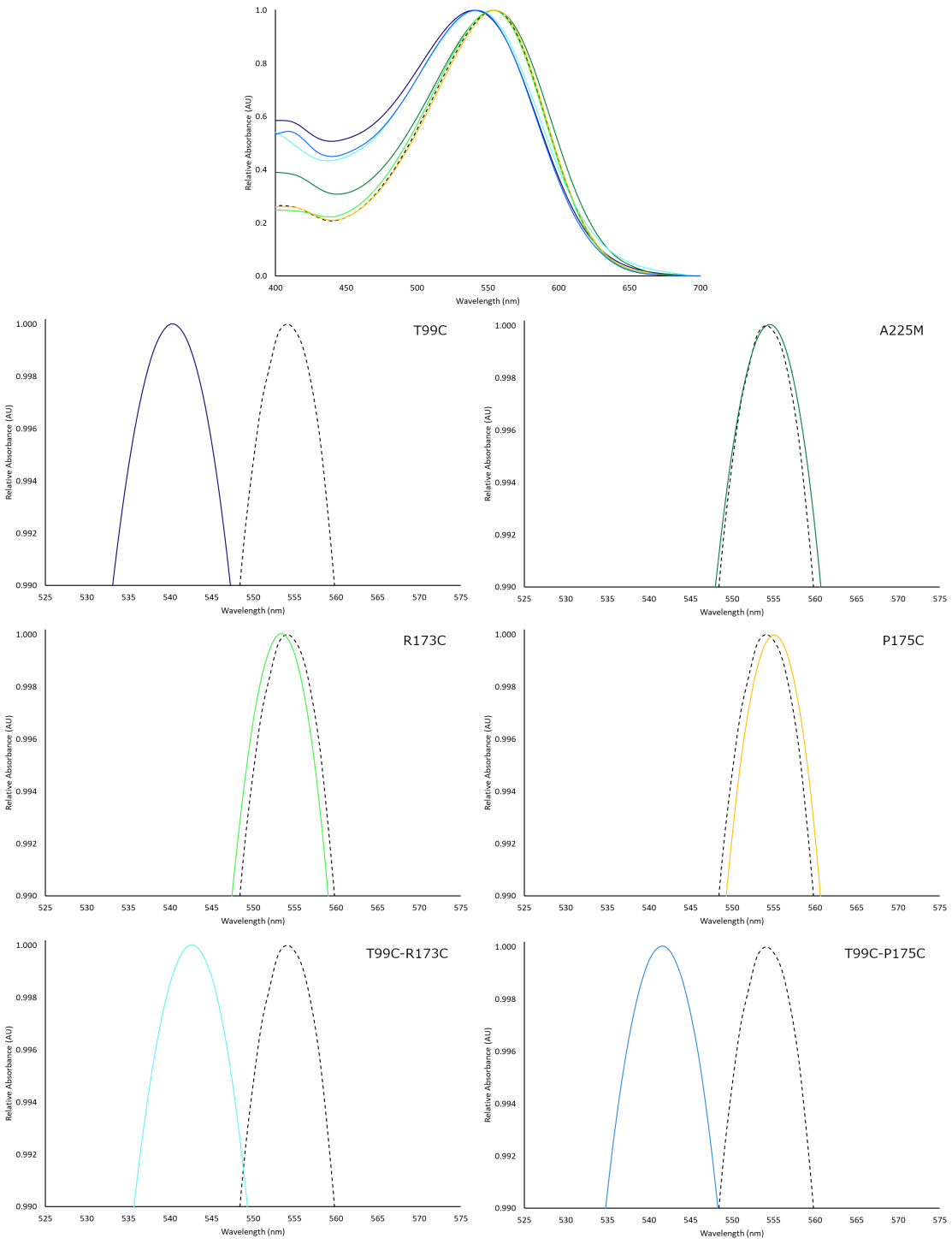


Figure 3.4: Absorption spectra of AR3 mutants purified from *E. coli* in OG (n-Octyl-β-D-Glucopyranoside) micelles. *Above,* A full UV-Vis spectrum, recorded between 400-700 nm. Each individual mutant peak is shown *below* compared to the *wt*AR3 spectrum (*dashed line*). For clarity, each peak is cut off at 0.990 AU. Maximum absorbances and shifts for each mutant are shown in Table 3.3.

Table 3.3: All Mutants successfully purified into OG micelles. Peak absorbances and shifts, relative to *wt*AR3 are shown.

Mutant	Peak Absorbance (nm)	Shift (nm)
*wt*AR3	554	-
T99C	540	-14
A225M	555	+1
R173C	553	-2
P175C	555	+1
T99C-R173C	543	-11
T99C-P175C	542	-12

3.3 S151C Purification

S151C was expected to give a large hypsochromic shift. However, during resolubilisation the sample turned from a purple to a yellow colour, indicative of the M state intermediate (Fig 1.9), absorbing predominantly around 400 nm (Fig 3.5 *Left*). Only when the same protocol was adapted to use different detergents (Triton^{TM} X-100, DDM or CHAPS) or a detergent-free approach was used, was a stable ground-state protein purified (Fig 3.5 *Right*).

Figure 3.5: S151C purification showing 'M' state and ground state. *Left.* S151C purified post HisTrap column exhibiting the characteristic yellow colour associated the M state photo-intermediate.*Right.* S151C purified in DDM detergent, the absorbance maximum at 544 nm is indicative of the ground state. Both samples exhibit fluorescence (Fig 3.7) which can only be achieved through full photocycle completion.

Circular Dichroism analysis[86] showed that the OG sample was properly folded, with an estimated secondary structure content of 74% α-helical, 6% β-sheet and 20% random coil (Fig 3.6). These values are similar to those of the other archaerhodopsin proteins. AR1 is composed of 65% α-helix and 4% β-sheet (PDB ID: 1UAZ) and AR2 has an estimated 68% α-helical and 3% β-sheet structure (PDB ID: 1VGO).

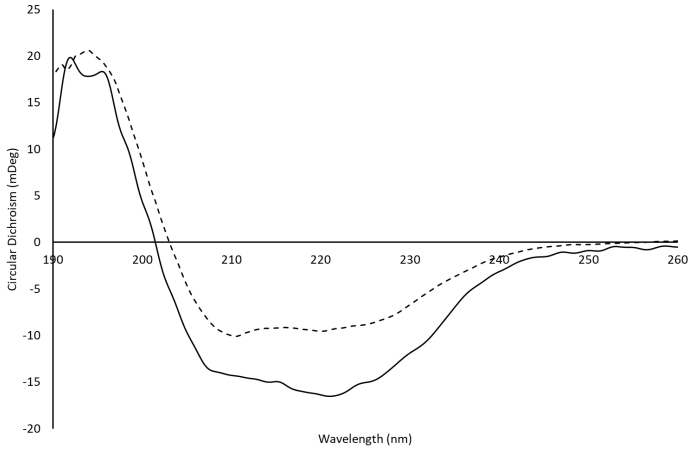


Figure 3.6: CD spectra of wtAR3 and S151C in OG. The mutant sample (*solid line*) differs slightly from the *wt*AR3 (*dashed line*)

When purified using OG and Triton^{TM} X-100, S151C was not stable in the ground state. The most prevalent photocycle intermediate appeared to be the M state due to the characteristic yellow colour and absorbance maximum at approximately 400 nm. This indicates that the M state, normally a short-lived photocycle intermediate, became rate-limiting due to the mutation. To test whether the protein was still functional and able to progress through the photocycle, fluorescence studies were carried out on the S151C mutant in OG (Fig 3.7). The fluorescence emission spectrum shows that the protein is still able to advance through the photocycle, producing a fluorescence emission due to the generation of the Q state (Fig 1.9), with a Stokes shift of approximately 140 nm.

Successful purification of S151C was only achieved when using either DDM or CHAPS detergent, as well as a detergent-free approach using SMA polymer to

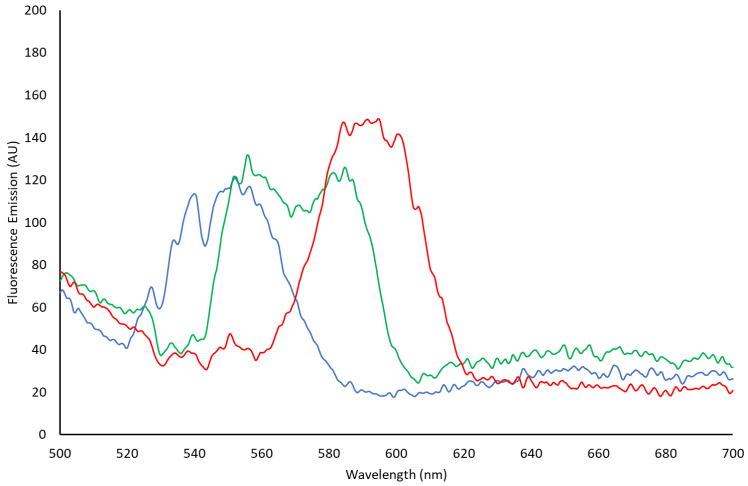


Figure 3.7: S151C fluorescence emission spectra in OG. S151C excited at 400 (*blue*), 420 (*green*) and 440 nm (*red*) and their fluorescence emission recorded. For fluorescence to be possible, S151C must be able to progress through the photocycle, thus the protein is not completely locked in the 'M' state in OG micelles.

purify AR3 into Lipodisqs^{TM}. The peak caused by the ground state S151C absorbs maximally close to 540 nm (Fig 3.8). Intriguingly, the shifts between solubilisation and reconstitution procedures displayed here show the effect of the protein-lipid environment on protein structure. Comparing the relative absorbances at 409 nm and 544 nm, which correspond to the peaks from M and ground state protein, respectively, it can be estimated that in OG, 82% of the total protein content is in the M state, with only 17% in the ground state. In DDM, the majority of the protein content (70%) is in the ground state, with approximately 30% in the M state intermediate. These differing absorbance spectra are caused by the differing lateral pressure exerted by the lipid on the whole protein which, in turn, effects the precise interactions between the retinal and the opsin protein. These small changes in protein structure shift the spectra comparably with the mutations themselves, proving that the interactions between protein and lipid are not as idle as previously believed and deserve investigation[9, 90].

