Immobilizing strategies for membranes to screen against antibody mimetics with phage display

Vincent Ochejeni Agboh

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Biomedical Sciences
The Astbury centre of structural molecular biology

November 2015

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

Acknowledgments

I would like to thank my supervisor, Professor Stephen A. Baldwin, who remains instrumental to this PhD despite his unfortunate passing on 13th November 2014.

I would also like thank my supervisors Dr. Lars Jeuken and Dr. Stephen Muench for allowing me to work on this project, as well as their critical input, constant advice, guidance and support throughout. I am also thankful to Dr. Trevor Wilkinson, Dr. David Bannister, Dr. Stacey Chin and Dr. Benjamin Kemp, as well as all members of the biologics and protein sciences teams from MedImmune who trained me in phage display and the subsequent validation experiments. I am also grateful to MedImmune for providing funding.

Special thanks to all past and present members of the Baldwin, Henderson and Jeuken groups from the University of Leeds. Particularly Mrs Jocelyn Baldwin, Dr Cheng Ma, Dr Vincent Postis and Zhenyu Hao who trained me in stopped flow, provided samples of VcCNT and NupC SMALPs, as well as AcrB nanodiscs. Thanks also to David Sharples who performed 30 L fermentations of several *E.coli* host strains. Also, thank you to Ivona Petrache, for training me in the production of nanodiscs.

Thank you to Professor Tim Daffron, Dr Sarah Lee and Mrs. Rosemary Parslow from the University of Birmingham for providing the styrene maleic acid polymer and training me in its use.

I would finally like to thank my family, Dr. Chris Agboh, Mrs Regina Agboh, Dr. Kelvin Agboh, Mrs Katie Agboh-Davison, Mr. Michael Agboh-Davison, Frankie Agboh-Davison and Ms. Pat Huffer for all of their encouragement and support.

Abstract

Antibody mimetics are a novel class of potential therapeutic agents which improve on several limitations which hinder antibodies. As a result, they are the subject of increasing interest from both the pharmaceutical and academic sectors. They are commonly selected via phage display against an immobilised target. However the presentation of these targets, particularly in the case of membrane proteins, is often a rate limiting step in their production.

In this study, several commonly-used protein tags were compared in their ability to immobilise a model membrane protein - MPSIL0294, a metal ion transporter from *Enterococcus faecalis*, a streptavidin binding peptide tag, an Avitag which has undergone *in vivo* biotinylation and a chemically biotinylated C-terminal cysteine. A fourth method of immobilisation was also included, aspecific chemical biotinylation of lysine residues in AcrB, a component of an *E.coli* multidrug efflux pump. The tags were compared using a phage display assay, in which designed ankyrin repeat proteins (DARPins) from a naïve library were screened against both model membrane proteins, which were both solubilised in *n*-Dodecyl-β-D-maltopyranoside (DDM). With the exception of the C-terminal cysteine variant of MPSIL0294, no DARPins capable of binding could be selected by any of the other immobilisation techniques after two rounds of phage display. Furthermore, the C-terminal cysteine MPSIL0294 also yielded poor success rates and only two binding DARPins were identified. Despite showing a moderately high level of sequence similarity, these two DARPins showed different binding activities during validation experiments. The alternative tags under test however; do not appear to be beneficial in the selection of antibody mimetics against immobilised membrane protein targets.

Different methods of protein solubilisation were also compared in an identical manner: detergent solubilisation (in DDM), lipid nanodiscs and styrene maleic acid lipid particles (SMALP). These solubilisation methods were compared using a panel of membrane proteins composed of MPSIL0294, AcrB, NupC, an *E.coli* concentrative nucleoside transporter and VcCNT, NupCs homologue from *Vibrio Cholerae*. SMALPs of AcrB, VcCNT and NupC displayed the highest success rate for selecting DARPins capable of binding the respective targets. Subsequent validation tests showed that the best DARPins selected against AcrB SMALPs and nanodiscs behaved in a similar manner. SMALPs are relatively easy to produce, especially compared to nanodiscs so this study concludes that SMALPs are the best format for phage display.

Table of Contents

List of figures

List of tables

Abbreviations

AcrB - Acriflavine resistance protein B

AUC – Analytical ultracentrifugation

b – Cell path length

BV – Bed volume

C - Concentration

CAC – Critical association concentration

CDR – Complementarity determining region

cfu – Colony forming units

CMC – Critical micelle concentration

CNT – Concentrative nucleoside transporter

DARPins – Designed ankyrin repeat proteins

DMT1 – Divalent metal transporter 1

FACs - Fluorescence-activated cell sorting

HRP – Horseradish peroxidise

HRV-3C – Human rhinovirus 3C protease

IMAC - Immobilized metal ion affinity chromatography

ITC – Isothermal titration calorimetry

LeuT - Leucine transporter

MPSIL0294 – Membrane protein structure initiative at Leeds number 294

MSP – Membrane scaffold protein

NHS - nucleoside: H^+ Symporter

NMR – Nuclear magnetic resonance

NRAMP – Natural resistance associated macrophage protein

NupC – nucleoside permease

OD – Optical density

RND - Resistance–nodulation–division

SBP – Streptavidin binding peptide

ScDMT - *Staphylococcus capitis* divalent metal transporter

scFv – Single-chain variable fragment

SEC – Size exclusion chromatography

SLC11A1/SLC11A2 – Solute carrier family 11 member 1/ member 2

SMA – Styrene maleic acid

SMALPs – Styrene maleic acid lipid particles

SPR – Surface Plasmon resonance

SRP – Signal recognition particle

SUV – Small unilamellar vesicles

TM – Transmembrane helix

VcCNT – *Vibrio cholerae* concentrative nucleoside transporter

Chapter 1

Introduction

1.1 Membrane proteins in modern research

Membrane proteins are a class of biological molecules of intense interest in several different sectors of biological research today; none more so than the pharmaceutical sector. This is especially apparent since approximately 60% of all currently prescribed drugs target membrane proteins (Klammt, Maslennikov et al. 2012, Moraes, Evans et al. 2014). Of this 60%, increasing proportions are protein therapies (for example monoclonal antibodies) as opposed to small molecule drugs. In 2010 17% of the total profit made by the global pharmaceutical market was due to therapeutic proteins and this trend is predicted to continue (i.e. monoclonal antibodies and recombinant proteins combined)(Elvin, Couston et al. 2013). This amount of investment is indicative of the importance of treatments targeting membrane proteins.

There are several reasons why novel methods of selecting proteins capable of binding to membrane proteins would be of great interest, the level of investment being one. From a research point of view, novel methods of developing crystallisation chaperones (which are membrane protein binding partners) are of great interest, as they aid in crystallogenesis which in turn allows X-ray crystallography and protein structure determination to be performed. From the point of view of the pharmaceutical industry, improved methods of selecting antibodies and other binding proteins specific for membrane protein targets are of great interest as such methods would greatly benefit drug development. Their potential as drug targets is epitomised in the fact that 20 – 30% of the open reading frames in genomes across all life encode membrane proteins. For humans specifically it is estimated to be 26% (Krogh, Larsson et al. 2001, Fagerberg, Jonasson et al. 2010, Moraes, Evans et al. 2014).

The main difficulties that arise when attempting to research membrane proteins are their high sensitivity to environmental changes, due to the fact that they are amphipathic. In order to integrate properly into a cell membrane they require hydrophobic domains which can exist stably within the phospholipid bilayer. Membrane proteins also tend to have hydrophilic domains that exist outside of the membrane (either extracellularly, intracellularly or both), which will often play a role in protein function. Therefore they need to be kept within a membrane or membrane mimetic structure at all times. Furthermore, the surrounding lipids can interfere with experiments aimed at determining their structures such as x-ray crystallography or nuclear magnetic resonance (NMR) (Franzin, Gong et al. 2007). Detergents such as Triton X-100 or n-Dodecyl β-D-Maltopyranoside (DDM) are employed due to their ability to disorganize membranes, resulting in the creation of micelles (spherical aggregates formed whenever a membrane is treated with a surfactant) in which membrane proteins are

soluble when the detergent is present above a threshold concentration known as the critical micelle concentration (CMC). Thereby making them more easily extracted for subsequent techniques such as high throughput screens. However when membrane mimetic environments are employed (e.g. solubilisation by detergents) there is a risk that membrane proteins may adopt non-native conformations (Otzen 2011) (see **section 1.3.1**).

This chapter shall discuss the main themes which this thesis impacts, starting with a discussion of the most common methods of membrane protein solubilisation employed in current research including detergent solubilisation as well as detergent free methods such as the use of membrane scaffold proteins (MSP) in order to construct nanodiscs, or the use of polymers such as styrene maleic acid (SMA) or amphipols. This is followed by a brief profile of each of the membrane proteins under test in this study before several *in vitro* selection strategies are introduced. The strategies discussed include phage, bacterial and yeast display, all of which use entire cells in order to display proteins of interest. Followed by a discussion of cell free methods of selection, namely ribosome and mRNA display. . Various antibody mimetics are then discussed, namely single chain variable fragments (scFvs), nanobodies, designed ankyrin repeat proteins (DARPins) and adhirons before the chapter concludes with the overall aim of this study.

1.2 Solubilising membrane proteins

1.2.1 Detergent solubilisation

Solubilising detergents are chemicals which have the ability to mimic the lipid bilayer due to a hydrophilic 'head' region and hydrophilic hydrocarbon 'tail'. Above a certain threshold concentration (known as the critical micelle concentration – CMC), this amphipathicity allows detergents to form micelles into which membrane proteins can be encapsulated. The formation of micelles is spontaneous once a detergent is exposed to an aqueous environment as the hydrophobic tails move to minimise their interaction with water. Depending on the detergent used, solubilisation of membrane bilayers can either occur rapidly or slowly, depending upon the rate at which detergent is flipped from the outer membrane to the inner membrane (Lichtenberg, Ahyayauch et al. 2013). The overall mechanism of solubilisation differs between rapid and slow solubilisation. The rapid solubilisation process involves the initial rapid insertion of detergent monomer into the outer membrane, followed by the detergent monomers flipping into the inner membrane. Once both the inner and outer membranes have become saturated with detergent, compartments are then directly removed

either by the micelles in an 'all or nothing' manner via the trans-bilayer mechanism, or open vesicular intermediates. Thus, in rapid solubilisation it is believed that lipids and membrane protein are co-extracted with one another into thread like mixed micelles (Kragh-Hansen, le Maire et al. 1998, Stuart and Boekema 2007, Lichtenberg, Ahyayauch et al. 2013).

Significantly different mechanisms have been proposed in the literature for slow solubilisation, however, although it is generally agreed that it results in only part of the phospholipid in micelles. Micelle formation is thought to occur either as the result of the saturation of membrane vesicles from which micelles 'pinch off' (Stuart and Boekema 2007), or as result of direct binding of micelles to the bilayer and extraction of membrane components by the micelle itself (Kragh-Hansen, le Maire et al. 1998, Lichtenberg, Ahyayauch et al. 2013).

Detergents such as Triton X-100 and dodecyldimethylamine oxide (DDAO) are examples of fast solubilising detergents, while sodium dodecyl sulphate (SDS) and *n*-Dodecyl-β-D-maltopyranoside (DDM) are both examples of the slow solubilising class (Lichtenberg, Ahyayauch et al. 2013).

Detergents are classified based on their structure within four main classes. The first are ionic detergents which encompass detergents with a charged head group and a hydrophobic hydrocarbon chain such as SDS. These detergents are often effective solubilising agents, but also have a tendency to denature membrane proteins. The second class are the non-ionic detergents which have uncharged hydrophilic head regions composed of glycosidic groups, such as the DM, DDM and the other maltoside detergents, or polyoxyethylene group, such as Triton X-100. They are considered a milder class of detergents and are thus often capable of solubilising membrane protein in a functional state, although this effect is dependent upon the length of the hydrophobic chain. Thirdly, there are zwitterionic detergents such as DDAO, which exhibit features of both the ionic and non-ionic detergents, although they tend to inactivate membrane proteins more often. Finally, there are the bile acid salts such as sodium cholate, which exhibit a polar and apolar face instead of a hydrophobic head group. This is due to the steroidal groups which form their backbone. The bile acid salts are quite mild and these detergents have a lower tendency to denature proteins because their head groups are combined with their steroidal backbones (le Maire, Champeil et al. 2000, Seddon, Curnow et al. 2004).

The choice of detergent for protein solubilisation may have an effect on protein- protein interactions with membrane proteins. Depending on the length of the hydrophilic chain, the

resultant micelle may be large enough to engulf potential surface epitopes thus rendering them inaccessible to potential binding partners (Kunji, Harding et al. 2008). The effect of micelle size upon selection experiments, an overview of which is presented in **section 1.3,** has not yet been investigated in the literature and remains an unknown factor in the selection of binding proteins to protein targets.

As previously stated, detergents can interfere with several experimental techniques that would provide valuable information regarding membrane proteins, such as mass spectrometry, NMR and X-ray crystallography. Therefore, several alternative methods of membrane protein solubilisation have arisen which aim to more closely mimic the native membrane environment, several of which will now be discussed.

1.2.2 Nanodiscs

Nanodiscs are membrane mimetics composed of 130-160 phospholipids arranged in a bilayer surrounded by two membrane scaffold proteins (MSP) derived from apolipoprotein 1 in a double belt configuration. They form spontaneously without the need of detergent and appear as flat disks, approximately 10-20 nm in diameter (depending on the MSP used) with a mass of approximately 150 kDa. The entire nanodisc population adopts the same specific size and is monodisperse once their construction is complete (Bayburt and Sligar 2010, Rajesh, Knowles et al. 2011).

The MSP can be altered in several ways prior to nanodisc formation, including biotinylation, tagging via polyhistidine, FLAG tags etc. with cleavage sites for later removal. The size of the nanodisc is dependent on the length of MSP which is used in its construction. This has been exploited and led to the generation of several different MSPs with extra helices inserted in the centre of the protein, thereby increasing the size of the nanodisc (Denisov, Grinkova et al. 2004, Schuler, Denisov et al. 2013). A result of these MSP alterations showed that the first 22 residues of the MSP may not be involved the formation of the discs as their deletion failed to affect its size (Grinkova, Denisov et al. 2010).

To entrap a membrane protein within a nanodisc, the protein first needs to be solubilised from the cell plasma membrane with detergent and subsequently purified in detergent. The general procedure involves treating the purified membrane protein (in detergent) with MSP and phospholipids in a very precise ratio. This ratio is specific to the membrane protein of interest. Once the detergent is removed, the membrane protein is drawn into the nanodisc as it

spontaneously forms, as the protein, MSP and lipid all seek to minimise their hydrophobic regions interaction with ambient water (see **Figure 1.1A**).

In theory, insertion into a nanodisc should increase the likelihood of the protein retaining its native conformation and functionality due to the presence of a phospholipid bilayer, the lipids of which can be altered to better suit its native environment if need be (Bayburt and Sligar 2010). However, if the protein adopts a non-native confirmation as a result of detergent solubilisation, presumably this confirmation will be carried forward into the nanodisc.

Nanodiscs also aid in membrane protein research due to their homogeneity, consistent size and minimal light scattering. Most importantly, however, is the fact that once a membrane protein is encapsulated within them, they can be handled like soluble protein in an aqueous solution, thus making the use of techniques such as surface plasmon resonance (SPR) or NMR easier (Justesen Bo and Günther-Pomorski 2014).

Nanodiscs have the potential to be a very powerful tool not only in membrane protein research but also in drug discovery, as they not only stabilise the proteins but also should aid in their screening. For example, the fact that any lipid can be used in their production means that, excessive tagging of proteins may be unnecessary as they can be encapsulated in nanodiscs formed with biotinylated lipid and immobilised on streptavidin surfaces this way (Goluch, Shaw et al. 2008, Bayburt and Sligar 2010). They also allow access to both sides of the lipid bilayer (Justesen Bo and Günther-Pomorski 2014) and thus both sides of the encapsulated protein. This, in theory, maximises the selected of binding proteins against a target. To this end, nanodiscs containing bacteriorhodopsin have been used in a phage display screen using the random 12-mer library of binding peptides (a commercially available library consisting of 2.7×10^9 random dodecapeptides). This selection resulted initially in 69 peptides which were considered positive. After further validation, three were characterised further and shown to bind to the extramembranous regions of bacteriorhodopsin (Pavlidou, Hanel et al. 2013).

1.2.3 Polymer solubilisation

Amphipathic polymers are an alternative class of compounds, some of which have shown a capability to solubilise membrane proteins. An early example of these polymers are the amphipols which are composed of hydrophilic backbones grafted with hydrophobic side chains. Despite their ability to solubilise membrane proteins, they still require the proteins to initially be purified in detergent (see **section 1.2.3.2**). Alternatively In recent years, the

amphipathic co-polymer styrene-maleic acid has been shown to remove membrane proteins directly from membranes, solubilising them in a pH dependent manner (see **section 1.2.3.1**).

1.2.3.1 Styrene-maleic acid

Styrene-maleic acid (SMA), an amphipathic co-polymer, is capable of forming disc like structures around membrane proteins due to its ability to associate with lipids. The copolymer itself is composed of alternating hydrophobic styrene groups and hydrophilic maleic acid groups. The polymer is able to remove membrane proteins directly from membranes without the inclusion of detergent at any point (Rajesh, Knowles et al. 2011). SMA is actually the hydrolysed form of styrene-maleic anhydride copolymer, a commonly used substance in the manufacture of plastics.

The procedure required to encapsulate membrane proteins involves directly treating membrane preparations with the SMA polymer. This causes the spontaneous entry of SMA into the membrane in a pH-dependant manner creating disc like structures known as styrene-maleic acid lipid particles (SMALPs), which are 9-12 nm in diameter (see **Figure 1.1B**). When the pH drops below 6 the SMA becomes insoluble due to protonation of the maleic acid and the SMALPs fall apart; when the pH is increased again, the SMALPs reform.

There is a major advantage to using SMALPs over nanodiscs, namely the lack of an MSP. From a drug discovery point of view this gives advantages in terms of screening, as binding partners that associate with the MSP instead of the target protein will no longer be an issue. Also the fact that SMALPs are generally smaller and scatter less light, makes them better suited for biophysical studies such as circular dichroism (CD)(Knowles, Finka et al. 2009, Jamshad, Lin et al. 2011, Rajesh, Knowles et al. 2011). A major benefit of SMALPs is the fact that the co-polymer encapsulates membrane proteins directly from the membrane therefore, it presumably retains its native structure and function (Postis, Rawson et al. 2015). It is theorised that the benefits of SMALPS arise due to their ability to encapsulate the lipids in the immediate vicinity of the protein along with the target, although this is currently not known. Due to their novelty, investigations into their suitability as phage display platforms have not yet been published. They have, however, proven to be beneficial during electron microscopy (Postis, Rawson et al. 2015).

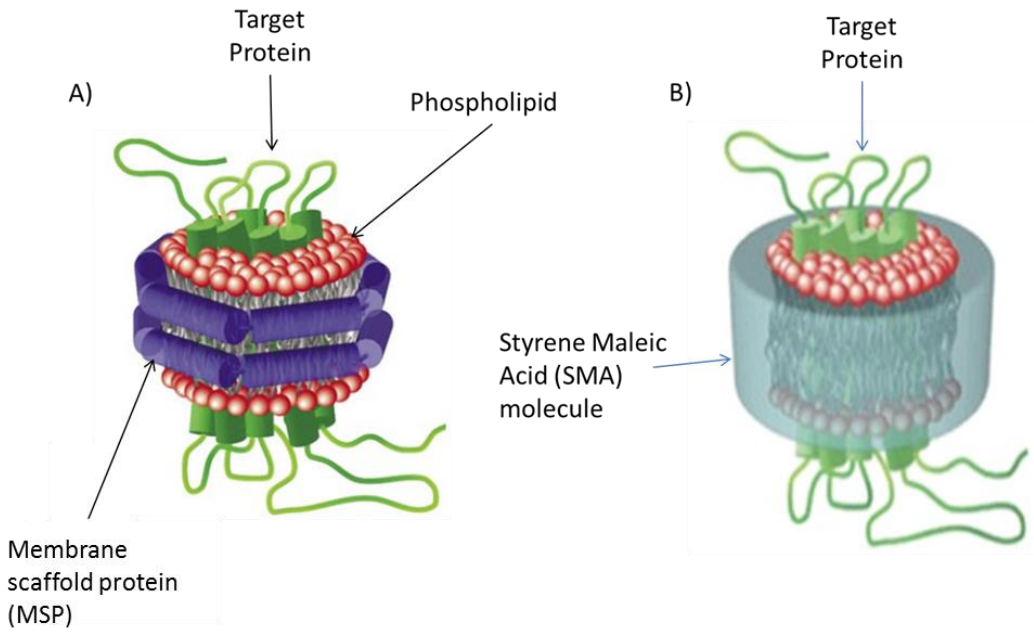


Figure 1.1: Cartoon models of a Nanodisc and a SMALP: A) a nanodisc encapsulating the target GPCR (Beta2-adrenergic receptor) and B) a SMALP encapsulating the same target protein. Original images taken from: A) the paper "Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology" (Leitz, Bayburt et al. 2006) and B) From the University of Warwick, Molecular Organisation and Assembly in Cells website: "Novel biophysical methods for membrane protein research" under the supervision of Dr David I Roper.

1.2.3.2 Amphipols

Amphipols are manufactured flexible amphipathic polymers consisting of a highly hydrophilic backbone grafted with many closely spaced hydrophobic side chains. This allows them to associate and solubilise membrane proteins by forming a belt around their hydrophobic regions, thereby protecting these regions from exposure to aqueous surroundings, while simultaneously burying its own hydrophobic side chains (Tribet, Audebert et al. 1996, Zoonens and Popot 2014). Amphipols are generally rather short in order to maximise the homogeneity of an amphipol/protein solution and minimise the amount of cross bridges between individual amphipol/protein complexes (Zoonens and Popot 2014).

Early amphipols consisted of a polyacrylic acid backbone grafted with octylamine and isopropylamine side chains in arbitrary positions. A grafting ratio of 25% was shown to confer a desirable level of amphipathicity while retaining the polymers solubility in water. This synthesis gave rise to the most prominent of the amphipols – A8-35 (35% free, ungrafted carboxylic acid groups), which has become the most extensively researched (Tribet, Audebert et al. 1996, Zoonens, Zito et al. 2014). In aqueous solution, A8-35 forms 40 kDa globular particles containing four A8-35 macromolecules and approximately 80 octyl chains above a critical association concentration (CAC) of 0.002 g L^{-1} (Gohon, Giusti et al. 2006, Giusti, Popot et

al. 2012). More recently, entire libraries of amphipols have been developed which aim to maximise their potential applications. For instance, substituting the isopropyl groups for sulfonate groups gave rise to a range of amphipols resistant to pH change, designated the sulfonated amphipols (Dahmane, Giusti et al. 2011). Alternatively, zwitterionic amphipols were synthesised by the inclusion of phosphocholine polar head groups which are resistant to low pH, high salt concentrations and the presence of divalent ions (Diab, Tribet et al. 2007). Non-ionic amphipols have also been produced by the incorporation of sugar groups (Sharma, Durand et al. 2012). Similarly the amphipol polymer can be directly tagged, biotinylated or associated with fluorophores (Le Bon, Popot et al. 2014).

However, amphipols are weak detergents, while they are able to permeabilize lipid vesicles and adsorb onto cell surfaces, they are unable to directly remove proteins from a membrane (Popot, Berry et al. 2003, Zoonens and Popot 2014). Although it has been suggested that A8-35 was capable of directly removing the *E.coli* maltose transporter as well as the human insulin transporter (which was over expressed in Chinese hamster ovary (CHO) cells) from their respective membranes, these instance are few and far between (Popot, Berry et al. 2003). In most cases, the solubilisation of proteins by amphipols is very similar to the production of nanodiscs, in that detergent solubilised protein is mixed with amphipol and the concentration of detergent is lowered below the CMC, often with bio-beads (Popot, Berry et al. 2003, Zoonens, Zito et al. 2014). Like nanodiscs, the optimal amphipol to protein ratio is required in order to keep the protein soluble, however for amphipols it is determined experimentally. The ratio is dependent on several factors including: the size of the protein and transmembrane region and the tendency of the protein to oligomerize. Initially, an excess of amphipol is typically added to the detergent-solubilised protein in order to form protein/micelle/amphipol complexes which are monodisperse, due to their weak dissociating power (Zoonens and Popot 2014, Zoonens, Zito et al. 2014). Once the amphipol/membrane protein complexes are formed, some polymer will be left as free particles in the sample. When large membrane proteins (over 40 kDa) are used, this free polymer can be removed by size exclusion chromatography (SEC). Alternatively the excess amphipol can be removed by affinity chromatography such as immobilised metal ion affinity chromatography (IMAC) if the protein is tagged. Thirdly the free polymer can be removed by analytical ultracentrifugation (AUC) (Zoonens, Zito et al. 2014). Despite the fact that amphipol protein complexes are more stable than their detergent micelle counterparts, they were not included in this study.

1.3 The membrane protein panel under test

The panel of membrane proteins selected for this study were chosen in order to provide a range of oligomeric states and extramembranous regions. The first is predicted to be a divalent metal transporter which was highlighted as part of the membrane protein structure initiative in Leeds (MPSIL) designated MPSIL0294. There is not currently a crystal structure for MPSIL0294 therefore it is difficult to predict protein's accessibility to binding probes during phage display. The second member of the panel is a component of a major *E.coli* multidrug efflux pump called AcrB. This protein (which functions as a trimer in *E.coli*) has previously been subjected to phage display with DARPins and isolated high affinity binders, some of which were used in co-crystallisation trails. Also AcrB is well characterised and has a large extracellular domain which should be highly accessible to potential binding partners (Brandstatter, Sokolova et al. 2011, Eicher, Cha et al. 2012, Johnson, Cheong et al. 2012). Finally, a concentrative nucleoside transporter (CNT) from *Vibrio Cholerae* (VcCNT) along with its *E.coli* homologue NupC are included in the panel of membrane proteins. The crystal structure of VcCNT show that it exists as a trimer and therefore provides a good model of multimeric proteins during phage display. The background of each of the membrane target will be introduced separately below.

1.3.1 MPSIL0294

MPSIL0294 is a divalent metal cation transporter from *Enterococcus faecalis* and is a member of the NRAMP family of metal transporters, although the metal ions which it is able to transport remains unclear. It is more commonly known as Ef MntH Cβ1 and research into its function have shown that MPSIL0294's overexpression results to an increase in sensitivity to cadmium, cobalt and iron metal ions in *E. coli*. Intriguingly, no increase in sensitivity was reported for manganese (which is thought to be a key ion transported by the MntH proteins) (Richer, Courville et al. 2003). Conversely, recent investigations carried out in *Enterococcus faecalis* showed that MPSIL0294s expression is downregulated in the presence manganese ions (and to a lesser extent copper ions) and upregulated in the presence of iron and zinc (Abrantes, Kok et al. 2013). This investigation also suggested that the MPSIL0294 gene is under the control of the EfaR regulator, a co-repressor which, in the presence of manganese, is able to bind to a DNA binding motif which encompasses the promotor, thus suggesting that MPSIL0294 is a manganese transporter in *Enterococcus faecalis* (Abrantes, Kok et al. 2013).

MPSIL0294 was included in this study due to the fact that both its N- and C-termini are cytosolic, potentially making them easier to attach affinity tags to without interfering with

biosynthesis or function. It has also proven difficult to produce crystals of MPSIL0294, therefore it is hoped that the use of crystallisation chaperones will aid in determining its structure. Its predicted 12 transmembrane helix (TM) topology differs from most of its bacterial homologues, which have an 11 TM topology and a periplasmic C-terminus. With this topology, it more closely resembles its human homologues in the divalent metal cation transporter or "Natural Resistance-Associated Macrophage Protein" (NRAMP) family with which it shares a 34% sequence homology (White, Stewart et al. 2004, Ma 2013). The latter, which undergoes extensive post-translational N-linked glycosylation, is more difficult to express and work with than MPSIL0294. In this study, MPSIL0294 was used to compare three of the four methods of immobilisation (the fourth is aspecific biotinylation, which was performed on AcrB, see **section 1.6**). Samples of MPSIL0294 were also encapsulated in nanodiscs and SMALPs, and compared after rounds of phage display using a 7.5×10^9 library of DARPins. A DARPin which successfully bound MPSIL0294 was also included in the proteins stopped flow Zn^{2+} uptake assay with the hope of testing its effect on protein function. Relatively little is known about MPSIL0294 specifically, so instead, the background of the NRAMP family will be discussed, in particular the human NRAMP.

While a structure of neither MPSIL0294 nor NRAMP is currently available, a crystal structure of a close homologue of NRAMP2 (also known as divalent metal transporter 1- DMT1) from *Staphylococcus capitis* (ScDMT) was solved by co-crystallisation, using nanobodies which bound to the exposed C-terminal of the protein (Ehrnstorfer, Geertsma et al. 2014).

Structurally, ScDMT is composed of eleven transmembrane helices, the first and sixth of which are interrupted by a short loop and are directly involved in substrate translocation. The structure displays the LeuT fold in which two five helix units are structurally related to one another forming a ten helix subdomain (the eleventh helix of ScDMT is not involved in the fold) (Khafizov, Staritzbichler et al. 2010). In ScDMT the five helix units show a twofold rotation around an axis in the centre of the membrane. It is formed of a sequence which is 37% identical to NRAMP2 and 59% homologous, its structure shows that it has a cytoplasmic N-terminal and a periplasmic C-terminal. Transport experiments with ScDMT which had a 44 residues from its N-terminal removed in order to stabilise it, showed that the protein was able to transport manganese and cadmium with micromolar affinity (similar to wild type) (Ehrnstorfer, Geertsma et al. 2014).

1.3.1.1 NRAMP

Two members of the NRAMP family are encoded within the human genome, NRAMP1 and NRAMP2 (also known as Solute Carrier Family 11 members 1 and 2 (SLC11A1 and SLC11A2)). The two are very similar in amino acid sequence (approximately 61% identity) and in hydropathy profiles (Pinner, Gruenheid et al. 1997, Mackenzie and Hediger 2004). Despite seemingly performing the same function of proton-dependant divalent metal ion transport, they differ from one another in some regards, the most notable being their localization in the body.

NRAMP1 is found on the phagolysosomal membrane within phagocytes and appears to play a role within the immune response (Lam-Yuk-Tseung, Govoni et al. 2003). It is likely that it functions primarily following phagocytosis of bacteria such as *Mycobacterium*. Phagocytosis leads to the formation of a phagosome which fuses with a lysosome to form a phagolysosome. The NRAMP1 protein is located in lysosomal membranes, as indicated by the presence of lysosomal targeting motifs within its sequence and through co-localisation experiments using full length and truncated NRAMP1 tagged with GFP and the lysosomal marker, lysosomal-associated membrane protein 1 (LAMP1), tagged with RFP (Cheng and Wang 2012). Its localisation suggests that NRAMP1 is placed into the phagolysosomal membrane upon phagosome/lysosome fusion.

NRAMP1 is subsequently thought to contribute to the establishment of detrimental environmental conditions within the phagosome which result in the encased pathogens adopting either a bacteriostatic or bacteriolytic state (Nada Jabado 2004). There are multiple theories concerning the mechanisms by which NRAMP1 establishes such conditions. For example, research suggests that recruitment of NRAMP1 results in the co-transport of protons and divalent metal ions such as Mn^{2+} or Fe^{2+}, essential for bacterial growth, out from the phagolysosome (see **Figure 1.2**). This is plausible as acidification of the phagolysosome could activate the efflux activity of NRAMP1 (Mackenzie and Hediger 2004, Fortin, Abel et al. 2007). It has also been reported that NRAMP1's presence within this membrane may help to prevent *Mycobacterium* from blocking the fusion of the phagosomes with lysosomes (a defence mechanism that several *Mycobacteria* strains are thought to employ) (Frehel, Canonne-Hergaux et al. 2002).

Alternatively, it has been suggested that with the influx of Fe^{2+} into the phagolysosome is the primary function of NRAMP. It is proposed that NRAMP1 is able to generate hydroxyl radicals via the Fenton/Haber-Weiss reactions using a proton/ferrous cation antiporter mechanism in

which Fe^{2+} ions are pumped into the phagolysosome (Lafuse, Alvarez et al. 2000, Goswami, Bhattacharjee et al. 2001). Despite the theoretical plausibility of this mechanism (contradictory evidence supporting both the co-transporter and antiporter activities of NRAMP1 is present in the literature), there are some fundamental problems with this hypothesis. Fe^{2+} is an essential metal ion for bacterial growth; importing it would thus aid bacterial survival within the phagolysosome. Also, more importantly, in order to achieve the aforementioned H^+/Fe^{2+} antiporter mechanism, NRAMP1 would need to be topologically opposite in the phagolysosome compared to the topology of NRAMP2 in the plasma membrane (see below). Due to the high level of structural similarity between NRAMP1 and NRAMP2, such a drastic difference in function and topology seems unlikely (Lam-Yuk-Tseung, Govoni et al. 2003, Nada Jabado 2004).

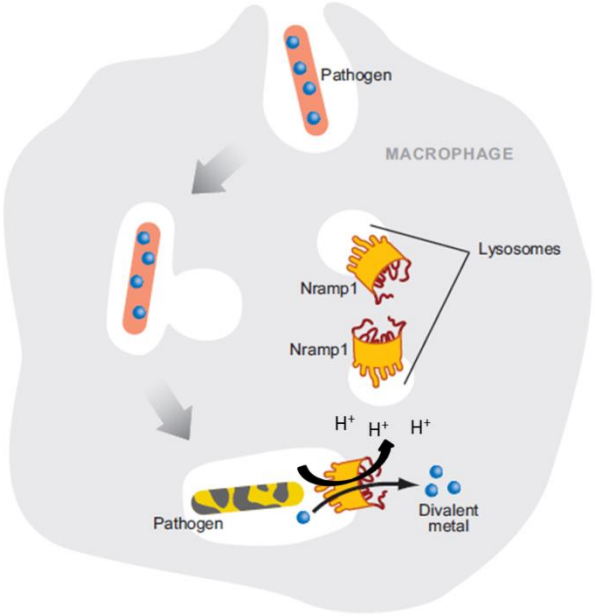


Figure 1.2: <u>NRAMP 2's role in the immune system:</u> A schematic of a possible mechanism by which NRAMP1 helps the immune system to kill phagocytised pathogens, using proton/divalent metal cation co-transporter activity. Note: Picture was adapted from a figure appearing in the paper "Host Genetics of Mycobacterial Diseases in Mice and Men: Forward Genetic Studies of BCG-osis and Tuberculosis" (Fortin, Abel et al. 2007)

NRAMP2 (also known as either Divalent Metal Transporter 1 (DMT1)) is far more ubiquitously found in the human body and its mechanism of transport has been better characterized. It is present in a variety of tissues ranging from the erythrocytes to intestinal cells (Lam-Yuk-Tseung, Govoni et al. 2003). There has even been some evidence to suggest that NRAMP2 may play a role in transporting Mn^{2+} across the blood brain barrier as it is found in the membranes of the astrocytes (Aschner 2006). In Mice, NRAMP2 was found to localize in the brush border of the absorptive epithelium in the duodenum, where it is thought to be responsible for the

cellular uptake of dietary Fe^{2+} (Jabado, Cuellar-Mata et al. 2003, Lam-Yuk-Tseung, Govoni et al. 2003, Mackenzie and Hediger 2004). An endosomal proton gradient is established by the Na^+/H^+ exchanger, thereby creating an acidic microclimate at the duodenum brush border membrane (Mackenzie and Garrick 2005). This microclimate provides NRAMP2 with plenty of available protons for co-transport with Fe^{2+} into the enterocyte where it can be stored in the intracellular iron pool (also known as the labile iron pool). However, since free iron in the blood is usually transported in its ferric form (Fe^{3+}) bound to the iron transporter transferrin and iron enters cells via the transferrin receptor (see **Figure 1.3**), it is unclear if NRAMP2 in the plasma membrane plays a role at cellular uptake of iron in tissues other than the duodenum (Trinder, Fox et al. 2002, Zhang, Canonne-Hergaux et al. 2008). NRAMP2 is more typically localised in endosomal membranes, where it is responsible for the translocation of essential metal ions such as Fe^{2+} from the endosome into the cytoplasm (see **Figure 1.3**). Ferritin is endocytosed with iron in its ferric form and the endosomal Fe^{3+} is reduced to Fe^{2+} prior to transportation by NRAMP2 (Sendamarai, Ohgami et al. 2008). The endosomal metalloreductase STEAP3 is thought to be responsible for the reduction of ferric iron (see **Figure 1.6**). However, STEAP3 has also been shown to have the ability to bind a large range of other divalent metals including Mn^{2+}, Co^{2+}, Cd^{2+}, Zn^{2+} and Pb^{2+}.