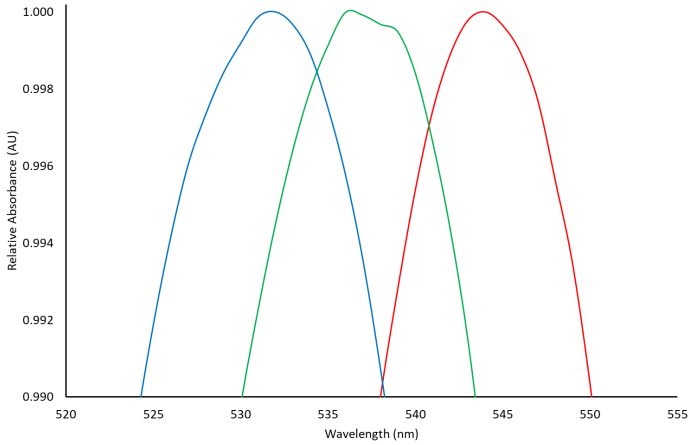


Figure 3.8: Absorbance maxima of S151C successfully purified utilising different detergents and SMA polymer. DDM (*red*) and CHAPS (*green*) detergent reconstitutions produced ground state protein that peaked absorbance at 544 and 536 nm respectively. Detergent-free purification using SMA polymer (*blue*) shifted the ground state absorbance maximum to 532 nm.

3.4 DDM Purification

After failing to purify the whole mutant library using OG, DDM was chosen as it has been demonstrated to be a suitable detergent for membrane proteins[41] and a more appropriate detergent for solubilisation of AR3, particularly for the spectral mutants that appear to affect the photocycle kinetics. The samples were resolubilised into 1% DDM and then diluted to 0.2% for purification on the IMAC column, the trace of which can be seen in Fig 3.9. The UV trace increases sharply as the protein is loaded, and then plateaus throughout loading. The wash step removes non-specifically bound contaminants, seen as peaks in the spectra. Imidazole elutes the AR3 protein due to the His_6 tag being outcompeted by the increased concentration of imidazole in the eultion buffer (Fig 3.9B).

Purification was carried out on all the previous single mutants, including an additional mutation, D222E, which has been previously shown to produce a red shift[89]. As can be seen in Fig 3.10, the proteins are sucesfully purified from the *E. coli* model organism. Using the HisTrap as the only chromatographic purification

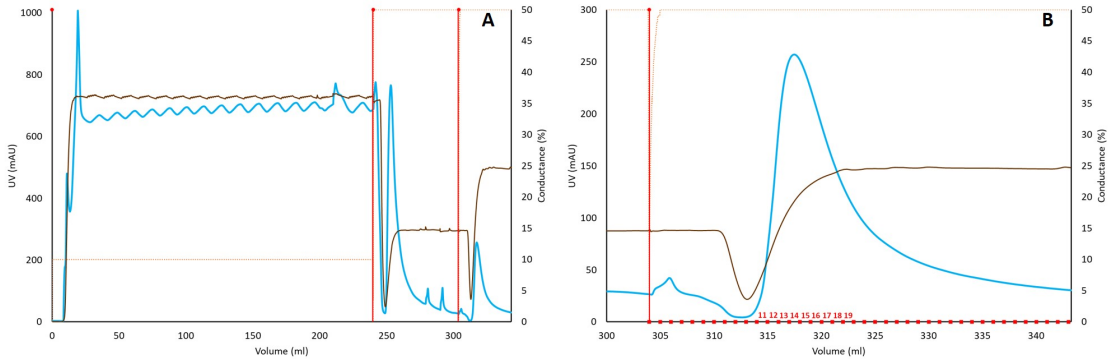


Figure 3.9: A trace of the HisTrap when purifying an AR3 sample. UV trace (*blue*) and conductance (*brown*) are used to track the protein loading, wash and elution. **A.** A flow rate (*orange dotted line*) of 1 ml/min is used for sample loading, 5 ml/min used for washing and eluting the sample. Washing commences at 240 ml and Elution at 303 ml (*red lines*). **B.** A zoomed in view of the elution step. Fractions (*red squares*) are collected and pooled for UV-Vis and further experiments.

step results in some contamination in the final sample, although not at a level that prohibits further experimentation in this report. This could be remedied using *Halobacterum salinarum* as the model organism, as it can be engineered to express only AR3 in its membrane.

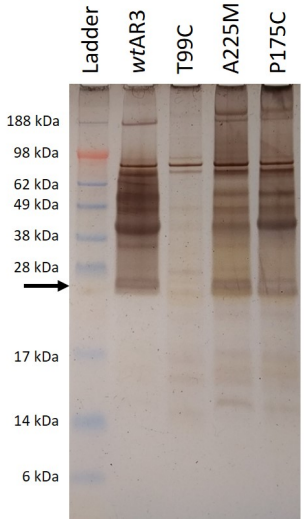


Figure 3.10: A silver stain protein gel of *wt*AR3 and three mutants, when resolubilised in DDM. Monomeric AR3 can be seen in the band highlighted by the *arrow* at approximatley 27 kDa.

*wt*AR3 absorbs maximally at 558 nm (Table 3.4), shifted by 4 nm compared to the OG sample. This 4 nm shift from OG to DDM is also seen with T99C, R173C, P175C and to a lesser extent A225M. Overall, all six single mutants, one double mutant, two triple mutants and a quadruple mutant were purified successfully, and their peaks are shown in Fig 3.11 and Table 3.4. With each of these mutants, the ground state protein peak round 550 nm can be discerned from the full spectra, with all single mutants producing a pure ground state protein. The more complex combined mutants do produce a considerable peak at 400 nm, corresponding to the expected M state region, although the ground state can still be identified. This indicates the mutations are not solely responsible for changing the photocycle kinetics but decrease the balance of the photocycle along with the detergent pressure on the opsin protein. The S151C mutant caused a blue shift of 14 nm compared to *wt*AR3, consistent with the literature[60]. T99C also caused a 14 nm hypsochromic shift. D222E resulted in a 7 nm red shift in the absorbance maximum (565 nm). The A225M 2 nm red shift seen in OG was not observed in DDM. In DDM the mutation shifts the absorbance maximum of the protein by a 1 nm blue shift, peaking at 557 nm. The cysteine mutants intended for gold-thiol bond immobilisation, R173C and P175C, produced 1 nm hypsochromic and bathochromic shifts respectively. This is in accordance with the data collected in OG (Fig 3.4).

The absorption maximum of the double mutant T99C-A225M (TA) purified in DDM was measured at 548 nm. This decreased the effect of T99C and does not directly coincide with the effects of the two single mutants combined.

The triple mutants T99C-S151C-P175C (TSP) and T99C-A225M-P175C (TAP) absorb maximally at 538 and 546 nm, respectively, corresponding to a 20 nm and 12 nm blue shift, respectively. As seen previously, this does not directly correspond to the addition of each single mutant equally.

A quadruple mutant containing all three original retinal pocket mutations (T99C, S151C and A225M) as well as an available cysteine for immobilisation (P175C) was purified and gave the highest hypsochromic tuning of the AR3 absorbance maximum, a shift of 25 nm relative to wtAR3 producing a peak at 533 nm. However,

there is a considerable peak at around 400 nm seen from this mutant, although the ground state protein absorbing at 533 nm can still be observed.