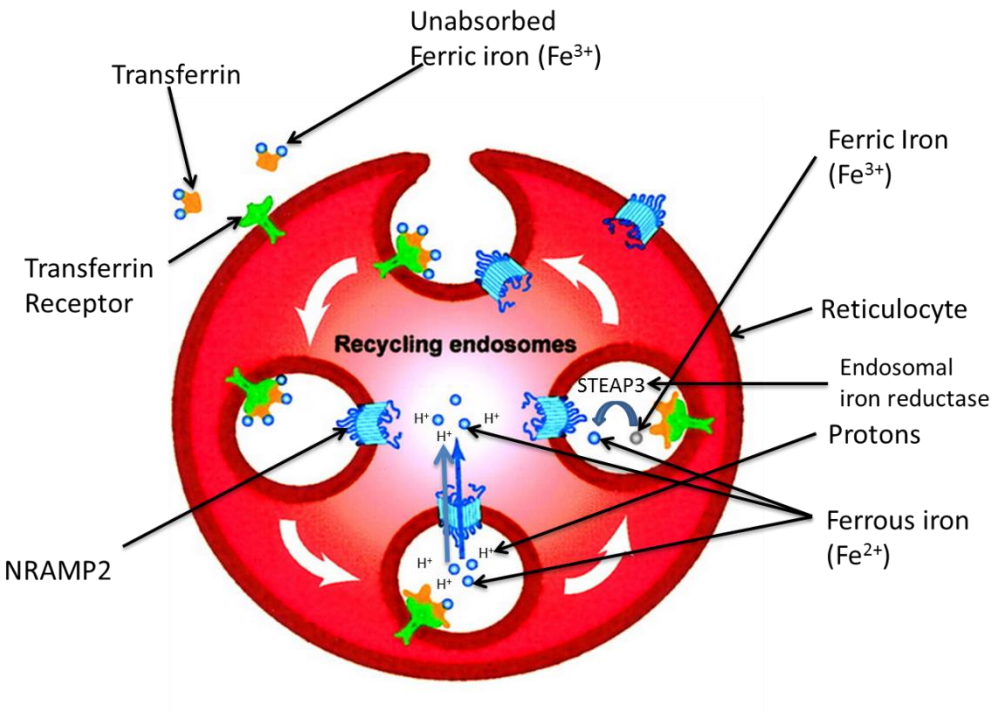


Figure 1.3: NRAMP2's role in iron absorption: A diagrammatic representation of the role of NRAMP2 in co-transporting protons and ferrous iron from the endosome into the cytosol, following liberation of ferric iron from endocytosed transferrin/transferrin receptor complexes and its reduction by STEAP3. Note: this picture was adapted from an image found in the paper "Single gene effects in mouse models of host: pathogen interactions" (Fortier, Min-Oo et al. 2005).

NRAMP2 activity is pH dependant and is activated by an acidic pH, thereby providing further evidence for the direction of transport. A high concentration of protons is required before NRAMP2 is able to co-transport metal ions out of the endosome, utilising the electrochemical gradient which is the cumulative force produced by the different concentration of metal ions across the membrane and the electrical difference produced by the charge of the ions (Lam-Yuk-Tseung, Govoni et al. 2003).

1.3.2 AcrB

AcrB is a large component of a multi drug efflux pump in *E. coli,* alongside the outer membrane channel TolC and the membrane fusion protein AcrA. AcrB is composed of a 12 helix transmembrane domain and a very large extracellular loop which forms the porter domain and the TolC docking domain. The AcrB/TolC/AcrA complex is a member of the resistance–nodulation–division (RND) family of transporters which are spread throughout Gram-negative bacteria. As such, it is able to export a wide array of substrates including cationic, anionic, zwitterionic and neutral compounds and is major cause of drug tolerance in *E. coli* (Murakami, Nakashima et al. 2006).

AcrB self-associates into a homotrimer, with each monomer providing components of the various extracellular domains. The porter domain consists of a bundle of helices which play a role in the proteins function. In the TolC docking domain, the protein opens like a funnel into which TolC (a trimeric β-barrel channel protein which extends from the outer membrane into the periplasm) docks directly. The AcrB/TolC complex is surrounded by AcrA which interacts with both components of the complex and acts as a linker between the two (Mikolosko, Bobyk et al. 2006, Murakami, Nakashima et al. 2006, Pos 2009).

The efflux mechanism is energised by a proton motive force and occurs in a substrate/proton antiporter manner. The leading theory of drug export is named the peristaltic pump mechanism, and proposes a three part rotary procedure conceptually similar to that of ATPase, without the physical rotation. The three protomers of AcrB are thought to adopt three distinct states; the 'loose' state, the 'tight' state and the 'open' state. Substrates from the outer leaflet of the inner membrane bind to the 'loose' protomer which, as a consequence, undergoes a transition into the 'tight' conformation. This transition is dependent upon an 11 amino acid residue switch loop, the flexibility of which appears to determine the ease of this transition and the size of compound which can be translocated (Eicher, Cha et al. 2012, Cha, Muller et al. 2014). Finally by utilising energy procured from the translocation of periplasmic protons by the transmembrane domain, the 'tight' conformation transitions into the open conformation and the drug is released into a central funnel towards TolC. This transition is also thought to be dependent on substrate binding into the adjacent 'loose' site. The porter domain is then thought to constrict, thereby driving unidirectional translocation of the substrate toward TolC and out of the cell (Murakami, Nakashima et al. 2006, Pos 2009, Nakashima, Sakurai et al. 2011, Eicher, Cha et al. 2012).

AcrB was selected for this study due to the ease with which it can be overexpressed and purified in *E.coli*. More importantly, the large extracellular loop of AcrB increases the likelihood of selecting DARPins which bind tightly to this domain. Due to a cluster of histidine residues in the C-terminal of AcrB, it is a very common *E.coli* contaminant when IMAC is used (Veesler, Blangy et al. 2008). Therefore a DARPin capable of binding AcrB tightly may be used to develop a column capable of specifically removing it from *E.coli* membrane preparations. As part of this study, samples of purified AcrB were aspecifically biotinylated via their lysine residues using NHS-biotin. They were also encapsulated in nanodiscs constructed with an MSP which itself had been biotinylated and prepared in SMALPs. Like MPSIL0294, each AcrB sample was compared with one another using two rounds of phage display with a DARPin library.

1.3.3 NupC and VcCNT

In *E. coli*, nucleosides can be used as a precursor to nucleotides or as a carbon and nitrogen source. Nucleosides are able to freely diffuse across the *E. coli* outer membrane via the specific porin, Tsx (Bremer, Middendorf et al. 1990). In low concentrations however, nucleosides have to be actively translocated across the inner membrane via two unrelated families of transporters; the concentrative nucleoside transporters (CNT) and the nucleoside: H^+ Symporter (NHS) family.

NupC is a member of the CNT family of transporters and is a proton linked nucleoside symporter. Experimental evidence has highlighted the properties which nucleosides must have in order to be successfully translocated by NupC. Firstly, C-3' on the sugar group must be in the α-configuration with a hydroxyl group. Secondly, the glyosidic bond must be in a β-configuration. It has also been shown that, for as yet unknown reasons, NupC is unable to translocate guanosine or deoxyguanosine and is poor at translocating inosine. Due to NupCs 22 and 26% sequence homology with the human CNT proteins (hCNT1 and hCNT2 respectively), a crystal structure of NupC is more therapeutically relevant than any of its *E.coli* homologues, as it is the only *E. coli* nucleoside transporter which shows such a similarity (out of both the CNT and NHS families) (Patching, Baldwin et al. 2005, Vaziri, Baldwin et al. 2013). It also transports anti-viral drugs such as azidothymidine (AZT) which is used in the treatment of AIDS and gemcitabine which is a chemotherapy agent (Loewen, Yao et al. 2004).

While there is no structure of NupC currently available, a crystal structure for its homologue in *V. cholerae,* designated VcCNT has been elucidated. Unlike NupC, VcCNT is a sodium linked concentrative nucleoside transporter which shows 39% sequence homology with hCNT3. The VcCNT structure suggests that the protein exists as a homotrimer composed of protomers with eight transmembrane helices each, two helix-turn helix hairpin motifs and three helices which run along the membrane known as interfacial helices (Johnson, Cheong et al. 2012).

The VcCNT protomers are separated into different sub-domains based on the positions of the helices relative to the centre of the transporter. The first is located toward the outside of the protein and is known as the scaffold domain. It is responsible for maintaining the transporters structure and plays a role in the trimerization process. The domain is composed of four transmembrane helices and a single aliphatic interfacial helix. Secondly, each protomer has a transport domain toward the centre of the protein which is composed of two structural groups, each composed of two transmembrane helices, a single helix-turn-helix hairpin motif

and an interfacial helix. The two structural groups are separated from one another by the sixth transmembrane helix which sits at an axis across the membrane (Johnson, Cheong et al. 2012).

Based on the high degree of conservation, the two turn-helix-turn motifs and the two transmembrane helices immediately adjacent to them (the 4th and the 7th respectively) are believed to be essential in substrate translocation. The substrate binding site is formed by the two structural groups within the transport which form deep clefts toward the centre of the protein (Johnson, Cheong et al. 2012).

Both NupC and VcCNT were selected in this project due to the extensive amount of characterisation which had already been performed, including established expression and purification procedures for both proteins and a transport assay for NupC. As previously stated while a crystal structure is available of VcCNT, none yet exists for NupC. Therefore the selection of DARPins which could be subsequently used in co-crystallisation trails with NupC were also of interest. During this study, samples of NupC and VcCNT SMALPs were provided and compared with one another during a phage display screen using the aforementioned DARPin library, with the aim that DARPins capable of binding the targets could be used during future co-crystallisation trails.

1.4 *In vitro* selection strategies

Several different examples of *in vitro* selection strategies currently exist and each strategy differs from one another in their efficiency and reliability (see below). In this study, phage display was chosen due to the technique's robustness and in the extended literature that is available in which this technique is used for the selection of antibody mimetics. This includes antibody fragments such as scFvs and nanobodies as well as DARPins which are discussed in **Section 1.5**, and includes a wide array of targets like cell surface receptors and toxins. Finally, phage display often results in antibody mimetics with a sub-nanomolar affinity (Steiner, Forrer et al. 2008, Bazan, Calkosinski et al. 2012). For example, DARPins have been isolated against many membrane targets, for example, a DARPin inhibitor was used in the co-crystallisation of AcrB using ribosome display. This DARPin had a low nanomolar affinity and was isolated after four selection cycles on detergent solubilised AcrB immobilised via biotinylation (Sennhauser, Amstutz et al. 2007).

1.4.1 Phage display

In principle, phage display technology aims to fuse filamentous bacteriophage (viruses that infect bacterial cells) coat proteins (usually coat protein P3 from the bacteriophage M13) with proteins of interest (in this instance, Designed Ankyrin Repeat Proteins or DARPins), so that the phage is able to display them on its surface. This thereby links the phenotype of a particular peptide with its genotype. Once display has been achieved these phage particles can be selected for their ability to bind to an immobilised protein target (Smith 1985). As can be seen in **Figure 1.4** a typical solid surface phage display selection, first requires the target protein of interest to be immobilised onto a solid surface. As is the case in this study, this immobilisation often relies on the biotinylation of target proteins and the use of streptavidin coated surfaces. In instances such as this, naïve libraries of putative binding partners displayed on the surface of phage particles are deselected against streptavidin in order to remove proteins capable of binding directly. Phage which does not bind during the deselection are removed by washing and used to select against the immobilised target protein. The fusion proteins which fail to bind to the target are removed by washing and discarded, while those that successfully bind are subsequently eluted (phage elution can be achieved via several different methods such as a pH change with trimethylamine and glycine or by trypsinization) and used to infect a culture of host bacteria (most often *E.coli*). Once infected the cells can either be used to amplify the amount of phage (and thus the fusion protein which successfully bind to the target) with the use of 'helper phage' or the DNA which encode the binding protein fused to the bacteriophage coat protein (the phagemid) can be recovered from the bacterial cells and sequenced. If the

phage is amplified then it can be used in further rounds of selection against the immobilised target protein. If the DNA is recovered, the binding protein DNA can be sub-cloned into an expression vector and subsequently expressed and purified (see **Figure 1.4**). An very similar methodology is employed for solution based selections, the difference being that target proteins are immobilised on beads (for example streptavidin coated magnetic beads) as opposed to a plate as shown in **Figure 1.4**.

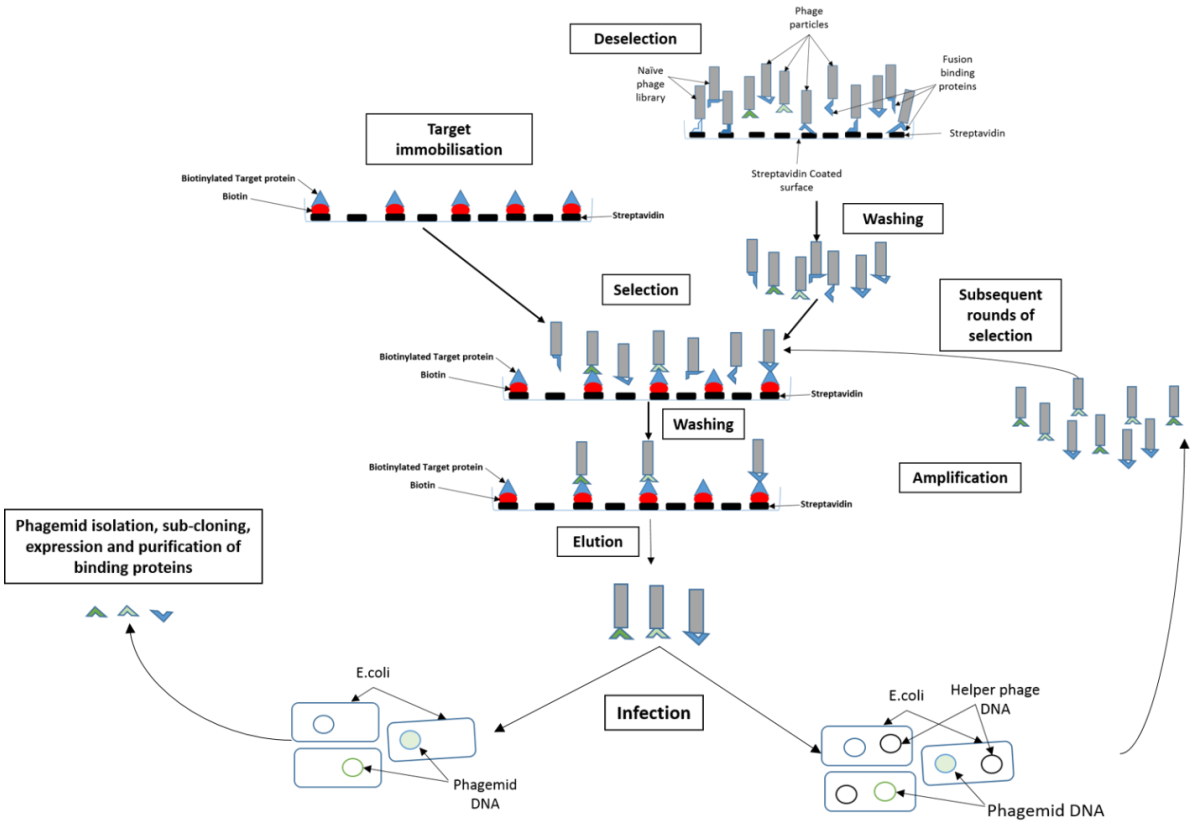


Figure 1.4: <u>The schematic of a typical phage display experiment</u>: Initially libraries of potential binding proteins are deselected against streptavidin in order to remove any binding proteins capable of binding to it directly, target proteins also undergo biotinylation and subsequent immobilisation on a streptavidin coated surface. Phage that are incapable of binding streptavidin are washed used in a selection against the immobilised target protein. Phage incapable of binding the target are then washed while the binding phage are eluted and used to infect *E.coli*. The infected cells are then superinfected with helper phage which facilitates the amplification of the eluted phage. This eluted phage can then be used in subsequent rounds of selection against the biotinylated target protein. Alternatively infected *E.coli* can be used to isolate the phagemid DNA in order to express and purify the binding proteins.

As mentioned above, the phagemid requires a 'helper phage' in order to produce full phage particles. This is because the phagemid lacks all of the genes required to produce viable phage molecules in order to avoid bacterial cell stress which can lead to unstable phage libraries. Hence these components are provided by the 'helper phage' which itself lacks the coat protein contained in the phagemid and thus facilitates the production of phage molecules and their subsequent presentation of binding protein on their surface (Baek, Suk et al. 2002, Soltes, Hust

et al. 2007). To this end, during phage display screens, bacterial hosts are transformed with both phagemid and 'helper phage'.

Premature folding of antibody mimetics has been shown to reduce the amount of protein successfully translocated to the phage surface. Therefore, their open reading frames are often preceded by the signal sequence from *E.coli* namely the signal recognition particle (SRP). The SRP ribonucleoprotein is required because (unlike the Sec pathway), upon binding nascent polypeptides containing the appropriate signal sequence, it arrests translation. In this way it prevents the premature folding of DARPins before they can be translocated to the plasma membrane (Valent, Scotti et al. 1998, Steiner, Forrer et al. 2008). The SRP sequence of DsbA is the most commonly used SRP in plasmids designed for phage display (phagemid) and has been previously shown to very effective in the co-translocation of mimetics to the periplasm (Schierle, Berkmen et al. 2003, Steiner, Forrer et al. 2008).

An amber stop codon is also often included within the phagemid, separating the binding protein from the phage coat protein. This stop codon is only recognised by bacterial amber suppressor strains which are able to read through it due to a mutation in their tRNA. Therefore its presence confers a level of control over which form of binding protein is expressed (either the protein alone or the protein bound to the phage coat protein) by transforming particular strains of expression host (mostly *E.coli*) with the phagemid (Soltes, Hust et al. 2007).

Phage display technology takes advantage of several fundamental characteristics of the filamentous bacteriophage genome, in particular, their tolerance to insertions within non-essential regions and the fact that alterations to the phage coat proteins does not affect the organism's pathogenicity toward its bacterial hosts (Russel, Lowman et al. 2004, Steiner, Forrer et al. 2008, Thie, Schirrmann et al. 2008). However, the fact that bacterial cells must be transformed during the phage display procedure lowers the diversity of the library which can be presented to immobilised targets.

1.4.2 Bacterial display

Bacterial display procedures are similar to phage display in that, peptides of interest are presented at the surface of bacterial cells by fusing them to anchoring motifs. Proteins of interest can either be inserted into the centre of the surface protein or as a terminal fusion. However, fusion of binding proteins onto surface proteins can have undesirable physiological effects on the host cells which may be detrimental to the selection process (Lee, Choi et al. 2003).

Research has now highlighted a wide range of carrier proteins which are resistant to the integration of binding proteins and have thus been used to display peptides of interest on the surface of bacterial cells. Although, Gram negative bacteria are the most extensively used, particularly *E.coli* due to its rapid growth rate and ability to allow the presentation of larger libraries (Daugherty 2007). Carrier proteins which have been used in this manner include outer membrane proteins such as OmpA and OmpX or flagellar proteins such as FliC (Getz, Schoep et al. 2012). Selection experiments using bacterial display are usually linked with techniques such as fluorescence-activated cell sorting (FACs) in order to enrich clones which bind to an immobilised target. The use of FACs also provide a real time quantitative analysis of the library as it effects a target and increase the contexts in which bacterial display is applicable, for example the proteins' stability in the presence of proteases can be monitored as can its binding affinity (Getz, Schoep et al. 2012). The size of an antibody mimetic is an important factor in the likelihood of an antibody mimetic to be successfully displayed on the surface of a bacterial host, as well as its disulphide bond content and charge (Lee, Choi et al. 2003).

Despite the fact that high affinity binders to a variety of targets (which had been previously subjected to phage display) have successfully been isolated using this method, its applicability to this study is limited. For instance, with OmpA as the carrier protein and a library of 5×10^5 small 15 residue monomer peptides, several high affinity binders were selected against five unrelated target proteins which have previously been used in phage display. These targets included human serum albumin, a monoclonal antibody (mAb) against an epitope on T7, human C-reactive protein, the HIV-1 surface protein GP120 and streptavidin (Bessette, Rice et al. 2004). However, due to the unpredictable consequences displaying full length fusion proteins such as DARPins may have on the host cell and the fact that smaller libraries of proteins give rise to lower quality binders, this selection strategy is not appropriate for this study.

1.4.3 Yeast display

Yeast display has many conceptual similarities to both bacterial and phage display in that proteins of interest are fused to the surface of yeast cells and used as binding probes during selection experiments. As with bacterial display, a wide variety of surface proteins can have binding proteins integrated within them in order to display them on the cell surface, the original being α-agglutinin, a cell wall protein in *Saccharomyces cerevisiae* involved in mating. Many members of the glycosylphosphatidylinositol (GPI) family such as of cell wall proteins can also be used for display procedures in yeast (Pepper, Cho et al. 2008). The most commonly

used carrier protein is a C-terminal fusion with the α-agglutinin protein Aga2p which is anchored to the yeast surface via a disulphide bridge with Aga1p and is also involved in mating (Pepper, Cho et al. 2008, Gera, Hussain et al. 2013). However, it has recently been shown that use of this surface protein may have many negative effects on the cells including either altering its growth kinetics or the hydrophobicity of its cell wall (Andreu and Del Olmo 2013).

S. cerevisiae cells produce Aga2p fusions through the use of shuttle plasmid vectors, while Aga1p is stably integrated within the chromosome. Both proteins are under the control of the GAL1 promotor, therefore the protein of interest is only displayed in the presence of galactose (Gera, Hussain et al. 2013).

The major benefit of yeast display is that the selection is performed in a eukaryotic system. This allows any complex post-translational modifications to be performed to the antibody mimetic. Also, like bacterial display, they are compatible with FACs analysis, although neither display technologies are able to present as large a library as phage display (Gera, Hussain et al. 2013).

1.4.4 Ribosome display

Phage display has become one of the fundamental methods for the elucidation of both antagonistic and agonistic drug candidates against a wide variety of target proteins. The robustness of the technique when compared to other screening methods makes it particularly useful. Ribosome display, however, is superior in certain key aspects, namely the ease with which binding protein libraries can be altered through the use of error prone PCR and the removal of the transformation efficiency of a bacterial host cell as a factor in the success of selections (thereby allowing the display of larger libraries). Phage display on the other hand depends upon the *in vivo* production of the binding proteins and their presentation on the surface of phage particles (Schaffitzel, Hanes et al. 1999, He and Khan 2005). The main limitations with ribosome display lie with its dependency on stable mRNA as well as the ribosome itself and reaction conditions must therefore be closely monitored with regard to the presence of RNases, denaturing conditions and temperature (Steiner, Forrer et al. 2008); issues which are exacerbated when the target under test is a membrane protein.

In principle, ribosome display relies on the *in vitro* translation of an mRNA molecule encoding an antibody mimetic of interest. In order to stabilise the ribosome/mRNA/ folded polypeptide complex, the mRNA is designed with a C-terminal spacer sequence without a stop codon. When translated, this spacer sequence remains bound to the appropriate tRNA and thus remains within the ribosome, while the correctly folded binding protein protrudes from the

ribosomal surface. These complexes can thus be used as binding probes to an immobilised target creating a direct link between the phenotype and genotype (Zahnd, Amstutz et al. 2007, Pluckthun 2012). The mRNA of binding proteins can then be easily recovered and its cDNA found via reverse transcription and PCR amplified so they can be used in subsequent rounds of screening (Pluckthun 2012).

DNA constructs for ribosome display are designed with a strong promotor upstream from either a prokaryotic ribosome binding site or a eukaryotic Kozak sequence, depending upon the subsequent expression system that is required. The DNA constructs are also designed to place a 'stemloop' at both the 3' and 5' ends of the RNA in order to increase its resistance to RNase mediated degradation (Pluckthun 2012).

1.4.5 mRNA display

mRNA display is a conceptually similar selection technology to ribosome display in so far as it takes advantage of the translation process. However, while ribosome display aims to form complexes between the ribosome, the mRNA and the fully folded polypeptide (Zahnd, Amstutz et al. 2007), mRNA display is designed to form polypeptide/mRNA fusions which are released from the ribosome and subsequently used as screening probes. This is achieved through the use of the 3' tRNA analogue puromycin which is able to prematurely terminate translation due to its structural similarity to adenosine. This similarity allows puromycin to occupy one of the binding sites within the ribosome after which it is incorporated into the polypeptide and translation is immediately terminated resulting in proteins covalently linked to the mRNA which encodes it (Keefe 2001).

In order to perform mRNA display, DNA libraries are first established with constant regions at both ends of the DNA which provide features such as the promotor region, affinity tags and ribosome binding sites. Once the DNA is transcribed, the 3' end of the mRNA is modified with an oligonucleotide which houses the puromycin binding site as well as a poly-adenosine tail. The oligonucleotide acts as a spacer, which ensures the puromycin properly fuses to the protein while the adenosine tail facilitates protein/mRNA purification (Keefe 2001, Seelig 2011).

Like ribosome display, mRNA display shares several distinct advantages over other display strategies, namely the ease with which libraries of binding proteins can be altered via error prone PCR or mutagenesis, the large libraries which can be screened and the option to include unnatural amino acids. However, due to the fact that mRNA display does not need to maintain

the ribosome protein complex, the experimental conditions required are less stringent compared to ribosome display.

Despite their benefits, neither ribosome nor mRNA display are appropriate selection strategies for this study due to their reliance on *in vitro* translation. This expression technique requires the inclusion of Mg^{2+} ions as they prevent the undesired hydrolysis of tRNA and are therefore added during both ribosome and mRNA display procedures (Keefe 2001, Pluckthun 2012). Observational data has shown that magnesium ions are detrimental to the stability of SMALPs, which are one of the solubilisation methods under test in this study.

Display Technology	Maximum Library size	Suitability	Example of its use
Phage Display	10^{10} - 10^{11}	Suitable	Affibody against EGFR (Friedman, Nordberg et al. 2007)
Bacterial display	10^{9}	Unsuitable as the library under test is a commercial phage display DARPin library	Higher affinity binders found against human C-reactive protein, the HIV-1 surface protein GP120 (Bessette, Rice et al. 2004)
Yeast	10^{7}- 10^{9}	Unsuitable as the library under test is a commercial phage display DARPin library	Affinity maturation of a scFv against the Vβ8 T cell receptor (Kieke, Cho et al. 1997)
Ribosome	>10^{12}	The presence of Mg^{2+} makes this unsuitable for SMALPs	A DARPin crystallization chaperone for AcrB (Monroe, Sennhauser et al. 2011)
mRNA	>10^{12}	The presence of Mg^{2+} makes this unsuitable for SMALPs	88 amino acid peptide apatamer against streptavidin (Wilson, Keefe et al. 2001)

Table 1.1: A comparison of the different display technologies

1.5 Antibody mimetics

Antibodies are involved in the immune response and are responsible for binding to foreign particles, highlighting them for degradation by components of the immune system such as the phagocytes. Antibodies come in several different classes, but the class that is typically used for research and therapeutics is known as Immunoglobulin G (IgG). IgGs are Y-shaped molecules consisting of 2 long heavy chains associated with 2 shorter light chains by disulphide bonds (Ruigrok, Levisson et al. 2011). These chains are made up of several units known as immunoglobulin folds composed of several anti-parallel β-sheets (Bork, Holm et al. 1994). Both the heavy and light chains are separated into constant and variable regions. The constant regions are found in all classes of antibody within a particular species and play roles in antibody stability and signal transduction. The variable regions determines antibody specificity through the complementarity determining region (CDR) (Ruigrok, Levisson et al. 2011).

With regard to the drug discovery pipeline, any antibody or antibody mimetic which has been isolated against a target of interest via a selection experiment are subject to extensive characterisation using a wide array of techniques. These techniques include ELISA and size exclusion chromatography, which aim to confirm its ability to bind to target as well as identify which epitope the antibody or antibody mimetic binds to. ELISAs with potential drug candidates also try to use a wide enough range of concentrations in order to establish an IC_{50} and thereby estimate an affinity of the binding protein for its target. This can then be subsequently confirmed using techniques such as Isothermal titration calorimetry (ITC). These binding affinities can then be optimised via site directed mutagenesis if desired, although identification of the essential residues is a time consuming process which requires the random substitution of several residues. Once binding has been confirmed and a desirable affinity has been obtained, the *in vivo* pharmacokinetics and efficacy of the candidate are investigated. In the case of bio conjugates, residues in the candidate can once again be altered via site directed mutagenesis in order to introduce residues amenable to biochemical alteration such as a cysteine residue (Simon, Frey et al. 2013). Pharmacokinetics can be explored via the use of GFP fusion proteins or the use of fluorophores and fluorescence microscopy. If a cell surface receptor has been targeted, the level of receptor at the surface of the cell can also be monitored (Boersma, Chao et al. 2011).

Antibody mimetics are small proteins which replicate the binding activity and specificity of full length antibodies, while being smaller and, typically, more stable. They also have the added benefit of relatively easy expression and purification system, with a short production time that does not require the use of animals. As such, antibody mimetics are of particular interest to

the pharmaceutical industry. Typically, they are protein constructs that utilise the immunoglobulin fold. However, in recent years a range of mimetics based upon different protein scaffolds which facilitate protein-protein interactions have arisen. These stable scaffold motifs can be used as a starting point onto which diversity can be introduced in the sections of the protein which actually bind targets. For example, in the case of DARPins, they each have the same ankyrin motif but the actual residues responsible for binding differ from one another (Steiner, Forrer et al. 2008). As previously mentioned, the displayed antibody mimetics also often have the SRP signal sequence which is recognised by the *Escherichia coli* (see **section 1.3.1**) (Steiner, Forrer et al. 2008). In this section, we will discuss a number of the most common antibody mimetics.

1.5.1 Single chain variable fragments

As previously stated, antibodies are composed of a heavy and light chain separated into modules known as immunoglobulin domains. Single chain variable fragments (scFv) are antibody mimetics made up of the variable regions of the light (V_L) and heavy chains (V_H) linked with one another by either a disulphide bridge or a flexible linker peptide which stabilizes them (Glockshuber, Malia et al. 1990, Jung, Pastan et al. 1994). The length of this linker is important; a linker above 3.5 nm ensures that the scFv will fold correctly. A linker that is shorter than this is used to produce diabodies, in which the variable regions dimerize (Holliger, Prospero et al. 1993). Stretches of glycine and serine are often employed in order to maximise the linker's flexibility. Similarly, it is important that the linker is hydrophilic in order to avoid intercalation with the variable domains. Therefore, they are often designed with a large number of lysine and glutamic acid residues thereby maximising the proteins solubility. scFvs are therefore able to maintain the specificity of fully formed antibodies while being easier to produce as they can be produced within bacteria such as *E. coli* (Ahmad, Yeap et al. 2012).

Typically, the construction of scFVs utilises the V_H and V_L produced by hybridomas, the spleen cells from immunised mice or from human B lymphocytes. Their construction involves the isolation of the V_H and V_L mRNA and its subsequent conversion to cDNA via reverse transcription. This cDNA is then used as a template for PCR amplification, thus resulting in a library of V_H and V_L genes, which can then be randomly combined via PCR assembly and expressed in, for instance, phagemids in order to perform subsequent binding analysis via phage display or other display techniques are described above (McCafferty, Griffiths et al. 1990, Marks, Hoogenboom et al. 1991). Alternatively, naïve scFvs can be formed by isolating

the variable region mRNA of non-immunised animals and subjecting the subsequent cDNA to error prone PCR in the V_H and V_L CDR regions (Silacci, Brack et al. 2005).

Several applications of scFVs have been demonstrated and of particular interest are their potential as cancer therapies. For example, cytotoxic effects could be engineering by using a scFV capable of binding the interleukin receptor CD123, a common marker of acute myeloid leukaemia. By conjugating the scFVs with a truncated form of the immunotoxin *Pseudomonas* exotoxin-A, targeted cell death was observed (Stein, Kellner et al. 2010). To this end, scFVs have also played an important role in the development of immunocytokines. These are a class of promising therapeutic agents, which fuse antibodies or antibody mimetics with immune mediators such as interleukins, resulting in the selective removal of targets by the hosts own immune system. Immunocytokines can be classified in two distinct groups; firstly the large fusion proteins in which cytokines are fused to the entire heavy chain of an antibody. Secondly, the small fusion proteins in which cytokines are fused to antibody fragments such as scFvs or diabodies. While the large immunocytokines have a longer half-life in blood, the increased concentration of cytokines can result in dangerous side effects which can be avoided by the use of the small fusion immunocytokines. A fusion of IL-15 and a human antibody scFv specific for an angiogenesis marker which is upregulated in solid tumours (EBD domain of fibronectin) showed potent anti-cancer activity when tested in mice (Kaspar, Trachsel et al. 2007, List and Neri 2013).

1.5.2 Nanobodies

Nanobodies are another class of antibody fragments which are significantly smaller than full length antibodies. They are composed of a single variable immunoglobulin domain and exhibit a significant increase in thermostability, as well as an increased resistance to denaturants compared to full length antibodies and scFvs. Nanobodies are typically derived from camelids due to their atypical immunoglobulins which do not have light chains. Instead, camelid immunoglobulins are produced as heavy chain homodimers (V_HH) and have the smallest antigen binding site of all natural antibodies (Goldman, Anderson et al. 2006, Fridy, Li et al. 2014).

The isolation and selection of nanobodies capable of binding a single target is similar to that off scFVs, in that it relies on phage display (or other display techniques). Generally, in order to produce nanobodies, camelids (typically a llama) are first immunised and the variable gene repertoire from the resultant blood lymphocytes are sub cloned into phagemids. Alternatively, if a naïve library of nanobodies is required, the variable gene regions of non-immunized

camelids can be isolated, subjected to reverse transcription and sub-cloned into phagemid vectors. In order to maximise the number of clones within the library, the genes are subjected to hypermutation (using error prone PCR within the CDR region) in order to account for the diversity which is usually introduced by exposure to an antigen in a similar fashion as is performed to produce naïve scFv libraries (Goldman, Anderson et al. 2006, Monegal, Ami et al. 2009).

Nanobodies have been demonstrated to be proficient candidates for many applications, mainly due to their high specificity and stability. Their small size means that they are more susceptible to renal clearance than either scFvs or whole antibodies, but it also increases their tissue permeability. Nanobodies against GFP conjugated with fluorophores have been described which allow single molecule localisation experiments to be carried out on GFP tagged proteins (Ries, Kaplan et al. 2012). As potential therapeutic agents, nanobodies have been demonstrated to exhibit a high neutralisation activity in the presence of bacterial toxins and snake venom. Similarly, bispecific nanobodies (two nanobodies capable of binding two different epitopes bound to one another with a peptide linker) have been engineered which provided a high degree of protection to mice treated with lethal doses of *Androctonus australis hector* scorpion venom (Aah venom) compared with the commonly used polyclonal antibody based anti-venom. Bispecific nanobodies were of particular use, as the Aah venom is composed of three different small basic toxins the pharmacokinetics of which have shown that they can rapidly spread throughout the body compared to its anti-venom. Thus bispecific nanobodies capable of targeting the two most toxic elements the venom were desirable. Previous attempts to neutralise this venom used scFvs derived from mouse hybridomas, however their therapeutic use opens the possibility of anti-mouse antibodies limiting their efficacy. Nanobodies are able to avoid this shortcoming due to the similarity between camelid and human antibody variable regions and their small size suggests that their pharmacokinetics should be similar to the venom (Hmila, Saerens et al. 2010). They have also shown some promise in the treatment of inflammation and due to their rapid clearance from blood they have some potential as *in vivo* imaging tools (Muyldermans 2013).

1.5.3 Designed ankyrin repeat proteins and antibody fragments

Designed Ankyrin Repeat Proteins (DARPins) are small binding proteins which benefit from lower production costs (as they can be produced in bacteria) and increased robustness (they can be used in chemically harsh environments) when compared to full length antibodies. Their structure is based upon the widespread ankyrin repeat, a 30-34 residue protein motif arranged

in two antiparallel alpha helices, separated by an outward projecting loop. It facilitates many protein-protein interactions (Li, Mahajan et al. 2006) and occurs as several tandem repeats which bundle with one another in a similar fashion as a β-hairpin sheet within proteins. Typically, proteins contain 4 – 6 ankyrin repeats which form a grooved, solenoid structure with a hydrophobic core (Pluckthun 2015).

One of the main benefits of DARPins is their high degree of stability which is a result of several different factors. The most important being the coupling between the internal repeats, which arises due to a combination of electrostatic interactions and hydrophobic shielding (Tamaskovic, Simon et al. 2012). DARPins can contain any number of ankyrin repeats, it has been shown that the more repeats present, the more stable the DARPin is. Once a DARPin has a minimum of three internal repeats, there is a marked increase in its thermal stability as well as its resistance to denaturants. These internal repeats are flanked by an essential C- and N-terminal cap which ensure proper folding and prevent aggregation of the DARPin in bacterial cells (see **Figure 1.5**). The C-terminal cap also plays a role in DARPins high degree of stability due to its close packing with the internal repeats, which is a consequence of their sequence similarity. This stability and robustness allow them to withstand conditions which would be considered too harsh for antibodies - such as in the presence of a reducing agent (antibodies are held together by disulphide bonds) or a detergent. Their complete denaturation requires incubation at high temperature in 5 M guanidine hydrochloride. The diversity of DARPin libraries comes from the randomization of residues located within a single repeat. Depending on the number of repeats within a DARPin it may become grooved (like ankyrin bundles) therefore complementing folded proteins and increasing the likelihood of correct binding (Gronwall and Stahl 2009, Milovnik, Ferrari et al. 2009, Boersma and Pluckthun 2011, Tamaskovic, Simon et al. 2012). The DARPin library under test is composed of three internal repeats capped with an N and C-terminal cap and 21 randomised residues in the loop regions (see **Figure 1.5)**.