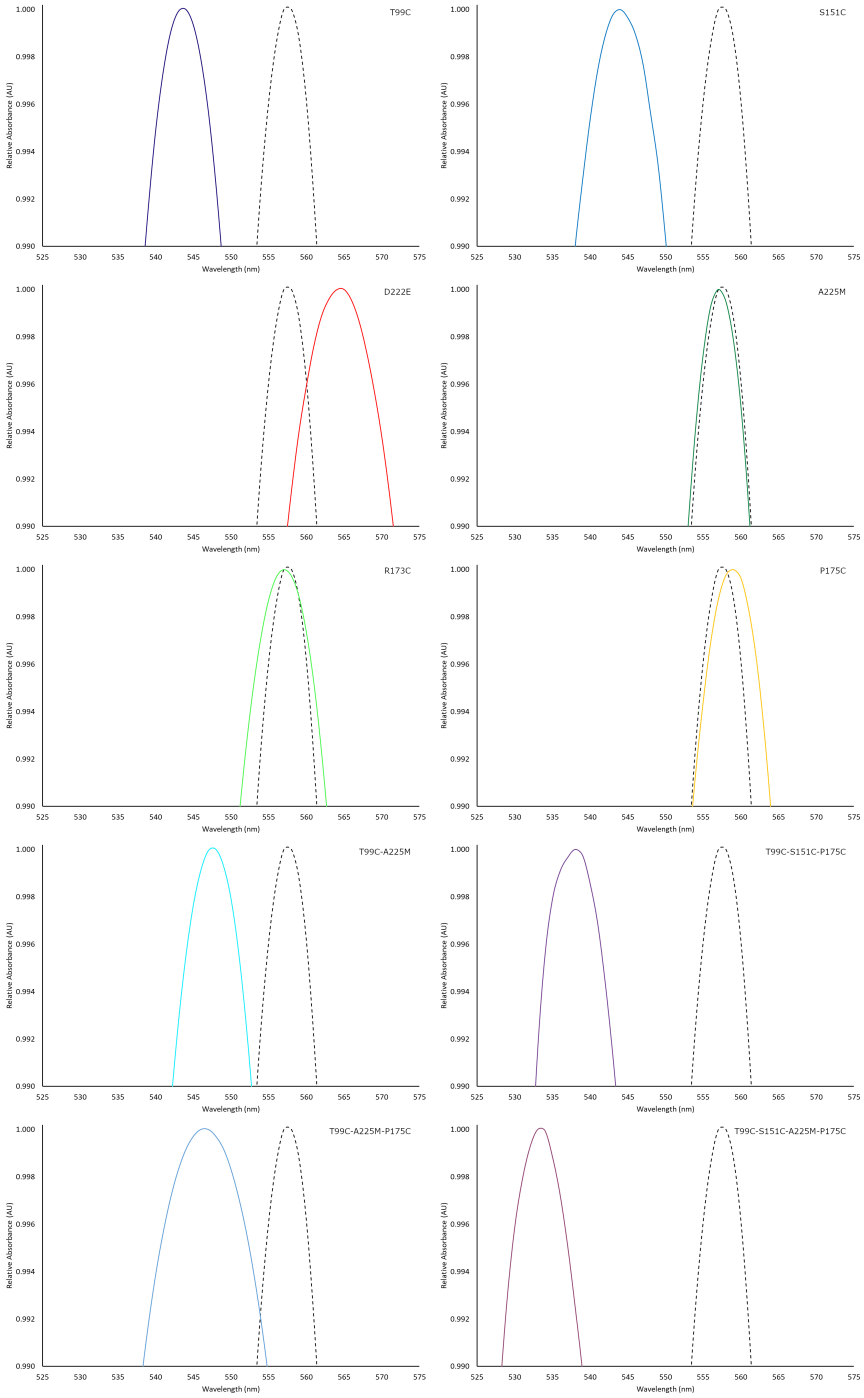


Figure 3.11: Absoption spectra of AR3 mutants resolubilised from *E. coli* in DDM detergent. Each individual mutant peak is shown compared to the *wt*AR3 spectrum (*dashed line*). For clarity, each peak is cut off at 0.990 AU. Maximum absorbances and shifts for each mutant are shown in Table 3.4.

Table 3.4: All Mutants successfully purified into DDM micelles. Peak absorbances and shifts, relative to *wt*AR3 are shown.

Mutant	**Peak Absorbance (nm)**	**Shift (nm)**
*wt*AR3	558	-
T99C	544	-14
S151C	544	-14
D222E	565	+7
A225M	557	-1
R173C	557	-1
P175C	559	+1
T99C-A225M	548	-10
T99C-S151C-P175C	538	-20
T99C-A225M-P175C	546	-12
T99C-S151C-A225M-P175C	533	-25

4

Results - AR3 Reconstitution for Biotechnological Applications

4.1 Visible CD

To assess the effect of AR3 solubilisation in detergent micelles on protein oligomerisation, circular dichroism in the visible spectrum (400-700 nm) was used. This technique utilises the retinal molecule, which is optically active in the presence of circularly polarised light. The environment of the retinal dictates the spectrum, with multiple chromophores in a relatively close proximity, as is seen in trimeric bR, producing a distinct spectra. The CD spectra of monomeric and trimeric bR are distinct in this region[91–94]. Archaerhodopsin 3 is thought to be arranged as a trimer, with bacterioruberin bound at the interface between subunits, forming hydrophobic interactions with both the helices and the omega loop of each protein, as in AR2 (Fig 1.8). In the *E. coli* system used here, there is no bacterioruberin present, and as such it is expected that AR3 is found in its monomeric state. As seen with many other membrane proteins, and specifically with rhodopsin homologues of AR3, reconstitution of the protein into a detergent micelle abolishes the interactions between protein subunits necessary for multi-subunit complexes to be formed. In the case of the archaerhodopsins, there are no known covalent interactions between the subunits that would otherwise resist the extraction from the native membrane

to detergent micelles.

Fig 4.1 shows the visible CD and UV-CD spectra for a number of the AR3 mutants, compared to a sample of *wt*bR extracted directly from the purple membrane of *Halobacterium salinarum* and a *wt*AR3 sample from the archaeal membrane of *Halorubrum sodomense*. All of the samples are properly folded, as shown by the UV-CD. As can be seen from the *wt*bR sample which has been shown to be in a trimeric formation by AFM[95], there is a distinct biphasic curve, peaking at approximately 530 nm before dropping to a negative minimum point at 600 nm. The *wt*AR3 sample extracted from the archaeal membrane of *H. sodomense* does not show a biphasic curve, instead three distinct peaks can be seen, likely a result of the bacterioruberin pigment. Compared to these spectra, the AR3 samples from *E. coli* do not show a distinct biphasic curve, thus suggesting that AR3 is monomeric in detergent. The *wt*AR3 samples in both OG and DDM have similar curves, with a single broad peak over the range of 450-550 nm. This is similar to the existing spectra from monomeric bR, in which a single broad hump is seen over the same wavelength range[94]. The spectrum is also clearly matched by the R173C sample when incorporated into DDM micelles, suggesting that the addition of an accessible cysteine is unlikely to result in disulphide bond formation between AR3 monomers. The S151C sample purified with SMA produces a folded sample, as shown by the UV-CD, but does not give an intense enough response in the visible CD spectrum to be clearly evaluated against the noise of the system. In the published article that include Lipodisq purified bR samples directly extracted from the purple membrane, there is a distinct sinlge peak in the spectrum that has a higher intensity than the detergent purified sample, along with a sightly shifted peak[94]. This is not seen with the AR3 sample extracted from *E. coli*, potentially due to the lower concentration of the sample or due to interactions between the AR3 and other contaminant proteins.

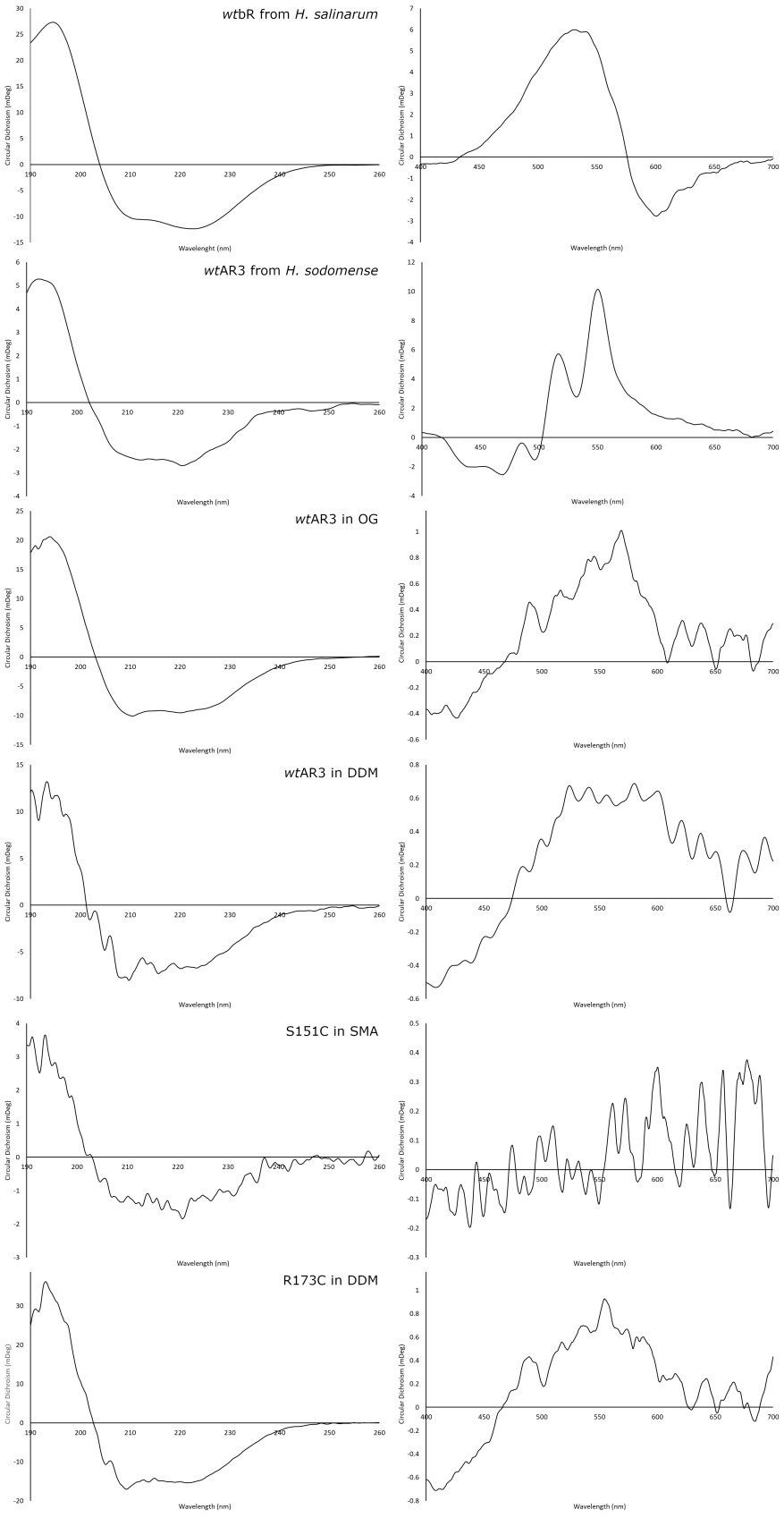


Figure 4.1: Circular Dichroism of bR and AR3 in the wavelength range of 190-260 nm and 400-700 nm. *Left.* Far-UV CD spectrum of each sample, this shows that each is folded correctly irrespective of purification protocol. *Right.* Visible CD can uncover the oligomeric state of the protein in various environments. The *wt*bR and *wt*AR3 samples prepared from archaeal cells were purified in collaboration with J. Viñals Camallonga.

4.2 Native PAGE

To further ascertain if AR3 can be resolubilised into detergent systems as a multimeric complex from *E. coli*, native PAGE was performed on each S151C sample generated from the different purification approaches. The samples were run against a positive control of bR extracted directly from the membrane of *H. salinarum* in which bR is found in the purple membrane patches as a trimer, as well as the same sample solubilised in OG, in which bR is monomeric. To help identify the contamination present from other proteins expressed in the *E. coli* membrane, a sample of purified *E. coli* BL21 without the AR3 plasmid insertion was also used. This allowed for the identification of the bands from *E. coli* and those from the AR3. From the gel image (Fig 4.2) it can be seen that monomeric AR3 is present in all of the *E. coli* samples as shown by the stained band at the very bottom of the gel.