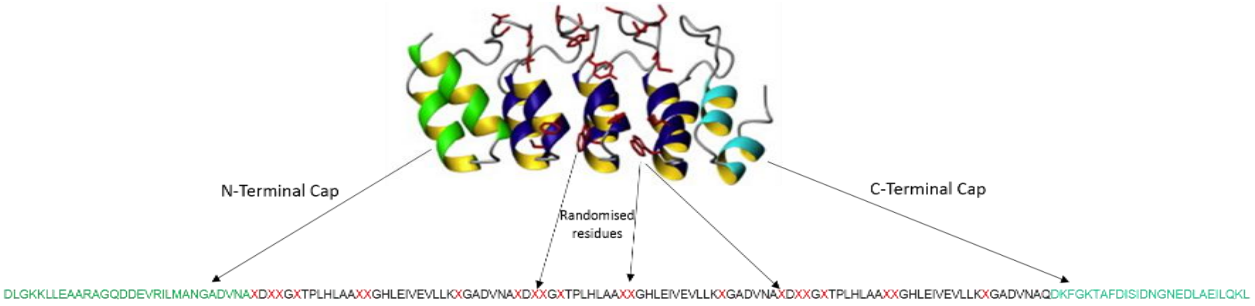


Figure 1.5: <u>The 3D structure of a DARPin</u>: The model depicts three internal ankyrin repeat units flanked by an N-terminal cap in green and a C-terminal cap in cyan. Randomised residue side chain are highlighted in red. The 3D structure has been placed over a generalised sequence of the DARPins used in this study which has been colour coded according to the structure. Each of the DARPins have three internal repeats and 21 randomised residues. Picture taken from the Chapter five "Designed Ankyrin Repeat Proteins (DARPins): From Research to Therapy" of the book "Methods in Enzymology" (Tamaskovic, Simon et al. 2012)

DARPins have shown promise in several different aspects of therapeutics in recent years. For example, a DARPin with pico-molar affinity to the human epidermal growth receptor 2 (HER2) was able to detect elevated levels of the receptor in complex tissue with a specificity that surpassed a pre-existing therapeutic antibody, while showing a very similar level of sensitivity (Theurillat, Dreier et al. 2010). Thus highlighting its potential as either a powerful diagnostic tool or a therapeutic in patients with breast cancer, as elevated levels of HER2 occurs in 20 – 30% of cases. The specificity of DARPins also made them suitable components as biological recognition elements in biosensors. For instance, the DARPins were shown to be capable of differentiating between non-phosphorylated and phosphorylated extracellular signal-regulated kinase (ERK), thus allowing the protein activation to be monitored in mouse embryo fibroblasts when the DARPins were conjugated with a thiol-reactive merocyanine dye (Kummer, Hsu et al. 2013, Pluckthun 2015).

1.5.4 Adhirons

Adhirons are a novel class of antibody mimetic based upon a scaffold derived from a cysteine protease inhibitor in plants called phytocystatin. This inhibitor is a member of the cystatin family, all of which are identified by a highly conserved fold, comprised of a central α-helix with four anti-parallel β-sheets wrapped around it (see **Figure 1.6**). In order to construct adhirons, the N-terminal of the scaffold is truncated and its two inhibitory loops (one of which contains a QXVXG active site) are altered with a random sequence of nine residues. This insertion gives rise to two variable regions in the loops located between the β-sheets. Like DARPins, adhirons exhibit a vastly superior degree of thermostability compared to full length

antibodies with melting temperatures of 101°C and high level of expression in *E. coli* (Tiede, Tang et al. 2014).

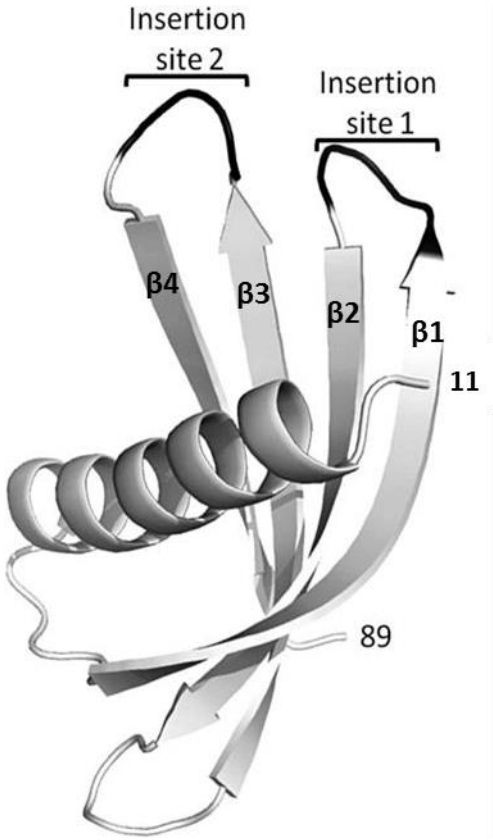


Figure 1.6: <u>The X-ray crystal structure of an adhiron:</u> The structure is at a resolution of 1.75 Å and depicts a truncated adhiron, from residue 11 – 89. The positions of the randomised amino acid residue insertions are marked in black. Original image was taken from the paper "Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications" (Tiede, Tang et al. 2014).

Phytocystatin is a very important protein in plants and has been shown to play roles in plant defence against proteases released by pathogens as well as environmental stress such as drought. It also plays a role in the regulation of proteases during seed maturation and is involved in programmed cell death (Chan, Abu Bakar et al. 2014). Its small size, stability and natural lack of either cysteines or glycosylation sites, make it ideal for the development of a protein library capable of binding targets of interest.

Adhirons have been shown to bind several different targets specifically, such as a chemically biotinylated yeast small ubiquitin-like modifier (SUMO) protein with low nanomolar affinity. They have also been shown to bind magnetite nanoparticles with high affinity in the presence of casein (Rawlings, Bramble et al. 2015).

1.6 Project Aim

As this chapter has discussed, detergent solubilisation is the most common method of membrane protein solubilisation, however the effect that detergent micelles have on *in vitro* selection experiments with antibody mimetics is poorly understood. Therefore this study, aims to compare the potential of different methods of membrane protein immobilisation and solubilisation with regard to their ability to select an antibody mimetics using a panel of membrane proteins. The methods of membrane protein immobilisation under test were (a) the cloning of a streptavidin binding protein (SBP) tag into the C-terminal of a model membrane protein, for binding to a streptavidin coated surface (a staple of the pharmaceutical industry); (b) the cloning of a C-terminal avitag which can undergo enzyme-catalysed biotinylation of its lysine residue and subsequently bind to streptavidin via this biotin moiety; (c) the introduction of an accessible cysteine residue by mutagenesis into the C-terminus of the model membrane protein, enabling chemical biotinylation with the reagent biotin maleimide; and (d) aspecific biotinylation of a model membrane protein by the biotinylation reagent NHS-biotin.

The methods of solubilisation under test included typical detergent solubilisation with DDM compared with detergent free methods of solubilisation such as (a) the encapsulation of purified membrane protein in nanodiscs; and (b) the use of the co-polymer SMA on bacterial membranes in order to purify membrane proteins in styrene-maleic acid lipid particles (SMALPs).

The panel of membrane proteins discussed in **section 1.3** were used as models in this study and were subjected to two rounds of phage display on a streptavidin coated surface, using a naïve library of DARPins. DARPins which successfully bound to the targets were further characterised via phage ELISA with the hope that they may be used in future co-crystallisation trails. In a typical phage display experiment a minimum of 3 rounds would be used and the cross-reactivity of the proteins which successfully bound the target would be tested, however this was not possible due to time constraints.

Ultimately this study aimed to determine which method of membrane protein immobilisation and solubilisation is the best in regard to their ability to isolate DARPin binders using phage display.

Chapter 2

Materials and Methods

2.1 Materials and Suppliers

All solutions (including growth media) where made using ultrapure water, which had been purified and deionized with a resistivity of 18.2 MΩ cm in a Milli-Q Plus system manufactured by Millipore Corporation. Where appropriate, buffers were sterilised either by autoclaving at 121°C or by filtering through a 0.22µM filter. Suppliers from which each chemical was purchased have been outlined in **Appendix 1**

2.2 Bacterial strains, Growth and Expression Media

2.2.1 Bacterial strain genotypes

E.coli strains served several functions including plasmid amplification and manipulation, protein expression and phage amplification. The genotypes of each strain used are listed below:

Omnimax™2 - *F^- [proAB$^+$ lacI^q lacZΔM15 Tn10(TetR) Δ(ccdAB)] mcrA Δ(mrr-hsdRMS-mcrBC) φ80(lacZ)ΔM15, Δ(lacZYA-argF) U169, endA1, recA1, supE44 thi-1, gyrA96, relA1, tonA, panD*

BL21 star™ - *F^- ompT, $hsdS_B$ (rB-, mB-) gal dcm rne131 (DE3)*

BL21™ - *F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3)*

AVB101 (Avanti) **-**

TG1 - F' [*traD36 proAB$^+$ lacI^q lacZΔM15*]*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, ($r_K^- m_K^-$)*

2.2.2 Growth media

Bacteria were routinely grown in either LB (1% Tryptone, 1% NaCl, 0.5% w/v yeast extract) or2xTY (1.6% Tryptone, 0.5% NaCl, 1% w/v yeast extract) media. They were adjusted to pH7 with NaOH prior to their sterilisation by autoclaving. LB agar was prepared by adding 1.5 g agar powder per 100 mL LB media prior to sterilisation.

LB agar plates were then prepared by melting pre-prepared stock LB agar in a microwave. 20 mL of the molten agar was then poured into a falcon tube and (once it had sufficiently cooled) 1/1000th of the required antibiotic was added (i.e. 20 µL). The agar was then poured into a petri dish and allowed to set for 30 mins.

2.2.3 Buffers and Media for expression via auto-induction

Expression of recombinant proteins via auto-induction was typically carried out in SB media (3.2 % Tryptone, 2% w/v yeast extract) supplemented with 1x auto-induction phosphate buffer (25 mM Na_2HPO_4, 25 mM, KH_2PO_4, 5 mM Na_2SO_4, 50 mM NH_4Cl) adjusted to pH 6.7, 1 mM $MgSO_4$, 1x 5052 (0.01% glycerol, 0.001% glucose, 0.04% α-lactose). Each constituent was sterilised by autoclave with the exception of the 5052 which was sterilised by filtration through a 0.22 μm filter. In order to avoid precipitation the $MgSO_4$ was added before the phosphate buffer. Once complete the media was immediately inoculated with bacterial culture and antibiotic was added at $1/1000^{th}$ the total volume of media.

2.2.4 Media for expression via Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction

Expression of protein via IPTG used autoclaved 2YT media (1.6% Tryptone, 1% yeast 0.5% w/v NaCl) with antibiotic added. IPTG was added to a final concentration of 0.6 mM after *E.coli* pre-cultures had reached the desired OD_{600nm}.

2.3 Recombinant DNA techniques

2.3.1 Amplification of plasmid DNA via the Polymerase Chain Reaction (PCR)

PCR was performed using KOD hot start polymerase from Novagen. This amplification usually had a final volume of 50μL and uses a maximum of 50ng of template DNA. A typical reaction was set up as follows:

5 μL 10 x KOD hot start buffer (Final concentration of 1x)

1 μL Forward Primer (100 μM) (Final concentration 2 μM)

1 μL Reverse Primer (100 μM) (Final Concentration 2 μM)

3 μL 25 mM $MgSO_4$ (Final concentration 1.5 mM)

5 μL 2 mM dNTPs (Final concentration 0.2 mM)

1 μL Template DNA (~ 50 ng/μL) (Final quantity 50 ng)

1 μL KOD hot start polymerase, 1U/μL (Final concentration 0.02 U/μL)

33 µL ultrapure, deionized water

Once mixed, the PCR samples were loaded into a DNA engine Dyad™ peltier thermal cycler from MJ research and subjected to the following thermal cycles:

94 ˚C, 3 minutes

94 ˚C, 30 seconds
50 ˚C, 30 seconds
72 ˚C, 105 seconds
} 32 cycles

72 ˚C, 5 minutes

10 ˚C, hold

In order to assess the success of the PCR, DNA samples were then analysed electrophoretically on an agarose gel (**see section 2.3.4**).

2.3.3 Restriction digestion of plasmid DNA

Typically a restriction digestion included 5 – 10 µg of DNA mixed with 2 µL of the appropriate 10x restriction digest buffer (NEB 1 – 4), 1 µL of the desired restriction enzyme (purchased from New England Biolabs) and ultrapure deionized water to a final volume of 20 µL. The restriction digestion was then incubated at 37°C for a minimum of 3h.

In the case of PCR products quick digest enzymes (from Thermo scientific) were utilized. During these restriction digests the final volume of the solution was typically 100 µL and involved mixing the entire 50 µL DNA fragment, recovered using the Promega Wizard SV® Gel and PCR clean up system (**see section 2.3.9**), with 10 µL of 10x green fast digest buffer, 2 µL of the desired fast digest enzyme and ultrapure deionized water up to 100 µL. The fast digest reactions were then incubated for a minimum of 1h at 37°C. The resultant DNA fragments were then analysed via electrophoresis on an agarose gel (**see section 2.3.4**).

2.3.4 Separation of DNA fragments via electrophoresis on an agarose gel

All agarose gels were performed in the Embi Tec RunOne™ electrophoresis cell using tris acetic acid EDTA (TAE) buffer diluted to 1x from a 50x stock, a litre of which was made with:

- 242g Tris
- 2.92g EDTA
- 57.1 mL acetic acid
- Up to 1L ultrapure H_2O

In order to make the agarose gels, 1 – 2% (w/v) agarose was added to 30 mL TAE buffer in a 250 mL conical flask. Due to the insoluble nature of agarose at room temperature, the mixture needed to be heated in a microwave for approximately 1.5 minutes. Once all agarose had dissolved the flask was cooled under cold water before 2 µL SYBR® safe DNA gel stain was added and mixed by swirling. The agarose mixture was then poured into a gel cassette with an attached comb and allowed to set for 20 minutes.

DNA samples were mixed with 6x MassRuler™ DNA loading dye and loaded into the gel alongside 5 µL Fermentas MassRuler™ DNA ladder mix and run at 100V for approximately 35 minutes. The gel was then quickly visualised using the Invitrogen Safe Imager™ and the size of DNA confirmed by comparison with the ladder. The quantity of DNA was confirmed by slicing the desired DNA fragment from the gel, purifying it using the Promega Wizard SV® Gel and PCR clean up system (**see section 2.3.9**) and spectrophotometrically analysing the fragment with a microspectrophotometer.

2.3.5 Dephosphorylation of DNA fragments

In order to prevent DNA from self-ligating, phosphate groups were removed from their 5' end. This process required the use of the Antarctic phosphatase enzyme from New England Biolabs. Typically a dephosphorylation reaction consisted of adding 6 µL 10x Antarctic phosphatase buffer (diluted down to 1 x) was added to the entire 50 µL DNA sample followed by 3 µL of dH_2O and 1 µL (5 units) Antarctic phosphatase enzyme prior to incubation for 1 hour at 37°C. The phosphatase enzyme was then heat inactivated by incubation at 65°C for 5 mins. The newly dephosphorylated DNA could then be used in subsequent ligations (**see section 2.3.6**).

2.3.6 Ligation of DNA fragments

A typical ligation reaction had a maximum of 10 ng of restriction vector DNA (the concentration of which was determined via micro-spectrophotometer either the ND-1000 Nanodrop from Thermo Scientific or the DS-11 from DeNovix) mixed with enough restriction insert DNA (dephosphorylated if compatible restriction enzymes were used) to have a molar ratio of 1:4, 1:10 or 1:15 (vector: insert). Ligations were typically performed in a total volume

of 10 – 20 µL in the presence of 1x T4 DNA ligase buffer and 1 µL T4 DNA ligase (Invitrogen 1 U/µL). The ligations were then incubated for a minimum of 16 hours at 16°C.

In the case of quick ligase (New England Biolabs), the ligation reaction typically had a final volume of 20µL and the vector and insert were mixed with 10 µl 2x quick ligation buffer followed by 1 µL quick T4 ligase. The ligation was then incubated at 20°C for 30 minutes. Once the ligation was complete the resultant DNA was used to transform Omnimax cells (**see section 2.3.8**).

2.3.7 Preparation of competent bacterial cells

The production of competent cells used two buffers designated transformation buffer 1 (Tfb1- 3 mM, potassium acetate, 100 mM $RbCl_2$, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 15% glycerol) and transformation buffer 2 (Tfb2 – 10 mM MOPs, 75 mM $RbCl_2$, 10 mM $CaCl_2$, 15% glycerol. The pH of Tbf1 was adjusted to 5.8 using 0.2 M acetic acid prior to filter sterilisation using a 0.2 µm filter. Tbf2 was adjusted to 6.5 with KOH prior to filter sterilisation with a 0.2 µm filter. Tbf1 was stored in 20 mL aliquots while Tbf2 was stored in 5 ml aliquots, both buffers were stored at 4°C

Bacterial host cells were made chemically competent with a rubidium chloride based procedure. Firstly glycerol stocks of the required cells were streaked aseptically onto a LB-agar plate (see **section 2.2.2**) and grown overnight at 37°C in a stationary incubator. A single colony from the plate was then used to inoculate 5 mL of LB media (see **Table 2.1**) and grown overnight at 37°C in an orbital shaker at 200 rpm (1 g). Depending on the strain used the agar plate and media were sometimes supplemented with an appropriate antibiotic. 100 µL of the pre-culture was then used to inoculate 50 mL of pre-warmed LB media in a 250 mL conical flask (with antibiotic if required) and incubated at 37°C with shaking at 200 rpm until an OD_{600nm} of 0.4 – 0.6 was attained. The culture was then transferred to a 50 mL falcon tube and left on ice for 5 minutes, followed by centrifugation for 10 minutes at 3000 rpm and the removal of the supernatant. The cell pellet was gently resuspended in 20 mL of ice cold Tbf1 buffer (see **Table 2.8**) so as not to damage the cells a 5 mL pipette was used instead of vortexing. Once resuspended the cells were stored on ice for 5 minutes, before being once again subjected to centrifugation for 10 minutes at 3000 rpm. Once the supernatant was discarded for the second time the cells were once again resuspended gently in 2 mL Tbf2 buffer (see **Table 2.9**) and stored on ice for 15 minutes. 50 µL aliquots of the cells were then transformed to pre-cooled Eppendorf tubes (which had been sterilised by auto-clave) and snap frozen in liquid nitrogen and stored at -80°C.