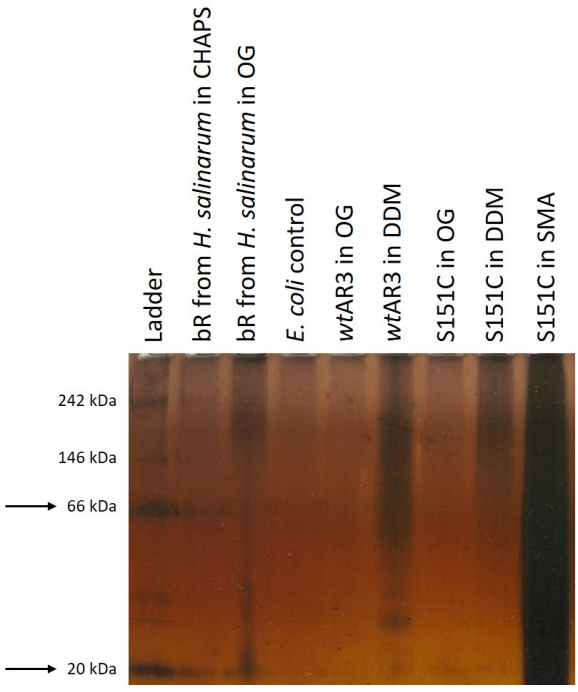


Figure 4.2: Silver stained native PAGE gel of various bR and AR3 samples produced by different preperations. bR from *H. salinarum* was prepared and delipidated using CHAPS to conserve the trimer, shown in by the *top arrow* at 66 kDa. The samples in OG produce bands at 20 kDa. The apparent higher molecular weight of the DDM sample may result from the larger micelle size.

The *wt*bR sample shows a band for trimeric bR at around 66 kDa, this is not seen in the AR3 samples, suggesting that AR3 is only solubilised as a monomer. The higher indicated molecular weight of the *wt*AR3 band in DDM is likely a result of the larger size of a DDM micelle compared to OG. The appearance of a smeared lane for the SMA sample may be a result of the nature of the polymer that does not allow it to run properly under native PAGE conditons. The contamination from the *E. coli* membrane control shows that the only remaining bands in the AR3 sample is the monomeric protein.

4.3 Solid Phase UV-Vis Spectroscopy

To confirm the adsorption of AR3 onto a gold surface, solid phase UV-Vis spectroscopy can be used as a first control. This technique requires the incubation of a gold on glass chip with a cysteine mutant of AR3 before having the absorption spectrum of the sample recorded. This has been achieved with a cysteine mutant of bR[75, 96], in which a hypsochromic shift in the absorbance occurs due to the plasmon effect of the gold surface[75]. The two single cysteine mutants tested here, R173C and P175C were both purified in OG micelles and produced an absorbance spectrum (Fig 4.3) absorbing maximally as 608 and 574 nm, respectively. The 55 and 19 nm shifts in R173C and P175C absorbance between the solution and solid phase UV-Vis readings confirm immobilisation. The appearance of a bathochromic shift instead of the hypsochromic shifts reported before with bR is likely due to differing gold nanochips. In the bR experiment previously carried out in the Watts lab, different gold on glass substrates were used which had a different fundamental absorption spectrum. The anchoring of the protein to the metal surface also restricts its conformation, therefore reducing the ability of the structure to adjust throughout the photocycle. This may cause a change in the retinal environment, leading to the absorbance shift once immobilised.

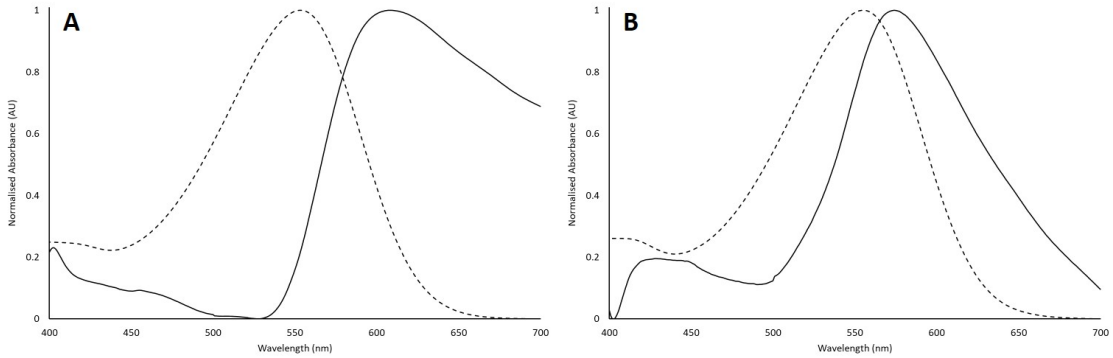


Figure 4.3: Solid phase UV-Vis spectroscopy of AR3 cysteine mutants. Solid phase UV-Vis samples (*Solid line*) plotted against the solution UV-Vis spectrum (*Dotted line*) for **A.** R173C and **B.** P175C. The bathochromic shift after immobilisation occurs due to the interacting metal ions. Immobilisation and spectroscopy carried out in collaboration with J. Viñals Camallonga.

This UV-Vis spectroscopy acts as an initial test to confirm cysteine mutation of AR3 and its ability to be adsorbed onto a gold surface. However, further testing with techniques such as C-AFM would provide more details as to the kinetics of the bound protein.

4.4 AFM

Atomic force microscopy (AFM) is one of the most valuable methods used in the research and development of bio-processors and sensors. For use in this field, the protein of choice must have the ability to be adsorbed onto a conductive surface, commonly gold, for the propagation of an electrical current. The AR3 mutants presented in this report, which contain an anchor point in the form of a cysteine mutant (R173C and P175C), are designed for this purpose, to allow the AR3 protein to be bound to the conductive surface and form a metal-molecule-metal (mMm) junction. AFM was performed on the single mutant P175C as well as with the *wt*AR3 as a preliminary investigation into the binding capacity of AR3 onto a mica surface. From the AFM images (Fig 4.4), it can be seen that both proteins adsorb onto the mica surface. The micelles can be seen as mountains in the image, with an approximate height of 5 nm, this corresponds closely to the height of AR3,

which is around 55 Å according to the crystal structure (PDB ID: 6GUX). As can be seen in the cysteine mutant image, there are large high peaks exceeding 15 nm, which may be be a result of aggregation of the sample, or possibly due to the accessible cysteine allowing the protein to form stacked dimers, although higher order oligomers are unlikely as shown by visible CD (Fig 4.1). There is a potential for other contaminant proteins to bind to the cystiene, resulting in *E. coli* proteins attaching to AR3 and causing the larger peaks in the image. To image individual protein oligomers, a higher resolution image would be required, using a much more precisely manufactured and smaller AFM tip.

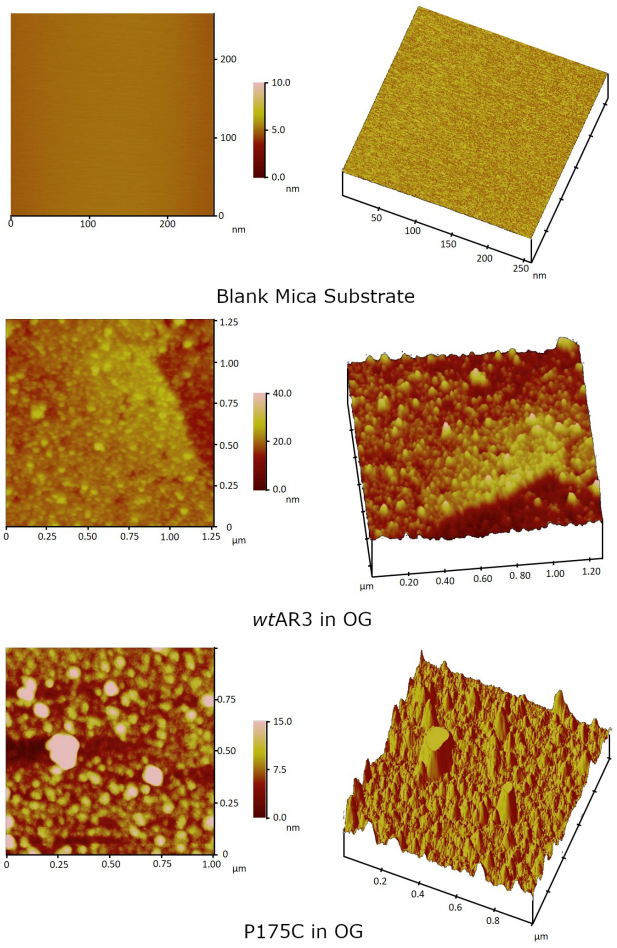


Figure 4.4: AFM images of AR3 deposited on mica. Analysis of the wild type sample showed peaks of 5 nm attributable to an OG micelle with adsorbed AR3 (*Round spots*). The cysteine mutant peaked at above 15 nm, potentially due to aggregation of the sample or contamination. Sample preperation and microscopy achieved in collaboration with J. Viñals Camallonga.

5 Discussion

5.1 Discussion

5.1.1 Spectral Tuning of AR3

The mutations selected in this thesis for AR3 have been chosen from previously published work investigating the effect of amino acid substitutions around the retinal chromophore in other rhodopsin proteins. The effects of a number of these mutations is well characterised, and several perturb the protein in more ways than simply tuning thier absorbance maximum. S141C has been shown to reduce the proton-pumping ability of bacteriorhodopsin to 20% compared to *wt*bR[60]. This may explain the lack of success purifying the AR3 S151C mutant in OG or Triton^{TM} X-100 detergent systems, seemingly as a result of a prolonged M state photo-intermediate. This mutation may perturb the photocycle kinetics of the protein, extending it, causing a delayed return to the ground state. Seemingly, any mutation at this conserved position in bR[60] or GR[57] results in a reduction of proton pumping when measured by light-dependent pH changes. S141A in bR caused a reduction to 80% of wild type proton pumping capacity[60], whilst the corresponding residue in *Gloeobacter violaceus* rhodopsin (GR), S181, caused a reduction to 50% of the wild type protein. In GR, mutations of serine to alanine, glycine or proline give a blue shift of 7, 12 and 18 nm, respectively[57].