2.3.8 Transformation of competent bacterial host cells

In order to transform bacterial host cells, 1 µL of DNA was mixed with 10 µL of competent cells freshly thawed on ice. The cells were then incubated for 30 minutes at 4°C before being subjected to heat shock at 42°C for 30 seconds and returned to 4°C for an additional 2 minutes. 100 µL of LB media was then added to the cells before incubating them for 1 hour at 37°C.

Once the 1 hour incubation was complete the entire bacterial culture was spread on an LB agar plate supplemented with the required antibiotic and incubated overnight at 37°C.

Resultant colonies were then either picked and grown in 5 mL LB media for plasmid purification (**see section 2.3.8**); or grown in 50 mL pre-cultures for protein expression (**see section 2.4.1**) overnight at 37°C with shaking (200rpm).

2.3.9 Purification of plasmid DNA

The desired plasmids were purified from the 5 mL cultures described above using the Promega Wizard® SV miniprep kit. The concentration of the isolated plasmids was determined via micro-spectrophotometer prior to characterisation by either restriction analysis or sequencing.

2.3.10 Site directed mutagenesis of plasmid DNA

Site directed mutagenesis was performed using the QUICKCHANGE® procedure from Strategene. A typical reaction had a final volume of 100 µL and was set up as follows:

- 10 µL of 10× KOD buffer – Final concentration 1x
- 50-100 ng of template DNA
- 3 µL Forward primer (10µM)- Final concentration 0.3 µM
- 3 µL Reverse primer (10µM) – Final concentration 0.3 µM
- 10 µL of dNTP mix (2mM) – Final concentration 0.2 mM
- 6 µL $MgSO_4$ (25 mM) – Final concentration 1.5 mM
- Ultrapure water up to 99 µL
- 1µL of *KOD* polymerase (1 U/µL) – Final Concentration 0.02 U/µL

The reaction was separated into 5x 20 µL aliquots in 0.2 mL PCR tubes, placed into the DNA engine Dyad™ peltier thermal cycler set to provide a temperature gradient between 50 - 70°C within a single block. In order to find the optimal temperature for the site-directed mutagenesis a temperature screen was established by incubating one of the five 20 µL

mutagenesis aliquots at a single 5°C incremental temperature within the aforementioned gradient:

95 °C 4 min	
94 °C 30 s	30 cycles
Gradient from 50 → 70 °C 30 s	
72 °C 6 min	
72 °C 10 min	
4 °C continuous	

In order to remove the template DNA a DpnI restriction digestion was used as it only the enzyme is only able to digest methylated DNA. Therefore 5 µL of each reaction was analysed electrophoretically while the remaining 15 µL was mixed with 2µL 10x NEB restriction buffer 4, 2 µL ultrapure water and 1 µL DpnI from New England Biolabs once the cycles were complete. The restriction digest was then incubated for a minimum of 3 hours at 37°C and used to transform competent 'Omnimax™2' cells selected by the appropriate antibiotics. The mutagenized plasmids were then purified as discussed in **section 2.3.8**.

2.3.11 DNA sequencing

Sequencing was performed by the Source Bioscience DNA sequencing service. The reaction required a minimum of 15 µL of 100 ng/µL template DNA and 15 µL of 3.2 pmol/µL sequencing primer. Typically one of the following primers were used:

- pTACF: TTGACAATTAATCATCGGC
- pTACR: CGCCAGCTGGCGAAAGGGG
- pC6F: CAGCTATGACCATGATTACG
- pC6R: GAATTTTCTGTATGAGTTTTG
- T7Fwd: TAATACGACTCACTATAGGG

2.4 General protein techniques

2.4.1 Expression of proteins in a bacterial host strain via auto-induction

To perform an auto-induction the 50 mL bacterial cultures described in **section 2.3.8** were used as a pre-culture for large scale protein expression. 3 - 4L of SB-5052 was prepared (**see table 2.6**) and separated into 500 mL aliquots in 2 L baffled conical flasks. These aliquots were then inoculated with enough pre-culture to have a starting OD_{600nm} of 0.1. To determine the volume of pre-culture a 1/10 dilution was made in LB media and the OD_{600nm} measured spectrophotometrically. The inoculated 500 mL culture was then incubated for a minimum of 20 hours at 37°C with shaking (200 rpm / 1 g). The cells were harvested via centrifugation at 2634 g for 30 minutes and the pelleted cell paste was transferred to falcon tubes. Before storage at -80°C their weight was determined.

2.4.2 Expression of proteins in a bacterial host strain via IPTG

IPTG inductions were performed in a similar manner as described in **section 2.4.1**, except that after inoculation of either LB or 2x TY (typically 3 – 4L in 500 mL aliquots, the OD_{600nm} was continually monitored until it reached between 0.6 and 0.8. At this point IPTG was added to a final concentration of 1 mM and the culture was incubated at 37°C with shaking (200 rpm/ 1 g) for a minimum of 4 hours. The cells were then pelleted, weighed and stored as described in **section 2.4.1**.

2.4.3 Disruption of bacterial cells and isolation of membrane extract

In order to disrupt bacterial cells harvested from induction experiments, the pellets were first resuspended in a volume of buffer (20 mM Tris, 0.5 mM EDTA, pH 7.4) five to six times the weight of the harvested cells. Once resuspended the cells were homogenized firstly with a hand held homogenizer, followed by the use of the Ultra-Turrax®. Once the cell solution was completely homogenous it was subjected to two runs through a TS5/40/AB/GA cell disrupter (from Constant Systems) at 30 Kpsi in order to break the bacterial cells. Once complete the disrupter and Ultra-Turrax® were cleaned with Teepol™.

The lysed cells were then centrifuged at 14,000 g_{av} for 45 minutes in order to remove unbroken cells and cellular debris. The supernatant was then ultracentrifuged at 131,000 g_{av} for 2 hours to isolate the bacterial membranes. Trace amounts of EDTA were removed from the membrane pellets by resuspending them in a 20 mM Tris buffer (pH 7.4) prior to ultracentrifugation at 131,000 g_{av} this wash was repeated twice. The membrane pellet was

then finally resuspended in a minimum of 2 mL 20 mM Tris (pH 7.4) and snap frozen in liquid nitrogen as beads. The beads were produced by syringing droplets of resuspended membrane directly into liquid nitrogen. The concentration of the protein in the membrane extracts was determined with a BCA assay (**section 2.4.9.1**) and typically would result with concentrations between 20 – 50 mg/mL.

2.4.4 Chemical Biotinylation of proteins

The following biotinylation reagents were used:

- EZ-Link® NHS-Biotin (Thermo scientific)
- EZ-Link® NHS-SS-Biotin (Thermo scientific)
- EZ-Link® Maleimide-PEG_{11}-Biotin (Thermo scientific)

The procedure for protein biotinylation followed the manufacturer's instructions and was identical for all the biotinylation reagents used. Once a molar excess of biotin reagent to protein had been selected (typically 2.5 – 5 times), the amount of reagent was calculated following the equations set out in the manufacturers instruction.

The biotinylation reagents were equilibrated to room temperature before a 10 mM stock solution (250 mM for biotin maleimide) was prepared in an appropriate volume in dimethyl sulfoxide (DMSO) immediately before use. For biotinylation with the maleimide reagent, proteins were first incubated for 30 minutes in 5 mM tris(2-carboxyethyl)phosphine (TCEP) at room temperature in order to reduce any cysteine residues. TCEP was then removed either by desalting column or dialysis (**section 2.5.1.1**).

An appropriate volume of biotinylation reagent (based on the earlier calculations) was then added to the protein and the protein solution was incubated for a minimum of 2 hours on ice (in the case of NHS-biotin and NHS-SS-Biotin more stable proteins such as the DARPins were incubated at room temperature for 30 mins). Excess or unreacted biotinylation reagent was then removed either by desalting column or by dialysis (**section 2.5.1.1**).

The labelling efficiency of biotin maleimide was tested using 4'-hydroxyazobenzene-2-carboxylic acid (HABA). Due to HABA's natural affinity for avidin, it can be used as a reporter molecule for biotin labelling efficiency. The presence of biotin will displace the HABA causing its adsorption at 500nm to decrease proportionally. To this end, a 10 mM stock solution of HABA was prepared by dissolving 24.2 mg in ultrapure water followed by 100 μL NaOH. This solution was then filtered through a 0.22 μm filter and stored at 4°C. In order to prepare the

HABA avidin solution 10 mg of avidin was added to 600 μL of the aforementioned HABA stock and 19.4 mL of PBS was added prior to storage at 4°C.

In order to test the labelling efficiency, the A_{500nm} of 900 μL of HABA/avidin was measured in a spectrophotometer and recorded. 100 μL of biotinylated protein was then added and mixed well before the A_{500nm} was measured once every 15 seconds until it remained constant (samples with A_{500nm} over 0.3 were diluted). The two absorbance readings were then used to calculate the number of biotin molecules per protein, in a series of calculations based upon the Beer Lambert law: A = ε x bC (where A is the absorbance, ε is the molar extinction coefficient – 67443 $cm^{-1}M^{-1}$ for MPSIL0294, b is the cell path length and C is the concentration in M) (https://tools.lifetechnologies.com/content/sfs/manuals/MAN0011200_HABA_UG.pdf).

2.4.5 Separation of proteins electrophoretically via sodium dodecyl sulphate polyacrylamide electrophoresis (SDS-PAGE)

SDS-PAGEs were run using the Bio-rad protean 2 system; typically with 12% (w/v) SDS-PAGE separating gels made with 3.3 mL ultrapure water mixed with 4 mL acrylamide/Bis-acrylamide (30% w/v solution), 2.5 mL separating buffer (1.4M Tris-HCl, pH 8.8), 100 μL 10% (w/v) SDS, 100 μL 10% (w/v) freshly prepared ammonium persulphate and 5 μL tramethylethylenediamine (TEMED). Once mixed, the separating gel was then left for 30 minutes at room temperature to polymerise within the gel apparatus. The stacking gel, composed of 3.025 mL ultrapure water, 625 μL acrylamide/bis-acrylamide (30% w/v), 2.5 mL stacking buffer (0.5M Tris-HCl, pH 6.8), 100 μL 10% (w/v) SDS, 100 μL 10% (w/v) ammonium persulphate and 5 μL TEMED was then made and overlaid atop the separating gel and left for 30 minutes to set with an appropriate comb.

SDS-PAGE protein samples were prepared by adding loading buffer (**0.**5M Tris-HCl, 10% w/v SDS, 10 mM EDTA, 5.4M glycerol, 10 mg/mL pryonin Y). The pH of the loading buffer was adjusted to 6.8 using 1M NaOH and stored at -20°C in 1 ml aliquots. For a 20 μL sample, 4 μL of loading buffer was added and 1 μL of DTT, for the SDS-PAGE of membrane proteins the sample were left at room temperature for a minimum of 20 minutes before loading. For globular proteins, the samples were incubated at 95°C for 15 minutes using the DNA engine Dyad™ peltier thermal cycler. The gels were then submerged in either 1x running buffer (19 mM glycine, 2.5 mM Tris, 0.01% w/v SDS) **or** 1x MOPs running buffer (2.5 mM 3-(*N*-morpholino) propanesulfonic acid, 2.5 mM Tris, 0.005% SDS, 50 μM EDTA, pH 7.7).) prior to sample loading

and run at a constant voltage of 90 V for 100 minutes until the dye front reached the bottom of the gel.

Protein bands were visualised by Coomassie staining using four solutions the first of which was a fixing solution (10% acetic acid and 25% isopropanol) which aimed to denature and entrap proteins within the gel matrix. Gels were incubated in this solution for a minimum of 4 hours before they were transferred to staining solution 1 (10% acetic acid, 25% isopropanol and 0.025% Coomassie). Once again the gels were incubated for a minimum of 3 hours before transfer to staining solution 2 (10% acetic acid, 10% isopropanol and 0.0025% Coomassie) for a minimum of 2 hours. Finally the gels were transferred to destaining buffer (10% acetic acid). This destain buffer was continuously replaced until all extraneous dye was removed.

2.4.6 Electrophoretic transfer and detection of proteins on an nitrocellulose membrane via western blotting

Six pieces of filter paper and a piece of nitrocellulose membrane were incubated in transfer buffer (25 mM Tris, 192 mM glycine, 20% v/v methanol) for a minimum of 1 hour. Once the SDS-PAGE gel had finished it was also incubated for a minimum of 30 minutes in the same buffer. Electrotransfer of protein samples was achieved using the TE77x semi-dry transfer unit (Hoefer Inc®) at a constant amplitude of 43 A for a minimum of 1 hour. The membrane was then blocked in 10 mL 3% BSA dissolved in TBST (50 mM Tris, 150 mM NaCl, 9, 2% w/v Tween®20) buffer for a minimum of 1 hour. After blocking the membranes, they were treated with the required antibody diluted to an appropriate dilution (as outlined in the manufacturer's instructions) in the same buffer used to block the membrane for 1 hour. After three 15 minute washes in 10 mL TBST buffer, the membrane was incubated in a freshly prepared equal volume mixture of SuperSignal® West Pico (Pierce) Luminol/Enhancer Solution and Stable Peroxide Solution for precisely 1 minute. The chemiluminescent bands of the blot were then visualized using a G-box from SynGene. Quantification by densitometry was performed using ImageJ.

2.4.7 Protein quantification

2.4.7.1 BCA assay

The concentration of membranes and purified protein was determined via the bicinchoninic acid (BCA) assay with a range of 0 – 10 μg of BSA used to determine the standard curve. All samples (and standards) had a final volume of 10 μL and were done in duplicate in a 96-well plate. Once all samples had been loaded, BCA assay reagent A (Pierce) was mixed with 4% $CuSO_4$ in a ratio of 50:1 and 200 μL was added to each well of the plate. In order to facilitate the BCA/peptide bond reaction the plate was incubated at 37°C for a minimum of 30 minutes before its absorbance at 570 nm was analysed using a plate reader (ICN Biomedical Titertek).

2.4.7.2 Spectrophotometric measurement

Alternatively protein samples could be quantified by analysing it's absorbance at 280 nm. This method relies on the ability of aromatic amino acids to absorb light at this wavelength and the protein concentration can be determined using the Beer-Lambert law (**see 2.4.6**). This method of quantification was used for several membrane proteins, namely AcrB (an extinction coefficient of 89855 $cm^{-1}M^{-1}$ was used), NupC (32890 $cm^{-1}M^{-1}$) and VcCNT (39545 $cm^{-1}M^{-1}$). DARPins (4470 $cm^{-1}M^{-1}$) were also quantified in this manner. Protein concentration was determined in this manner using either the Biowave II spectrophotometer (Biochrome WPA), the ND-1000 Nanodrop™ micro-spectrophotometer (Thermo scientific) or the DS-11+ spectrophotometer (DeNovix)

2.5 Purification of proteins via Immobilized metal ion affinity chromatography

2.5.1 Purification of MPSIL0294 variants

2.5.1.1 Purification of MPSIL0294-SBP

MPSIL0294-SBP was expressed using auto-induction with SB-5052 media (**see Table 2.6**) in BL21 star (DE3); the resultant membranes were isolated using a cell disruptor as discussed in **section 2.4.3.** MPSIL0294-SBPs subsequent purification typically involved taking less than 5 mL of membrane and solubilising them with 66.7 mL of solubilisation buffer (1.5% n-Dodecyl-β-D-maltoside (DDM), 50 mM HEPES, 20% glycerol, 150 mM NaCl, 10 mM Imidazole and protease inhibitor cocktail (EDTA free), pH 8), resulting in a final protein concentration of approximately 3 mg/mL. The membranes were then incubated at 4°C for 1 hour after a 100 μL sample was taken for later analysis. The solubilised membranes were then ultracentrifuged at 110,000g_{av}

for 1 hour, after which the insoluble pellet was resuspended in 1% SDS and a sample was taken for SDS-PAGE analysis. After a 100 µL samples was taken for further analysis, the supernatant was added to a BV of 1 mL of Hispur™ cobalt resin pre-equilibrated in wash buffer 1 (50 mM HEPES, 20% glycerol, 150 mM NaCl, 10 mM imidazole and 0.05% DDM, pH 8). The MPSIL0294 was bound to the resin by incubating at 4°C for 2 hours on a roller. The resin was packed into a disposable filtered column (Thermo-Pierce) by allowing the buffer to flow through it. A sample of this buffer was then taken for later analysis. The packed resin was then washed twice, initially with 10BV (i.e. 10 mL) wash buffer 1, followed by 7.5 BV wash buffer 2 (50 mM HEPES, 20% glycerol, 150 mM NaCl, 40 mM imidazole and 0.05% DDM pH 8); samples of each wash were taken for subsequent SDS-PAGE analysis in order to evaluate their effectiveness. 'Finally the protein was eluted from the resin in 1 mL fractions using elution buffer (50 mM HEPES, 5% glycerol, 150 mM NaCl, 200 mM imidazole and 0.05% DDM, pH 7.4). The amount of protein in each fraction was measured spectrophotometrically at 280 nm with the most proteinaceous fractions being pooled and subsequently dialyzed against dialysis buffer (50 mM MES, 5% glycerol, 150 mM NaCl and 0.05% DDM, pH6) with a minimum 333.3 fold excess in volume. In order to ensure a maximal amount of imidazole was removed a minimum of two rounds of dialysis was performed on MPSIL0294-SBP samples, after which the samples were concentrated in a Vivaspin® concentrator (Sartorius) with a 50 kDa molecular weight cut off. A 10 µL sample of the concentrated protein was then analysed on a 12% SDS-PAGE gel and western blotting (**see section 2.4.7 and 2.4.8**) alongside all of the other samples taken throughout the purification. The remaining protein was snap frozen in liquid nitrogen and stored at -80°C.

2.5.1.2 Purification of MPSIL0294-Avitag

Purification of MPSIL0294-avitag was identical to MPSIL0294-SBP. To biotinylate this protein *in vivo*, the avitag construct was, however, expressed in AVB101 with 50 µM biotin added to the auto-induction media (**see Table 2.6**).

2.5.1.3 Purification of MPSIL0294-V532C

MPSIL0294-V532Cs expression was identical to MPSIL0294-SBP also; it was however, solubilised and purified with 0.5 mM TCEP in all buffers. The addition of TCEP was designed to prevent MPSIL0294-V532C forming disulphide bridges with arbitrary endogenous *E.coli* proteins with an accessible cysteine leading them to co-purify. The TCEP was also useful as it reduces the cysteine residues thereby making them suitable for labelling with biotin-maleimide.