One of the possible reasons for failing to generate a mutant with an absorption maximum shift greater than 25 nm for AR3, may be the hydrophobicity of the retinal binding pocket of the protein compared to bR and other rhodopsins. A measure of the hydrophobicity of a series of amino acids, The grand average of hydropathicity (GRAVY)[97, 98] of the AR3 retinal pocket is -0.057, defined by the residues within 5 Å of the retinal chromophore, as observed in the crystal structure. In comparison, bR and GR have GRAVY values of -0.405 and -0.271, respectively. The residues involved in the retinal pocket are shown in Table 5.1.

Table 5.1: Amino acids within 5 Å of the retinal chromophore in AR3, bR and GR. Residues in *bold* show additonal residues in the AR3 pocket. Grand average of hydropathicity (GRAVY) calculated from[97, 98]

Protein	Retinal Pocket Residues	GRAVY
AR3	**VR**YDWTTLMI**T**G**LW**WST**C**MW**T**YPW**F**DAK	-0.057
bR	YDWTTLMIGWSTMWYPWDAK	-0.405
GR	DWTVLMIGEGSTFWYYPYDAK	-0.271

The red-shifting mutations D222E and A225M both increase the hydrophobicity of the retinal pocket (Fig 5.1). A 7 nm shift arises as a result of the increased hydrophobicity of aspartic acid compared to glutamic acid. The increased sidechain length of glutamate results in stronger hydrogen bonding between the negatively charged oxygen and the end of the retinal chain, neighbouring the Schiff Base. A225M produces a small shift due to the slight increase in hydrophobicity of methionine compared to the native alanine, as well as the introduction of a large sulphur-containing sidechain that induces steric hindrance in the pocket. This is consistent with the small shift seen in other published studies with AR3[56].

The blue-shifts of T99C and S151C mutants are mainly caused by the steric hindrance of the introduction of a large sulphur-containing group into the retinal pocket (Fig 5.1). The introduction of this relatively bulky group restricts the

rotational freedom and movement of retinal within the pocket, therefore altering its environment and giving rise to the shifted absorbance maximum. The loss of -OH groups from both threonine and serine for -SH groups in the mutants may also perturb the hydropgen bonding between protein and chromophore. As the introduced thiol group is less polar than original hydroxide groups, the hydrogen bonding network will be unsettled. Mutating a cysteine at position 151 may lock the retinal in place in the pocket of AR3, rendering it unable to progress through the photocycle, specifically slowing the decay of the M state by constricting the ability of the retinal to isomerise back to the all-*trans* isomer from 13-*cis*[99].

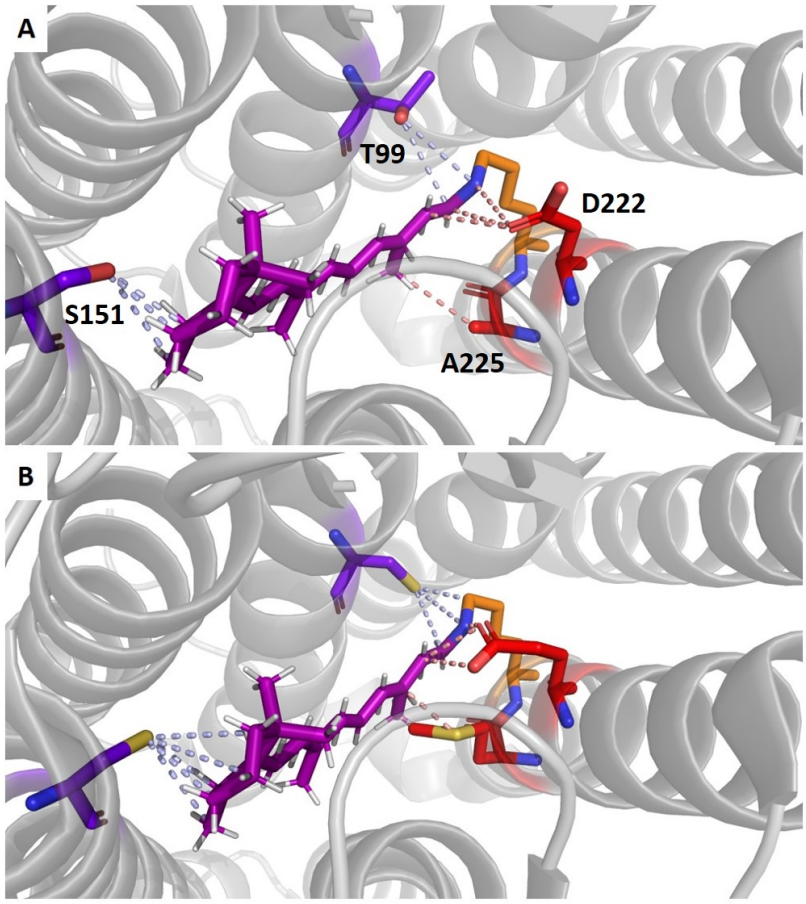


Figure 5.1: The retinal pocket of AR3. Retinal shown in *purple*, Schiff Base shown in *orange*. Mutated residues shown in *blue* and *red*. **A.** The native pocket with interactions between mutant residues shown as *dashed lines*. **B.** Schematic representation of the mutated pocket of AR3. Estimated interations between sidechains and retinal shown as *dashed lines*. Generated from PDB ID: 6GUX.

For effective use as either an optogenetic or molecular bio-sensor tool, AR3 mutants must have distinct absorption spectra that can be activated by specific wavelengths of light. Their absorbance maxima must be far enough apart, with minimal overlap, that activation at one mutant's maximal absorbance does not result in a large activation of another mutant also present. Wavelength-selectivity is a particularly desirable property for optogenetics, where multiple proteins are incorporated into a single cell, or within close proximity. The two most distinct mutants generated in this report, TSAP (λ_{max}= 533 nm) and D222E (λ_{max}= 565 nm) have a difference in maximal absorbance of 32 nm. The full spectra of both of these mutants show that there is some overlap, meaning the possibility for individual activation of the AR3 mutants produced in this thesis may be limited. By integrating the two curves to calculate the amount of overlap, it is found that 62% of the total area is overlapped between the two data sets. However, all of the mutants described previously have the potential to be successfully incorporated into a cell for neural imaging and optogenetics.

5.1.2 Membrane Protein Purification

One of the main challenges when purifying a membrane protein from a host organism is the difficulty in achieving the successful solubilisation of the protein in detergent to allow for further biochemical and biophysical characterisation. This is particularly apparent when working with multi-pass transmembrane proteins which often contain an intricate architecture within their transmembrane helices that relies on lipid-protein interactions in their native environment. In the case of retinal-containing proteins, the environment of the chromophore is delicately controlled by the surrounding amino acid sidechains, the precise structure of which is sensitive to the lipid content of the membrane environment.

The lateral pressure exerted on the protein in a natural bilayer cannot be accurately mimicked by a detergent micelle. Due to the single composition of the detergent micelle, there is currently no way to accurately mimic the precise differences in lipids

and their interactions that are normally found in the bilayer. Biological membranes contain different proportions of lipid classes, supplemented by other non-polar compounds. As these are not present in the simple single detergent micelles, there is no way for the protein to fully satisfy all of the lipid-protein interactions in the native environment. The curvature of a detergent micelle often differs greatly from that of a biological bilayer. In a spherical micelle, the embedded protein experiences a very different environment to a native membrane, due to the angle of the detergent molecules interacting with its hydrophobic regions. Commonly, hydrophilic amino acids allow for electrostatic and van der waals interactions with the charged lipid head groups. In a detergent environment, this interaction is unlikely to be fully satisfied as a result of the different orientations of the detergent molecule. DDM however, and to some extent OG, uses its unique sugar head group to create an aqueous-like microenvironment at the micelle interface, this mimics the conditions of the lipid bilayer more successfully than other detergents[39].

The complex lipid species found specifically in archaea, those bound by ether linkage, are not mimicked by either the *E. coli* expression system or the detergents used for purification. The complex archaeol (Fig 1.4), dimeric and ether-linked lipids needed for the archaea to survive in harsh environments are not present, these may have potential interactions with the membrane proteins and provide some stability in the bilayer.

5.1.3 Detergent Suitability

The aggregate number of each detergent (Table 5.2), the number of individual molecules required to create a micelle (at concentrations higher than the CMC) gives an indication as to the orientation of the lipid and its interaction with the protein. There are only 10 CHAPS molecules per micelle compared to 98 molecules for DDM. The smaller micelle and higher CMC of OG compared to DDM may explain the instability of the protein in this environment and is the reason for the migration differences seen in the native PAGE (Fig 4.2). The higher concentartion

needed to spontaneously form micelles, and the fewer monomers per micelle suggests that OG has difficulty producing stable micelles in solution, a factor that may affect its ability to solubilise proteins. OG monomers have also been suggested to be too small for efficient membrane solubilisation. It is thought that the small head group and chain can not just interact at the hydrophobic region. For example OG has been shown to bind to and occulde the proton channel of bR, deactivating the protein[100]. In an OG environment, bR undergoes irreversible light-induced denaturisation and is in general less stable[101]. For GPCRs, OG does not mimic the hydrophobic band around the protein as well as DDM, and therefore preserves fewer native interhelical contacts[102]. To successfully stabilise some complex membrane proteins, the addition of ligands[103, 104] or lipids[105] is often required as they help to lock the protein structure or induce some pressure.

Table 5.2: The aggregation number of commonly used detergents, calculated from respective suppliers listed in the materials and methods. Average protein-free micelle size reported in[41]

Detergent	**Aggregation Number**	**Average Micelle Size (kDa)**
Sodium dodecyl sulfate (SDS)	62	18
Triton^{TM} X-100	100-155	48-106
n-Octyl-β-D-glucopyranoside (OG)	27	25
n-Dodecyl-β-D-matloside (DDM)	98	98
3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate (CHAPS)	10	6

OG is widely used for crystallisation of membrane proteins[106], due to its propensity to induce rigidity in proteins that would otherwise be too flexible to form ordered crystals and its high CMC resulting in its easy dissociation from the protein. However, one of the inherent problems when producing ordered arrangements of membrane proteins is that this restricted structure is not representative of

protein behaviour in native environments. Membrane proteins undergo various structural changes throughout their functional roles in the cell, many of these cannot be accurately captured in a single conformation. The variable loops found in the protein structure are often crucial for identifying functional sites in proteins. However, these flexible regions can often be poorly resolved by X-ray crystallographic methods or may be locked in a non-native conformation by crystal contacts. The shorter chain of OG compared to other milder detergents of the same class, such as DDM, has been proposed to allow the protein to make crystal contacts with other asymmetric units. This is achieved as the detergent has fewer monomers per micelle and the hydrophobic belt of detergent does not extend to cover any hydrophilic sidechains[107]. Detergent purity also greatly affects the likelihood of success in crystallisation. Within the highly ordered crystals, even a small amount of a detergent enantiomer or contaminants can disrupt the crystal packing and thus reduce the resolution of the resulting diffraction pattern. This, alongside the variable chain length, is the main reason why Triton^{TM} X-100 is not suitable for crystallographic applications[108].

The most appropriate approach for biophysical and biochemical analysis of membrane proteins is to keep the protein its native state. Lipodisq^{TM} nanoparticles provide a novel way of generating a native-like discoidal membrane. Using an SMA polymer, a protein can be extracted directly from the host membrane, without the need for detergent solubilisation. This detergent free approach allows for a small amount of native lipid to be co-purified with the protein, allowing for lipid-protein interactions, that are frequently disrupted when transferring to detergent micelles, to be preserved. Excess lipids can then be removed, leaving only the native lipids directly interacting with the protein structure. One benefit of this technique compared to the use of membrane scaffold proteins (MSPs) to generate protein-lipid discs is that the MSP can interfere with downstream assays. The size of the SMA disc can also be controlled depending on the ratio of styrene:maleic acid. Commonly 3:1, used in this thesis, and 2:1 polymers are adopted as they entrap

enough lipid with the protein to create a stable disc which is stable at physiological pH[109]. With this approach, both hydrophilic faces of the protein are accessible, allowing characterisation of potential binding sites and functional regions of both the intracellular and extracellular portions of the protein[94].

5.1.4 AR3 Oligomerisation

The monomeric reconstitution of AR3 is somewhat expected when solubilising the protein from an *E. coli* model membrane into detergent micelles. If AR3 is natively in a trimeric arrangement in the membrane of *Halorubrum sodomense*, as has been proven with AR2 in its native archaeal membrane[53], it would likely require the presence of the pigment bacterioruberin to bind between subunits and facilitate oligomerisation. This chromophore is not present in the *E. coli* in which the studies in this thesis were performed. The appearance of an omega loop in the recently solved crystal structure of AR3 (PDB ID: 6GUX) may act as one possible binding site for a bacterioruberin molecule alongside the cleft between subunit helices. It may be possible that trimeric arrangements can form in a lattice arrangement without this pigment, as seen with bR. In the purple membrane of *H. salinarum*, bR can self-assemble into trimers without the intervention of any other molecules apart from the membrane lipids themselves[110, 111]. In AFM analysis the trimers were retained after protein extraction from purple membrane patches that were simply extracted directly from *H. salinarum* without detergent solubilisation. The distinct visible CD spectra (Fig 4.1) seen from the AR3 *E. coli* samples suggests that they are monomeric, a hypothesis that is reaffirmed with reference to the native PAGE (Fig 4.2). A monomeric reconstitution of rhodopsin proteins is favourable for their use in biotechnology, especially as a bio-sensor. Monomeric AR3 is the smallest possible functional unit of the protein and has been shown to still act at 100% functionality[112]. Therefore, a true single molecule processor can be developed. AFM could also confirm the arrangement of AR3 in the micelle, however the resolution of the image would need to be much higher than has been captured here.

For previous studies that have utilised AFM to distinguish monomeric and trimeric bR, high-speed AFM was employed that was able to clearly show the individual trimers which only span approximately 5 nm[95].

6 Conclusions and Future Work

6.1 Concluding remarks

Recent advances in the purification of membrane proteins has allowed for their structures to be solved in high-resolution. Many microbial rhodopsins have been resolved structurally through X-ray crystallography, including the protein of interest in this thesis, AR3. However, due to the nature of these proteins, there are still some hurdles to overcome to generate wild type structures that include fully native conformations, relevant for biotechnological and pharmaceutical utilisation. Chiefly, the lipid environment of the protein must be designed to mimic that of its native membrane as closely as possible to fully satisfy the lipid-protein interactions that contribute to the final protein conformation. The study of membrane proteins in model systems cannot fully replicate the intricacies of the native lipid membrane from which the protein is found naturally. Detergents have historically been used for protein solubilisation allowing downstream structural and functional analysis of the protein. However, it has become apparent that these systems may not be fully suitable for the analysis of unstable proteins that rely heavily on specific lipid-protein interactions, or with mutagenesis studies that put a strain on the protein folding or produce thermostabilised mutants[113] compared to the wild type. One potential remedy for this is the use of polymer discs which allow native lipids

to be purified with the protein, maintaining the important lipid-protein interactions that maintain the membrane protein in its native conformation.

Here, the initial aims of this thesis have been accomplished. A total of fifteen mutants were successfully produced, with 12 of these being successfully purified from the *E. coli* model organism into a detergent (either OG or DDM). The tuning of AR3 maximal absorbance was achieved, with the production of a 25 nm hypsochromic shifting mutant (TSAP) and an 8 nm bathochromic shifting mutant (D222E). The two cysteine locations chosen for immobilisation onto a conductive surface (R173C and P175C) allowed for solid phase UV-Vis spectroscopy to be carried out for the first time on AR3, generating a hypsochromic shift when immobilised. The suitability of AR3 as a biotechnological tool has been investigated in this thesis, although further experimentation could be carried out, especially in the assaying of proton pumping of the AR3 mutants.
The AR3 mutants produced here, with more development, have the potential to be applied to both optogenetic and bio-electronic applications. With the incorporation of a cysteine on the intracellular loop of the protein, AR3 can be adsorbed onto a conductive surface to function as a wavelength-specific single molecule processor, producing a proton current when activated by light. The red-shifted mutant D222E has interesting potential to act as a neural silencer in optogenetic studies. The tuned absorbance makes the protein more suited to these functions as it allows for distinct activation compared to the other mutants listed in this work.

6.2 Future Work

There are still some objectives to be completed in the development of AR3 as a successful molecular bio-sensor and optogenetic actuator.

1. More rationally designed mutants could be produced targeting the other amino acids in the retinal pocket, with the aim to shift the absorbance maximum further than the 25 and 8 nm hypsochromic and bathochromic shifts achieved

here. Producing a bigger wavelength difference between blue- and red-shifting mutants will allow for specific activation of one protein over the other, which can be used for the spatial analysis of neuronal networks.

2. The proton pumping ability of the mutants described in this report may need to be tested to assess their suitability as bio-sensors. There are a number of published protocols that assay the proton pumping of rhodopsins[55, 114], which may be of particular importance for S151C, which in bR has been reported to slow the proton pumping to 20% of the wild type protein[57].

3. The next steps for the development of AR3 as a potential bio-sensor is the assessment of proton conductance by C-AFM. The introduced cysteine on the EF loop of the protein allows covalent bonding to a gold surface. The immobilised protein can then be activated by a specific wavelength of light and the resulting proton current can be measured[71–73]. High-speed AFM with a more precise tip can give a resolution high enough (around 5 nm) to evaluate the oligomeric state of AR3 in the detergent environments compared to in *E. coli* and its native *H. sodomense* membranes[95, 115].

References

[1] Virchow Rudolf. “Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre”. In: August Hirschwald, Berlin, 1858.

[2] Jonathan Lombard, Purificación López-García, and David Moreira. “The early evolution of lipid membranes and the three domains of life”. In: *Nature Reviews Microbiology* 10.7 (2012), pp. 507–515.

[3] Bruce Alberts et al. “Molecular Biology of the Cell”. In: *Molecular Biology of the Cell* (2015).

[4] Suzanne Jackowski. “Cell Cycle Regulation of Membrane Phospholipid Metabolism”. In: *The Journal of Biological chemistry* 271.36 (1996), pp. 20219–20222.

[5] Jain M. *Introduction to biological membranes*. 2nd. Wiley, 188.

[6] S. J. Singer and Garth L. Nicolson. “The Fluid Mosaic Model of the Structure of Cel Membranes”. In: *Science. AAAS* 175.4023 (1972), pp. 720–731.

[7] Ken Jacobson, Erin D Sheets, and Rudolf Simson. “Revisiting the Fluid Mosaic Model of Membranes”. In: *Science* 268.5216 (1995), pp. 1441–1442.

[8] Erdinc Sezgin et al. “The mystery of membrane organization: Composition, regulation and roles of lipid rafts”. In: *Nature Reviews Molecular Cell Biology* 18.6 (2017), pp. 361–374.

[9] Takeshi Harayama and Howard Riezman. “Understanding the diversity of membrane lipid composition”. In: *Nature Reviews Molecular Cell Biology* 19.5 (2018), pp. 281–296.

[10] John E. Burke and Edward A. Dennis. “Phospholipase A2 structure/function, mechanism, and signaling”. In: *Journal of Lipid Research* 50.Supplement (2009), S237–S242.

[11] Arthur Kahlenberg and Batya Banjo. “Involvement of Phospholipids in the D-Glucose Uptake Activity of Isolated Human Erythrocyte Membranes”. In: *The Journal of Biological chemistry* 247.4 (1972), pp. 1156–1160.