2.5.2 Purification of AcrB

Membranes of AcrB were prepared by Dr. Pete Roach using a 30 L culture of C43 (DE3) produced in a fermenter. The AcrB gene was present in the plasmid pBPT0480-C0H which has been detailed previously (Glover, Postis et al. 2011). In order to purify AcrB, typically less than 5 mL of the membranes were diluted in 20 mL of ice cold 2x solubilisation buffer (2x PBS, 15 mM imidazole, 20% glycerol and EDTA free protease inhibitor cocktail, pH 7.4) and ice cold ultrapure water to a final volume of 38.4 mL. The membrane solution was then homogenised before 1% DDM was added (1.6 mL of a 25% stock), thereby resulting in a final protein concentration of approximately 2.5 mg/mL. The solubilising membranes were incubated on a roller mixer at 4°C for 1 hour before a sample was taken for further analysis. To remove any insoluble material the solubilised membranes were then ultracentrifuged at 100,000 g_{av} for 1 hour and samples of the pellet and supernatant were taken. The supernatant was then added to a BV of 0.6 mL of Hispur™ cobalt resin pre-equilibrated in wash buffer 1 (1x PBS, 7.5 mM imidazole, 10% glycerol, 0.05% DDM, pH 7.4) and incubated for 2 hours at 4°C on a roller mixer. The bound resin was then centrifuged in an Eppendorf 5804R centrifuge at 700 g_{av} for 5 minutes and the supernatant was discarded (after a sample was taken). The resin pellet was then washed twice in 10BV (6 mL) wash buffer 1 via centrifugation at 700 g_{av} for 5 minutes before resuspension in 3 mL of this same buffer and packing into a disposable column. The resin was then washed by gravity flow in 30BV (18 mL) wash buffer 1 followed by the same volume of wash buffer 2 (1x PBS, 20 mM imidazole, 10% glycerol, 0.05% DDM, pH 7.4). The wash buffers were discarded once samples were taken for further analysis. AcrB was then eluted in 1 mL fractions using elution buffer (1x PBS, 250 mM imidazole, 10% glycerol and 0.05% DDM, pH 7.4) in an identical manner as MPSIL0294-SBP (**see section 2.5.1.1**). The peak fractions of AcrB were then pooled and extensively dialysed in 1x PBS supplemented with 10% glycerol and 0.05% DDM, before they were concentrated in a vivaspin protein concentrator with a 100 kDa molecular weight cut off and snap frozen in liquid nitrogen. All of the samples taken throughout the purification were then analysed via SDS-PAGE and western blotting.

2.5.3 Purification of VcCNT

Once again membranes of VcCNT were provided by David Sharples from a 30 L *E.coli* culture using the plasmid pET26-MPSIL0206-VcCNT. The plasmid is designed to place a pelB signal sequence followed by a 10x polyhistidine tag and a maltose binding protein at the N-terminal of VcCNT. The protein and tags are separated by a HRV3C cleavage site. Purifying VcCNT was similar to AcrB in that less than 5 mL of membrane was diluted in 25 mL ice cold solubilisation buffer (100 mM Tris, 300 mM NaCl, 10 mM imidazole, 20% glycerol, pH 7.4) and made up to a

final volume of 47 mL in ice cold ultrapure water. The solution was then homogenised and 1.5% DDM (3 mL from a 25% stock) was added, bringing the final concentration of solubilisation buffer to 1x and the protein concentration to approximately 2 mg/mL. The membrane solution was then incubated for 2 hours at 4°C on a roller mixer. Once solubilisation was complete the insoluble fraction was isolated via ultracentrifugation at 100,000 g_{av} for 1 hour; the resultant supernatant was added to a BV of 1 mL HisPur™ cobalt resin pre-equilibrated in wash buffer (50 mM Tris, 150 mM NaCl, 15 mM imidazole, 10% glycerol and 0.05% DDM, pH 7.4). VcCNT was then allowed to bind to the resin by incubating the supernatant/resin solution at 4°C for 2 hours on a roller mixer, after which the resin was pelleted, washed in wash buffer and eventually eluted in elution buffer (50 mM Tris, 150 mM NaCl, 250 mM imidazole, 10% glycerol and 0.05% DDM, pH 7.4) in an identical manner as described in **section 2.5.1.** Peak fractions of VcCNT were dialysed in 50 mM Tris supplemented with 150 mM NaCl, 10% glycerol and 0.05% DDM (pH 7.4). The protein was then concentrated in a vivaspin concentrator with a 30 kDa molecular weight cut off. Samples were taken throughout the VcCNT purification and were once again analysed with SDS-PAGE and western blotting.

2.5.4 Purification of membrane protein via Styrene Maleic Acid co-polymer (SMA)

In order to produce styrene maleic acid lipid particles (SMALPs) 1.5g of bacterial membrane with a concentration of 20 – 40 mg/mL, was diluted in 32 mL of a buffer containing 50 mM tris, 10% glycerol, 500 mM NaCl (pH 8) and mixed with 2.5% (w/v) SMA co-polymer (provided by Tim Dafforn in the university of Birmingham). Once the entire SMA polymer was dissolved, the solution was incubated at room temperature for 2 hours on a rocker before it was subjected to ultracentrifugation at 145,000g_{av} for 45 minutes. The supernatant was then incubated with a bed volume (BV) of 2 mL pre-equilibrated cobalt resin overnight at 4°C. Once the SMALPs were bound to the resin, it was washed in 10 BV (i.e. 20 mL) of the previously described tris buffer supplemented with 10 mM imidazole. The protein SMALPs were then eluted from the resin in 6.25 BV (i.e. 12.5 mL) elution buffer (500 mM NaCl, 10% glycerol, 50 mM Tris and 250 mM imidazole, pH 8). The eluted SMALPs were collected in fractions and their A_{280nm} were monitored spectrophotometrically using a micro-spectrophotometer.

2.5.5 Preparation of protein nanodiscs

In order to construct nanodiscs, a precise molar ratio of membrane scaffold protein (MSP): membrane protein: lipid is required. To determine this ratio, a calculation which takes into

account the amount of lipid that is encapsulated in a single nanodisc, its occupancy, the area of the target protein and the subsequent amount of lipid displaced upon its entry into the nanodisc must be performed.

Once an appropriate ratio is estimated, the membrane protein must first be purified in detergent micelles. A stock solution of POPC lipid was also prepared by drying chloroform lipid stock (25 mg/mL) in a glass tube under a gentle stream of nitrogen gas until all chloroform was removed. The 100 mg of dry lipid was then solubilised in 657.83 µL of lipid solubilisation buffer (20 mM Tris-HCl pH 8, 150 mM NaCl and 400 mM sodium cholate) resulting in a 200 mM lipid solution in 400 mM sodium cholate. The lipid was the subjected to vortexing, incubation at 60°C and sonication in water bath sonicator until the solution was transparent, at which point it was stored at -80°C.

The membrane protein, lipid and membrane scaffold protein were then mixed according to the previously determined molar ratio. It was vital that the final concentration of sodium cholate remained between 12 mM and 40 mM (CMC = 9.5 mM). If required, extra sodium cholate was added from a 1M stock solution prepared prior to nanodisc construction. The nanodisc solution was then incubated at 4°C for 2 hours before the detergent was removed using pre-equilibrated bio-beads and gentle agitation in order to keep the bio-beads suspended. Detergent removal was continued until all detergent was removed (the level of detergent was tested by gently passing air through the solution) and the bio-beads were replaced with a fresh stock twice a day. Typically this process took a minimum of 1 week. Once complete the nanodiscs were stored at 4°C and analysed by gel filtration and western blotting (**see section 2.4.8**).

Alternatively nanodiscs were analysed via sucrose density gradient and involved adding 250 µL of nanodiscs to 250 µL 60% sucrose buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 60% sucrose). The resultant 30% nanodisc solution was then applied to a Beckman TLS-55 centrifuge tube (max capacity 2.2 mL) and gently overlaid successively with 0.5 mL samples of buffer containing 20%, 10% and 5% sucrose. The samples were centrifuged for 16 hours at 259,000 g_{av} before each fraction was collected and analysed via western blotting (see **section 2.4.6**).

In order to biotinylate the MSP, a buffer exchange is required in order to remove the protein from the Tris buffer it is eluted in (40 mM Tris, 0.3 M NaCl, 0.4 M Imidazole, pH 8) into PBS. This was accomplished via extensive dialysis and was necessary in order to avoid competition between the primary amines on lysine residues and the Tris for the biotinylation reagent.

Biotinylation was then carried out as described in **section 2.4.4.** The resultant biotinylated MSP was then used in the construction of AcrB nanodiscs.

2.5.6 Sub-cloning the DARPin gene into expression vector pET16b

The sub-cloning of the DARPin gene was designed to remove the gene from pC6 (**Figure 2.1**) into the expression vector pET16B, thereby removing the reliance on the DsbA signal sequence and the requirement of the amber stop codon and placing a His-tag on the N-terminal of the DARPin. In order to achieve this, an N-terminal NdeI and a C-terminal BamHI cut site were inserted at the two ends of the DARPin gene in the pC6 via PCR (**see section 2.3.1**) using the following primers:

- Forward primer (introducing the **NdeI** site) pET16b_DP_Fwd:
  - 5' – CGAT**CATATG**GATCTGGGAAAAAAACTGCTGGA – 3'
- Reverse primer (introducing the **BamHI** site) pET16b_DP_Rev:
  - 5' – ATCG**GGATCC**TCACTACAGTTTCTGCAGG – 3'

The amplified DARPin was inserted into an empty pET16B vector via double restriction digest with NdeI and BamHI followed by ligation (**see section 2.3.3** and **section 2.3.6**). The newly constructed plasmids were then used to transform the *E.coli* strain Omnimax™2 thereby amplifying the concentration of DNA.

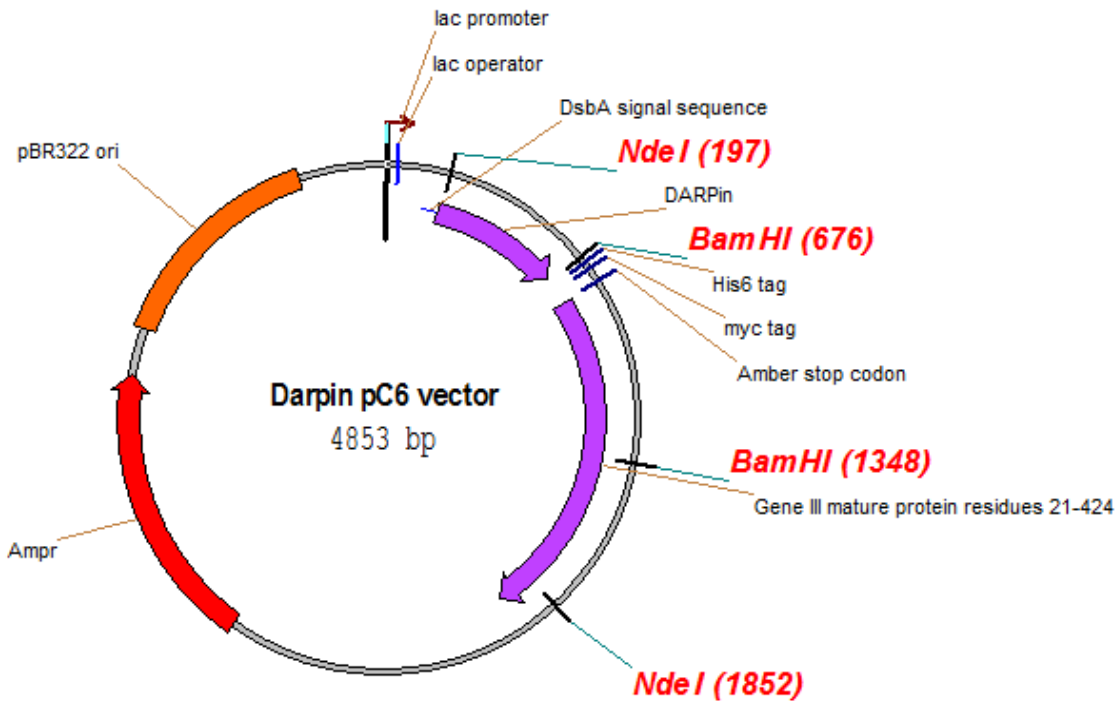


Figure 2.1 Plasmid map of the DARPin phagemid pC6: A diagrammatic cartoon of plasmid pC6 in which all DARPins are originally inserted after the PCR which places a novel BamHI cut site (676) as well as a novel NdeI (197) cut site at each end of the DARPin. The DARPin (represented by the smaller purple arrow) is fused to the coat protein pIII of the bacteriophage M13 (the larger purple arrow) thereby allowing its presentation on the surface of a phage particle. The two proteins are separated by a polyhistidine tag, a myc tag and an amber stop codon. The N-terminal of the DARPin has a DsbA signal sequence which facilitates its rapid translocation to the periplasm thereby preventing premature folding of the DARPin in the cytoplasm.

2.5.7 Periplasmic extraction of cells expressing DARPins

As previously mentioned the plasmid in which the DARPin was placed, was designed to facilitate its co-translational release into the periplasm therefore periplasmic extractions were carried out on the cultures. They were first pelleted and resuspended in 1 mL TSE buffer (20% (w/v) sucrose, 30 mM Tris/HCl, 1 mM EDTA, pH 8.3) before incubation at room temperature for 15 mins with gentle mixing on a roller. The samples were then centrifuged at 5000 g_{av} for 15 min and the supernatant removed and stored as periplasmic fraction 1. The pellet was then resuspended once again in a minimal amount of TSE buffer before being immediately diluted in 1 mL ice cold ultrapure water before being stored on ice for 15 min. The solution was then centrifuged at 16,000 g_{av} for 30 min and the supernatant taken and stored as periplasmic fraction 2. The resultant pellet was finally resuspended in 1 mL ice cold MilliQ water and stored as the cytoplasmic fraction. This procedure was performed on both IPTG induced and non-

induced samples and each fraction was subjected to SDS-PAGE and analysis by western blotting alongside a positive control of MPSIL2094.

2.5.8 Expression and purification of Designed ankyrin repeat proteins (DARPins)

DARPins were expressed in BL21 cells via IPTG induction in 500 mL of 2xTY media (**see Table 2.2**). The expression was done in a slightly modified form from that described in **section 2.4.2.** Initially 30 mL pre-cultures of transformed BL21 cells were grown overnight at 37°C with shaking at 280 rpm. 25 mL of this culture was then used to inoculate 500 mL of 2xTY media supplemented with 100 µg/mL carbenicillin resulting in a 1:20 dilution. The 500 mL cultures were then incubated for 2 hours at 37°C with shaking at 200 rpm, after which IPTG was added to a final concentration of 0.6 mM. Upon the addition of IPTG, the cultures were incubated for a further 4 hours at 37°C with shaking. The cells were then pelleted at 4,000 rpm for 20 minutes and resuspended in 2x PBS with 5 mM imidazole before immediate disruption in the cell disruptor (**see section 2.4.3).** In order to pellet the membranes, the output from the cell disrupter was then centrifuged at 150,000 g_{av} and the supernatant was added to 500 µL (BV) pre-equilibrated (in 2x PBS, 5 mM imidazole) Hispur™ Cobalt resin. The mixture was then incubated at 4°C overnight on a roller mixer. Once DARPin binding was complete the resin was packed into a disposable gravity flow column by allowing the buffer to flow through it. The resin was then washed in 10BV (i.e. 5 mL) 2x PBS buffer with 20 mM imidazole. The DARPin was then eluted (in 2xPBS with 250 mM imidazole) and dialyzed (in 2x PBS) in an identical manner a described in **section 2.5.1.1.** The DARPin was then concentrated in a vivaspin concentrator with a 10 kDa molecular weight cut off before its concentration was determined spectrophotometrically, the protein was snap frozen in liquid nitrogen and stored at -80°C.

2.6 Quality control of MPSIL0294 samples

The three MPSIL0294 protein constructs were designed to immobilise onto a solid surface via a biotin binding protein. In the case of the phage display this was streptavidin coated plates (Thermo scientific). Verification of the proteins ability to bind via their various tags used manually produced neutravidin coated plates in an ELISA procedure.

2.6.1 Producing Neutravidin Coated plates

In order to test the quality of protein samples in relation to their ability to present an epitope for binding they were tested in an ELISA on neutravidin coated plates. Initially a 10 mg/mL Neutravidin (Thermo scientific) stock solution was prepared in 50% glycerol. This stock solution

was then used to make a 2 μg/mL working solution, 100 μL of which was used to coat the wells of a Nunc Maxisorp® plate. The plate was then sealed with an adhesive cover and incubated overnight at 4°C, after which it was washed three times in 1xPBS and stored at 4°C.

2.6.2 Indirect ELISA of MPSIL0294 samples on Neutravidin Coated plates

A serial dilution of MPSIL0294 from 100 to 0.098 μg/mL was established in the 96-well neutravidin coated plate. The protein dilutions were then allowed to bind to the surface by incubating the plates (which were sealed with an adhesive cover) for 1 hour at 4°C. Once binding was complete the plates were washed three times in PBS before they were blocked in 200 μl 3% (w/v) dried skimmed milk (Marvel) in PBS for 1 hour at room temperature, after which the plates were washed three times once again in PBS. The plates were then probed with 50 μL of an anti-polyhistidine tag antibody conjugated with HRP diluted 1:10,000 in the 3% (w/v) milk block buffer. The plates were once again incubated for 1 hour at room temperature before it was washed three times in PBS supplemented with 0.05%Tween (PBST). 50 μL of tramethylbenzidine (TMB) was added and the plates were incubated at room temperature for a maximum of 20 minutes after which 50 μL of 0.5 M H_2SO_4 was added to stop the reaction. The absorbance at 450 nm was then measured using the Multiskan Ascent 96/384 plate reader (Thermo scientific).

2.6.3 Fluorescence-activated cell sorting (FACs) analysis of detergent solubilised MPSIL0294 samples immobilised on streptavidin coated magnetic beads

FACs analysis was performed in order to show that the MPSIL0294 detergent solubilised variants had the ability to bind to a streptavidin coated surface. In order to perform this analysis 10 μg/mL of protein was bound to streptavidin coated Dynabeads® (Thermo Scientific) and incubated for 1 hour at room temperature. The beads were then washed three times in PBS (using a magnetic Eppendorf stand) before being blocked in 3% (w/v) skimmed milk (Marvel) in PBS for 1 hour at room temperature. After blocking, the beads were once again washed 3 times in PBS before being probed with a mouse polyhistidine antibody (at a dilution of 1:1000) in block buffer and incubated at room temperature for 1 hour; after which the beads were washed three times in 0.05% PBST. A goat anti-mouse antibody conjugated with DyLight 488 was used as a secondary antibody (Jackson ImmunoResearch Ltd). The beads were then treated in an identical manner as they were after incubation with the primary antibody. Blank beads with no bound MPSIL0294 and a biotinylated and polyhistidine tagged control

protein (which had been characterised by MedImmune) was also prepared in an identical manner in order to act as a negative and positive control, respectively. The beads were then washed again in FACs buffer (10% FBS (Fetal bovine serum), 1 mM EDTA, 0.01% Azide in 1x PBS) prior to analysis with a FACs machine which measured the fluorescence at 488 nm and plotted the results on an overlay histogram.