[12] D Oursel et al. “Lipid composition of membranes of Escherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization”. In: *Rapid communications in mass spectrometry : RCM* 21 (2007), pp. 1721–1728.

[13] Christian Sohlenkamp and Otto Geiger. “Bacterial membrane lipids: Diversity in structures and pathways”. In: *FEMS Microbiology Reviews* 40.1 (2015), pp. 133–159.

[14] Avanti lipids. *E. coli extract total*. URL: `https://avantilipids.com/product/100500`.

[15] William J. Griffiths. “Tandem mass spectrometry in the study of fatty acids, bile acids, and steroids”. In: *Mass Spectrometry Reviews* 22.2 (2003), pp. 81–152.

[16] G. Sweetman et al. “Electrospray ionization mass spectrometric analysis of phospholipids of Escherichia coli”. In: *Molecular Microbiology* 20.1 (1996), pp. 233–238.

[17] Giuseppe Paradies et al. “Functional role of cardiolipin in mitochondrial bioenergetics”. In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1837.4 (2014), pp. 408–417.

[18] Author Musatov, Andrej and Author Robinson, Neal C. “Bound Cardiolioin Is Essential for Cytochrome c Oxidase Proton Translocation”. In: *Biochimie* 18.9 (2014), pp. 1199–1216.

[19] Eugenia Mileykovskaya and William Dowhan. “Visualization of Phospholipid Domains inEscherichia coli by Using the Cardiolipin-Specific Fluorescent Dye 10-N-Nonyl Acridine Orange”. In: *Journal of Bacteriology* 182.4 (2000), pp. 1172–1175.

[20] G Dennis Sprott. “Archaeal Membrane Lipids and Applications”. In: *eLS* (2011).

[21] Masateru Nishihara, Hiroyuki Morii, and Yosuke Koga. “Structure Determination of a Quartet of Novel Tetraether Lipids from Methanobacterium thermoautotrophicum”. In: *The Journal of Biochemistry* 101.4 (1987), pp. 1007–1015.

[22] M Kates. “Archaebacterial lipids: structure, biosynthesis and function”. In: 58 (Feb. 1992), pp. 51–72.

[23] G. Dennis Sprott. “Structures of archaebacterial membrane lipids”. In: *Journal of Bioenergetics and Biomembranes* 24.6 (1992), pp. 555–566.

[24] Angela Corcelli. “The cardiolipin analogues of Archaea”. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1788.10 (2009), pp. 2101–2106.

[25] M. Kates. “Biology of halophilic bacteria, Part II”. In: *Experientia* 49.12 (1993), pp. 1027–1036.

[26] Mary F Roberts. “Biology of Inositols and Phosphoinositides”. In: *Springer* 39 (2006), pp. 103–134.

[27] Mario De Rosa and Agata Gambacorta. “The lipids of archaebacteria”. In: *Progress in Lipid Research* 27.3 (1988), pp. 153–175.

[28] Matthieu Chavent et al. “Interactions of the EphA2 Kinase Domain with PIPs in Membranes: Implications for Receptor Function”. In: *Structure* 26.7 (2018), 1025–1034.e2.

[29] Erik Wallin and Gunnar Von Heijne. “Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms”. In: *Protein Science* 7.4 (1998), pp. 1029–1038.

[30] K. Lundstrom. “Structural genomics and drug discovery”. In: *Journal of Cellular and Molecular Medicine* 11.2 (2007), pp. 224–238.

[31] HM. et al Berman. “Protein Data Bank `www.rcsb.org`”. In: *Nucleic Acids Research* 28 (2000), pp. 235–242.

[32] Roger Pain. “Determining the CD spectrum of a protein”. In: *Curr Protoc Protein Sci* Chapter 7 (Jan. 2005), Unit 7.6.

[33] César Fernández and Kurt Wüthrich. “NMR solution structure determination of membrane proteins reconstituted in detergent micelles”. In: *FEBS Lett* 555.1 (Nov. 2003), pp. 144–50.

[34] Vincent Breukels et al. “Overview on the use of NMR to examine protein structure”. In: *Curr Protoc Protein Sci* Chapter 17 (Apr. 2011), Unit17.5.

[35] Philip L. Yeagle. “Non-covalent binding of membrane lipids to membrane proteins”. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1838.6 (2014), pp. 1548–1559.

[36] B Kronberg, K Holmberg, and B Lindman. In: *Surface Chemistry of Surfactants and Polymers*. Wiley-Blackwell, 2014. Chap. 4, pp. 75–94.

[37] M. Lever. “Peroxides in detergents as interfering factors in biochemical analysis”. In: *Analytical Biochemistry* 83.1 (1977), pp. 274–284.

[38] Isabel Moraes et al. “Membrane protein structure determination - The next generation”. In: *Biochimica et Biophysica Acta - Biomembranes* 1838.1 PARTA (2014), pp. 78–87.

[39] Junhao Huang et al. “Investigation of modified sodium alginate-Alkyl glycoside interactions in aqueous solutions and at the oil-water interface”. In: *RSC Advances* 6.56 (2016), pp. 51068–51077.

[40] Robert S. Cantor. “Lipid composition and the lateral pressure profile in bilayers”. In: *Biophysical Journal* 76.5 (1999), pp. 2625–2639.

[41] Marcella Orwick-Rydmark, Thomas Arnold, and Dirk Linke. *The use of detergents to purify membrane proteins*. Vol. 2016. April. 2016.

[42] Stephen A. Waschuk et al. “Leptosphaeria rhodopsin: Bacteriorhodopsin-like proton pump from a eukaryote”. In: *Proceedings of the National Academy of Sciences* 102.19 (2005), pp. 6879–6883.

[43] Jun Sasaki and John L. Spudich. “Signal Transfer in Haloarchaeal Sensory Rhodopsin– Transducer Complexes†”. In: *Photochemistry and Photobiology* 84.4 (2008), pp. 863–868.

[44] H G Khorana et al. “Amino acid sequence of bacteriorhodopsin”. In: *Proc Natl Acad Sci U S A* 76.10 (Oct. 1979), pp. 5046–50.

[45] Dieter Oesterhelt and W Stoeckenius. “Rhodopsin-like Protein from the Purple Membrane of Halobacterium halobium”. In: *Nature New Biology* 233.39 (Sept. 1971), pp. 149–152.

[46] Dieter Oesterhelt and Walther Stoeckenius. “Isolation of the cell membrane of Halobacterium halobium and its fractionation into red and purple membrane”. In: *Biomembranes Part A*. Vol. 31. Methods in Enzymology. Academic Press, 1974, pp. 667–678.

[47] R. Henderson et al. “Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy”. In: *Journal of Molecular Biology* 213.4 (1990), pp. 899–929.

[48] Thomas Schreckenbach, Baerbel Walckhoff, and Dieter Oesterhelt. “Specificity of the retinal binding site of bacteriorhodopsin: chemical and stereochemical requirements for the binding of retinol and retinal”. In: *Biochemistry* 17.25 (1978), pp. 5353–5359.

[49] Flickr. *ChatCatPR*. URL: %5Curl%7Bhttp://www.flickriver.com/photos/writeoutside/3674132996/%7D.

[50] Antoinette M. Dummer et al. “Bacterioopsin-Mediated regulation of bacterioruberin biosynthesis in Halobacterium salinarum”. In: *Journal of Bacteriology* 193.20 (2011), pp. 5658–5667.

[51] Kenneth J. Rothschild et al. “Vibrational spectroscopy of bacteriorhodopsin mutants: Evidence that THR-46 and THR-89 form part of a transient network of hydrogen bonds”. In: *Journal of Biological Chemistry* 267.3 (1992), pp. 1615–1622.

[52] Oliver P. Ernst et al. “Microbial and animal rhodopsins: Structures, functions, and molecular mechanisms”. In: *Chemical Reviews* 114.1 (2014), pp. 126–163.

[53] Keiko Yoshimura and Tsutomu Kouyama. “Structural Role of Bacterioruberin in the Trimeric Structure of Archaerhodopsin-2”. In: *Journal of Molecular Biology* 375.5 (2008), pp. 1267–1281.

[54] D. Maclaurin et al. “Mechanism of voltage-sensitive fluorescence in a microbial rhodopsin”. In: *Proceedings of the National Academy of Sciences* 110.15 (2013), pp. 5939–5944.

[55] Keiichi Inoue et al. “Converting a light-driven proton pump into a light-gated proton channel”. In: *Journal of the American Chemical Society* 137.9 (2015), pp. 3291–3299.

[56] R. S. McIsaac et al. “Directed evolution of a far-red fluorescent rhodopsin”. In: *Proceedings of the National Academy of Sciences* 111.36 (2014), pp. 13034–13039.

[57] Martin K.M. Engqvist et al. “Directed evolution of gloeobacter violaceus rhodopsin spectral properties”. In: *Journal of Molecular Biology* 427.1 (2015), pp. 205–220.

[58] Nicholas C. Flytzanis et al. “Archaerhodopsin variants with enhanced voltage-sensitive fluorescence in mammalian and Caenorhabditis elegans neurons”. In: *Nature Communications* 5 (2014), pp. 1–9.

[59] Sergei P. Balashov et al. “Glutamate-194 to cysteine mutation inhibits fast light-induced proton release in bacteriorhodopsin”. In: *Biochemistry* 36.29 (1997), pp. 8671–8676.

[60] Thomas Marti et al. “Bacteriorhodopsin mutants containing single substitutions of serine or threonine residues are all active in proton translocation”. In: *Journal of Biological Chemistry* 266.11 (1991), pp. 6919–6927.

[61] Federico Melaccio, Nicolas Ferré, and Massimo Olivucci. “Quantum chemical modeling of rhodopsin mutants displaying switchable colors”. In: *Physical Chemistry Chemical Physics* 14.36 (2012), pp. 12485–12495.

[62] Nicola Coppedè et al. “An in vivo biosensing, biomimetic electrochemical transistor with applications in plant science and precision farming”. In: *Scientific Reports* 7.1 (2017), pp. 1–9.

[63] Juras Požela. "Technological and Physical Limitations on Transistor Miniaturization". In: *Physics of High-Speed Transistors*. Boston, MA: Springer US, 1993, pp. 35–47.

[64] Max Schulz. "The end of the road for silicon?" In: *Nature* 399.6738 (1999), pp. 729–730.

[65] Robert R. Birge. In: *Molecular and Biomolecular Electronics*. American Chemical Society: Washington D.C, 1994, p. 595.

[66] T Sienko et al. In: *Molecular Computing*. The MIT Press, 2003, p. 257.

[67] G E Moore. "Cramming more components onto integrated circuits". In: *Proceedings Of The Ieee* 86.1 (1965), pp. 82–85.

[68] J. J. Davis et al. "The tunnelling conductance of molecularly ordered metalloprotein arrays". In: *Journal of Molecular Recognition* 17.3 (2004), pp. 167–173.

[69] David J. Lavrich et al. "Physisorption and Chemisorption of Alkanethiols and Alkyl Sulfides on Au(111)". In: *The Journal of Physical Chemistry B* 102.18 (1998), pp. 3456–3465.

[70] Jason J. Davis et al. "Genetic modulation of metalloprotein electron transfer at bare gold". In: *Chem. Commun.* (5 2003), pp. 576–577.

[71] Amol V. Patil et al. "Enhanced Photocurrent in Engineered Bacteriorhodopsin Monolayer". In: *The Journal of Physical Chemistry B* 116.1 (2012), pp. 683–689.

[72] Olivia Berthoumieu et al. "Molecular Scale Conductance Photoswitching in Engineered Bacteriorhodopsin". In: *Nano Letters* 12.2 (2012), pp. 899–903.

[73] Amol V. Patil et al. "Engineered bacteriorhodopsin: A molecular scale potential switch". In: *Chemistry - A European Journal* 18.18 (2012), pp. 5632–5636.

[74] D Nicholls and S Ferguson. In: *Bioenergetics 3*. Academic Press, 2002, pp. 189–193.

[75] Olivia Berthoumieu. "Single Molecule Studies of Seven Transmembrane Domain Proteins". In: Doctoral Dissertation (2011).

[76] Feng Zhang et al. "Optogenetic interrogation of neural circuits: Technology for probing mammalian brain structures". In: *Nature Protocols* 5.3 (2010), pp. 439–456.

[77] Brian Y. Chow et al. "High-performance genetically targetable optical neural silencing by light-driven proton pumps". In: *Nature* 463.7277 (2010), pp. 98–102.

[78] Xue Han et al. "A High-Light Sensitivity Optical Neural Silencer: Development and Application to Optogenetic Control of Non-Human Primate Cortex". In: *Frontiers in Systems Neuroscience* 5 (2011).

[79] Viviana Gradinaru et al. "Molecular and Cellular Approaches for Diversifying and Extending Optogenetics". In: *Cell* 141.1 (2010), pp. 154–165.

[80] addgene. *WT Arch-3 in pET28b*. URL: `%5Curl%7Bhttps://www.addgene.org/58487/%7D`.

[81] Thermo Fisher. *Oligo Design Tools*. URL: `https://www.thermofisher.com/uk/en/home/life-science/oligonucleotides-primers-probes-genes/custom-dna-oligos/oligo-design-tools.html`.

[82] R Hevner. In: University of Washignton, 2001, pp. 1–3.

[83] Source Bioscience. *Cambridge UK*. URL: %5Curl%7Bhttps://www.sourcebioscience.com/services/genomics/sanger-sequencing-services/overnight-sequencing-service/%7D.

[84] Dmitry Bratanov et al. "An approach to heterologous expression of membrane proteins. The case of bacteriorhodopsin". In: *PLoS ONE* 10.6 (2015), pp. 1–16.

[85] Sarah C. Lee et al. "A method for detergent-free isolation of membrane proteins in their local lipid environment". In: *Nature Protocols* 11.7 (2016), pp. 1149–1162.

[86] Lee Whitmore and B A Wallace. "Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases". In: *Biopolymers* 89.5 (May 2008), pp. 392–400.

[87] W. Curtis Johnson. "Analyzing protein circular dichroism spectra for accurate secondary structures". In: *Proteins: Structure, Function, and Bioinformatics* 35.3 (1999), pp. 307–312.

[88] Ali Abdul-Gader, Andrew John Miles, and B A Wallace. "A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy". In: *Bioinformatics* 27.12 (June 2011), pp. 1630–6.

[89] M Duñach et al. "Uv-visible spectroscopy of bacteriorhodopsin mutants: substitution of Arg-82, Asp-85, Tyr-185, and Asp-212 results in abnormal light-dark adaptation." In: *Proceedings of the National Academy of Sciences of the United States of America* 87.24 (1990), pp. 9873–9877.

[90] Michael F Brown. "Curvature forces in membrane lipid-protein interactions". In: *Biochemistry* 51.49 (Dec. 2012), pp. 9782–95.

[91] C. A. Hasselbacher, John L. Spudich, and T. G. Dewey. "Circular Dichroism of Halorhodopsin: Comparison with Bacteriorhodopsin and Sensory Rhodopsin I". In: *Biochemistry* 27.7 (1988), pp. 2540–2546.

[92] Shuguang Wu and Mostafa A. El-Sayed. "CD spectrum of bacteriorhodopsin: Best evidence against exciton model". In: *Biophysical Journal* 60.1 (1991), pp. 190–197.

[93] Gennaro Pescitelli and Robert W. Woody. "The exciton origin of the visible circular dichroism spectrum of bacteriorhodopsin". In: *Journal of Physical Chemistry B* 116.23 (2012), pp. 6751–6763.

[94] Marcella Orwick-Rydmark et al. "Detergent-free incorporation of a seven-transmembrane receptor protein into nanosized bilayer lipodisq particles for functional and biophysical studies". In: *Nano Letters* 12.9 (2012), pp. 4687–4692.

[95] Mikihiro Shibata et al. "High-speed atomic force microscopy shows dynamic molecular processes in photoactivated bacteriorhodopsin". In: *Nature Nanotechnology* 5.3 (2010), pp. 208–212.

[96] I Lee et al. "Scanning Surface Potential Microscopy of Bacteriorhodopsin Mutant D96N". In: *Microscopy and Microanalysis* 11.S02 (2005), pp. 356–357.

[97] J Kyte and R F Doolittle. "A simple method for displaying the hydropathic character of a protein". In: *J Mol Biol* 157.1 (May 1982), pp. 105–32.

[98] ExPASy. *Protparam tool*. URL: https://web.expasy.org/protparam/.

[99] V. I. Prokhorenko. "Coherent Control of Retinal Isomerization in Bacteriorhodopsin". In: *Science* 313.5791 (2006), pp. 1257–1261.

[100] S Lund et al. "Detergent structure and associated lipid as determinants in the stabilization of solubulized Ca2+-ATPase from sarcoplasmic reticulum". In: *J. Biol. Chem.* 264.9 (1989), pp. 4907–4915.

[101] Y Mukai, N Kamo, and S Mitaku. "Light-induced denaturation of bacteriorhodopsin solubilized by octyl-beta-glucoside." In: *Protein engineering* 12.9 (1999), pp. 755–759.

[102] Sangbae Lee et al. "How Do Short Chain Nonionic Detergents Destabilize G-Protein-Coupled Receptors?" In: *Journal of the American Chemical Society* 138.47 (2016), pp. 15425–15433.

[103] Xianjun Zhang, Raymond C. Stevens, and Fei Xu. "The importance of ligands for G protein-coupled receptor stability". In: *Trends in Biochemical Sciences* 40.2 (2015), pp. 79–87.

[104] Francesca Magnani et al. "Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor". In: *Proceedings of the National Academy of Sciences* 105.31 (2008), pp. 10744–10749.

[105] Md. Iqbal Mahmood et al. "Influence of Lipid Composition on the Structural Stability of G-Protein Coupled Receptor". In: *Chemical and Pharmaceutical Bulletin* 61.4 (2013), pp. 426–437.

[106] Gilbert G. Privé. "Detergents for the stabilization and crystallization of membrane proteins". In: *Methods* 41.4 (2007), pp. 388–397.

[107] H Michel. "International Tables for Crystallography Volume F". In: Dordrecht: Kluwer Academic Publishers, 2001, pp. 94–100.

[108] Sue-Hwa Lin and Guido Guidoti. "Purification of Membrane Proteins, Chapter 35". In: *Methods in Enzymology*. New York: Elsevier, 2009, pp. 619–629.

[109] Jonas M. Dörr et al. "The styrene–maleic acid copolymer: a versatile tool in membrane research". In: *European Biophysics Journal* 45.1 (2016), pp. 3–21.

[110] Hayato Yamashita et al. "Role of trimer–trimer interaction of bacteriorhodopsin studied by optical spectroscopy and high-speed atomic force microscopy". In: *Journal of Structural Biology* 184.1 (2013), pp. 2–11.

[111] Anthony Watts. "Bacteriorhodopsin: the mechanism of 2D-array formation and the structure of retinal in the protein". In: *Biophysical Chemistry* 55.1 (1995), pp. 137–151.

[112] Takashi Tsukamoto, Makoto Demura, and Yuki Sudo. "Irreversible Trimer to Monomer Transition of Thermophilic Rhodopsin upon Thermal Stimulation". In: *The Journal of Physical Chemistry B* 118.43 (2014), pp. 12383–12394.

[113] Salem Faham et al. "Side-chain Contributions to Membrane Protein Structure and Stability". In: *Journal of Molecular Biology* 335.1 (2004), pp. 297–305.

[114] Keiichi Inoue et al. "A natural light-driven inward proton pump". In: *Nature Communications* 7.May (2016), p. 13415.

[115] Mikihiro Shibata et al. “Oligomeric states of microbial rhodopsins determined by high-speed atomic force microscopy and circular dichroic spectroscopy”. In: *Scientific Reports* 8.1 (2018), pp. 1–11.