2.7 Assay for Zn^{2+} uptake by MPSIL0294 via Fluozin-1

2.7.1 Preparing a stock solution of small unilamellar vesicles

An assay was developed in order to test MPSIL0294s ability as a divalent metal cation transporter. Initially a 50 mg/mL stock solution of small unilamellar vesicles (SUV) was prepared by drying 100 mg of *E.coli* polar lipids dissolved in chloroform in a round bottom Corex glass tube under a gentle stream of nitrogen. In order to remove any residual traces of chloroform the tubes were left in a desiccator under continuous vacuum overnight. The lipids were then rehydrated in 1x Bis-Tris-Mes (BTM - 20 mM Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane with the pH adjusted to 6.8 with 20 mM Mes) supplemented with 2 mM 2-mercaptoethanol by vigorous vortexing. The lipid solution was then sonicated in a bath sonicator until it became a translucent suspension of SUVs.

2.7.2 Reconstitution of MPSIL0294 for stopped flow analysis

The reconstitution of MPSIL0294 was based of the work performed by Dax Fu (Chao and Fu 2004), as such detergent solubilised MPSIL0294 was diluted in protein dilution buffer (20 mM BTM, pH 6.8, 2 mM 2-mercaptoethanol, 1% (w/v) n-Octyl-β-D-Glucoside (β-OG), 50 mM K_2SO_4) down to a concentration between 0.26 and 0.52 mg/mL. Lipids were prepared by adding 100 µL 10% β-OG to 130 µL (6.5 mg) of the 50 mg/mL SUV stock solution in BTM buffer. The complete transparency of the lipid solution was indicative of complete lipid solubilisation. 770 µL of diluted MPSIL0294 was added and the protein-lipid-detergent mixtures were incubated at 4°C for 20 minutes. The samples were then applied to a PD-10 column pre-equilibrated in assay buffer (20 mM BTM, pH 6.8, 50 mM K_2SO_4) and the cloudy void volume was collected and ultracentrifuged at 140,000 g_{av} for 45 minutes. The pellet was then resuspended in 200 µL of assay buffer and 22 µL of the Zn^{2+} sensitive dye Fluozin-1. In order to maximise the amount of dye uptake into the proteoliposomes, the proteoliposome/dye solution was sonicated in a bath sonicator for 10 seconds, subjected to a single freeze thaw cycle in liquid nitrogen, subjected to extrusion through a 1 µm filter and finally sonicated for an additional 10 second.

The dye loaded proteoliposomes were then applied to a second PD-10 column in order to remove excess dye and the cloudy void fraction was once again collected.

The proteoliposomes (and control liposomes which were formed in an identical manner in the absence of MPSIL0294) were then tested against varying concentrations of $ZnCl_2$ in a stopped flow machine (Applied photophysics) measuring the fluorescence at 490 nm (F). The background readings were also measured by measuring the fluorescence when no Zn was present (F_0) and the maximum possible amount of fluorescence (F_{max}) was measured by mixing the proteoliposomes with 20 μM pyrithione (a Zinc ionophore) in the stopped flow mixing cell. Due to the fact that F_{max} and F_0 should be a single figure the mode of each data set was used as the true F_{max} and F_0. The data was then analysed using the following formula and $\Delta F/\Delta F_{max}$ was plotted on a graph as a function of time:

$$\text{Fraction } Zn^{2+} \text{ uptake (\%)} = 100 \times F - F_0/F_{max} - F_o$$

2.8 Screening of immobilised membrane protein samples via phage display

2.8.1 First round selections on streptavidin coated plates

Phage display was carried out onsite in MedImmune's screening labs in Cambridge. In order to screen immobilised proteins against a library of potential binding partners expressed on the surface of bacteriophage, a culture of TG1 cells were first prepared. This culture was established by inoculating 50 mL 2xTY media with a single colony and incubated at 37°C with shaking at 300 rpm (3 g). This incubation was maintained until the culture had an OD_{600nm} of 0.5 – 1 (i.e. exponential phase), this typically took 4 -5 hours and once the required OD_{00nm} was obtained the culture was placed on ice until it was needed.

For a single phage display screen against a single target protein, one 50 μL aliquot of the naïve DARPin library (which had a size of 7.5×10^9) was blocked in an equal volume of 'phage blocking buffer' (2x PBS, 6% w/v skimmed milk powder). Simultaneously a 100 μL aliquot of streptavidin coated Dynabeads® were vortexed in order to resuspend the beads. The beads were then captured using a magnetic Eppendorf stand and subsequently blocked in 100 μl 'well blocking buffer' (1x PBS, 3% w/v skimmed milk powder). The aliquots of phage and beads were then placed separately in an orbital mixer for 1 hour at room temperature, after which the beads were once again placed in a magnetic stand (in order to separate the beads from the

blocking buffer) and the supernatant was discarded. The entire 100 μL blocked phage aliquot was then added to the beads, vortexed gently and placed back into the orbital mixer for an additional hour at room temperature in order to deselect the phage against streptavidin.

125 μL of target protein diluted to a concentration of 50 - 100 μg/mL in PBS (with an appropriate concentration detergent if necessary) was added to a single well of a streptavidin coated plate (Pierce™) and incubated for 1 hour at 4°C. After the target protein had been immobilized on the plate's surface, it was washed three times with PBS (with detergent if required), blocked in 300 μL of 'well blocking buffer' and incubated once again at 4°C for 1 hour. After blocking, the plate was washed once in PBS, then the beads were placed in a magnetic stand once again and the 100 μL deselected phage in the supernatant was pipetted onto the appropriate well of the streptavidin coated plate. The plate was then incubated at 4°C for 1 hour in order to allow any phage capable of binding to do so.

The plates were then washed 5 times in PBST (1x PBS, 0.05% Tween) prior to elution of the phage via trypsinization (Thomas and Smith 2010). This was achieved with a 10 μg/mL working solution of trypsin (made from a 5 mg/mL stock diluted in 0.1M sodium phosphate buffer) 0.001 μg of which (100 μL) was pipetted onto the appropriate well on the plate. The eluted phage was then used to inoculate a 5 mL aliquot of the exponential phase TG1 cells. An additional 100 μL of 0.1 M sodium phosphate buffer was then used to wash out the well and added to the culture as well. This 5 mL culture was then incubated for 1 hour at 37°C with shaking at 150 rpm. 100 μL of a serial dilution of the TG1 cells (from pure to 10^{-5}) were plated on 2x TY agar plates supplemented with 100 μg/ml ampicillin in order to calculate the phage titre.

The remainder of the inoculated TG1 culture left was centrifuged for 5min at 2000 g_{av} after which the cells were resuspended in 1ml TY media. This entire 1ml sample was then spread on a large 24.5 cm x 24.5 cm square bio-assay plate, onto which 2x TY agar supplemented with 100 μg/mL ampicillin had been poured. The plates were then incubated overnight at 37°C.

2.8.2 Phage rescue of selection output

In order to recover the phage produced by the TG1 cells spread on the bio-assay plate, a phage rescue was necessary. Firstly 2xTY media was mixed with 50% (v/v) glycerol in a 2:1 ratio and supplemented with ampicillin. 10 mL of this 2x TY-glucose media (TYAG) was then spread across the surface of the bio-assay plate (in a sterilised hood) and the cells (and phage) were gently scraped using a disposable cell scraper. The cells were then transferred into a 50 mL

falcon tube and fully resuspended by placing the tubes into an orbital mixer for 10 minutes. Two 1 mL aliquots were taken from this cell suspension, pipetted into cryotubes (Nunc) and stored at -80°C as a selection backup. The OD_{00nm} of the undiluted cell suspension was then calculated by spectrophotometrically analysing a 1/100 dilution in 2x TY media and used to inoculate 25 mL of TG1 culture in 2x TYAG media (2x TY media as before, 2% w/v glucose, 100 μg/mL ampicillin) at an OD_{600nm} of 0.1. This 25 mL culture was grown in a 250 mL conical flask at 37° with shaking at 280 rpm until it reached an OD_{600nm} of 0.5 – 1. At this point the cultures were superinfected with 2.5 μL M13K07trp helper phage (from a stock containing 7.5 $x10^{10}$ pfu) which carries essential genes for the production of fully formed phage particles (Carmen and Jermutus 2002). The cultures were then incubated for an additional hour at 37°C with gentle shaking 150 rpm.

In order to express the DARPin proteins on the surface of the phage particles, due to glucoses ability to repress the lacZ promoter it needed to be removed from the growth media in order to allow expression of the DARPin, coat protein fusion (Carmen and Jermutus 2002). To this end the cultures were transferred into 50 mL falcon tubes and centrifuged at 2000 g_{av} for 10 minutes. The supernatant was subsequently discarded and the pellet was resuspended in 25 mL 2x TYAK media (2xTY as before, 100 μg/mL ampicillin, 50 μg/mL kanamycin). The cultures were then grown overnight in 250 mL conical flasks at 37°C with shaking at 280 rpm (1 g_{av}). Once the overnight incubation was complete 1 mL of this culture was centrifuged at 27,000 g_{av} (max speed) in a microfuge and the phage contained within the supernatant was stored on ice for the second round of selection (**see section 2.8.4**).

2.8.3 Calculating phage input and output titres

The input and output titre acts as a preliminary measure of the success of a phage display screen. A typical output titre is between 10^3 – 10^6 cfu (colony forming units) and was calculated by counting the colonies on the serial 2x TY dilution plates which were grown alongside the bio-assay plate after a round of phage selection (**see section 2.8.1**). Colonies were counted using the Acolyte automated colony counter and the output titre was calculated using the following formula:

- Number of colonies x dilution factor of plate x 10 = cfu/mL

A typical input titre on the other hand measures the number of phage particles which enter a subsequent cycle of phage display. The input titre is typically between 5 x 10^{11} – 10^{13} and was

calculated at the same time as the second round of selection using the phage solution isolated from the selection rescue as follows (**see section 2.8.2**).

A colony of TG1 cells is established in an identical manner as discussed in **section 2.8.1.** Once the TG1 cells reached exponential phase (OD_{600nm} = 0.5 – 1) a 990 µL aliquot was placed in a single well of a sterilized Greiner Masterblock® (Sigma), 10 µL of 10^{-4} dilution of the phage was then added to this aliquot (giving a 10^{-6} dilution) and the block was incubated for 1 hour at 37°C with shaking at 150 rpm. After preparing further dilutions of the TG1/phage mixture, 100 µL of the 10^{-8} – 10^{-10} dilutions were then spread on 2x TY agar plates supplemented with 100 µg/mL ampicillin and incubated at 30°C overnight. The resultant colonies and subsequent titre was calculated in an identical manner as discussed for the output titre.

2.8.4 Second Round selections on streptavidin coated plates

The second round of selection was performed simultaneously with the input titre; a 50 µL aliquot of the isolated phage solution from the selection rescue (**see section 2.8.2**) was blocked in an equal volume of 'phage blocking buffer' (**Table 2.14**). The selection was then performed in an identical manner as discussed in **section 2.8.1**.

Once the second round of selection was complete, a 96-well master plate was established for the output. To do this, 120 µL aliquots of 2xTYAG was pipetted into a 96 well round bottomed plate (Costar). 88 colonies were used to inoculate 88 of the wells (the final row was left with media alone in order to act as a negative control). The plate was then sealed and stored in a box with a moist tissue to prevent evaporation and placed in a slow rotating incubator at 30°C overnight. Once incubation was complete 40 µL was transferred to a fresh 96 well plate (ensuring that each culture was transferred to the same well) and sent for sequencing by MedImmune's internal sequencing lab. 40 µL 50% glycerol was added to the remaining cultures in the first plate (henceforth known as the master plate) and the plate was subsequently stored at -80°C.

2.9 Analysing the phage outputs via phage ELISA

2.9.1 96 well phage preparation

Master plates (**see section 2.8.**4) were removed from storage and thawed at room temperature. A separate 96 deep well plate was loaded with 500 µL 2xTYAG media and (once the master plate was completely thawed) inoculated with all 96 TG1 cultures of the master plate (including the negative control) using a sterile 96 well plate inoculator. The master plate

was returned to storage while the 'daughter' plate was incubated at 37°C at 280 rpm for 5 hours. The helper phage was then prepared by adding 5 μL M13K07 per 10 mL 2xTYAG media (a single daughter plate requires 10 mL). Once the 5 hour incubation was complete the plate was visually checked for growth in each well before 100 μL of diluted helper phage was added in each well thereby superinfecting the cultures. It was then incubated at 37°C for 1 hour at 150 rpm. Once again in order to develop binding partners on the surface of the phage, glucose needed to be removed from the media. Therefore the plates were centrifuged in a benchtop centrifuge for 10 mins at 2000 g_{av} at room temperature. Following centrifugation the supernatants of each well were discarded and the pellets were resuspended in 500 μL 2xTYAK media. The plate was then incubated overnight at 25°C, 280 rpm. Finally the phage cultures were blocked in 500 μL 'well blocking buffer' and incubated for 1 hour at room temperature. The plate was then centrifuged for a second time at 2000 g_{av} and the supernatants were ready to be used in phage ELISAs (**see section 2.9.2**).

2.9.2 Phage ELISAs on blank streptavidin plates

Despite deselecting the phage against streptavidin during each round of selection, an ELISA was still performed to ensure that any potential binding partners which showed an ability to bind directly to streptavidin were removed. Streptavidin coated plates were first blocked in 300 μL 'well blocking buffer' for one 1 hour at room temperature. Once blocking of the plate was complete the buffer was discarded and the plate was washed three times in PBS. 50 μL of each phage supernatant described in **section 2.9.1** was then transferred from the daughter plate to their corresponding positions on the streptavidin coated plate which was then incubated at room temperature for 1 hour. Once phage binding was complete the plates were washed three times with PBS then treated with 1/5000 anti-M13-HRP antibody conjugate (Thermo scientific) diluted in well blocking buffer, capable of binding to the surface of bacteriophage particles. The plates were once again incubated for 1 hour at room temperature before being subsequently washed three times in 0.1% PBST. The presence of phage was then detected via 5- 20 minute incubation with 50 μL tetramethylbenzidine (TMB) solution followed by the addition of 50 μL 0.5 M H_2SO_4 to stop the reaction, A_{450nm} was then measured in a plate reader.

2.9.3 Phage ELISAs on streptavidin coated plates in the presence of target protein

Following the ELISAs described in **section 2.9.2** all binding proteins which showed an ability to bind directly to streptavidin were removed from further analysis. ELISAs were then repeated in

an identical manner on the remaining binding candidates in the presence of the immobilized target protein. The procedure for these ELISAs was identical as those described in **section 2.9.2** with the exception that the streptavidin coated plates were first coated in 50-100 μg of target protein prior to blocking.

2.10 Confirming the DARPins ability to bind the various target proteins via ELISA

After the DARPins were subcloned, expressed and purified (**see section 2.5.6 and 2.5.7)**, ELISAs were designed to confirm the ability of the most promising DARPins to bind to their various targets. In order to do this they were first biotinylated using the reagent NHS-biotin (the procedure for which was described in **section 2.4.6**) with a 2.5 – 5: 1 molar excess of biotin.

The actual ELISA was performed by adsorbing 10 μg of purified detergent solubilized target protein onto Nunc Maxisorp plates. This was accomplished by diluting the various proteins down to the desired concentration in 2x PBS supplemented with 0.05% DDM, pipetting 50 μL into the required wells of the plate and incubating them overnight at 4°C. Once adsorption was complete the plates were washed three times in 200 μL wash buffer (2 x PBS, 0.05% DDM) prior to incubation at 4°C for 1h in 200 μL blocking buffer (2 x PBS, 3% BSA, 0.05% DDM). During the blocking process, a serial dilution of the biotinylated DARPin ranging from 10 μM to 1 pM was performed in wash buffer. Once the incubation was complete the plates were once again washed three times in 200 μL wash buffer before they were incubated for 1h at 4°C in the presence of 50 μL of the biotinylated DARPin. The plates were then washed for a third time in 200 μL wash buffer after which 50 μL streptavidin-HRP (diluted to a concentration of 1:20,000 in blocking buffer) was added to each well and the plate was incubated at 4°C for 1h. The plates were then washed in 200 μL 0.1% PBST three times before 50 μL TMB solution was added and the plates were incubated for 15 mins at room temperature. To end the ELISA reaction 50 μL of 0.5 M H_2SO_4 was added before the A_{450nm} was measured using a plate reader. For a negative control, the methodology was identical as just described except no target protein was adsorbed to the plates initially.

Chapter 3

A comparison of alternative methods of protein immobilisation using phage display with a DARPin library

3.1 Introduction

The production of novel therapeutic agents against membrane proteins is difficult, in part due to their amphipathic natures. One of the main requirements for the elucidation of a therapeutic agent is their specificity to the target. Therefore prior to screening against binding protein libraries, membrane proteins must be over-expressed in an appropriate expression host, solubilised using detergents such as n-Dodecyl β-D-maltoside (DDM), thereby placing them into micelles and subsequently purified.

In regard to their over-expression *Escherichia coli* has been shown to be a robust host for the production of a wide variety of prokaryotic membrane proteins. This is due to its high growth rate and our extensive understanding of its physiology (Makino, Skretas et al. 2011). Studies have shown however that the success of this over-expression can be influenced by factors such as the topological characteristics of the membrane protein and the position and length of the tag which has been used (Mohanty and Wiener 2004).

During phage display, typically a target protein will be first immobilised onto a solid support, often by a tag. There are several problems with immobilising membrane proteins, primarily due to their propensity to unfold and adopt non-native conformations. Or the adsorption of non-specific proteins to the surface as seen with earlier attempts to immobilise membrane proteins via a Histidine-nitrilotriacetic acid (NTA) interaction (Friedrich, Kirste et al. 2008, Vaish, Silin et al. 2013).

To this end, this chapter shall focus on the development of two membrane proteins which were prepared for phage display screening trails. A single model membrane protein from the BBSRC funded membrane protein structure initiative at Leeds (MPSIL) designated MPSIL0294 - a divalent metal ion transporter from *Enterococcus faecalis* and part of the MntH family of transporters. It was chosen due to the fact that topological predications show that it has 12 transmembrane helices thereby making it more structurally similar to its human homologue when compared to the other bacterial members of this transporter family. Its human homologue - the natural resistance associated macrophage protein (NRAMP), shares a 35% sequence homology with MPSIL0294 and plays a pivotal role in the transport of iron in the body. Due to the12th helix of MPSIL0294 predictive models place its N and C- termini within the cytoplasm therefore making it more amenable to tagging, unlike its homologue MntH from *E.coli*. It has been shown that tagging an extracellular N-terminal can hinder membrane proteins insertion into the plasma membrane and their expression even with a short polyhistidine tag (Rahman, Ismat et al. 2007, Kang, Lee et al. 2013). These three variants of

MPSIL0294 were designed to compare alternative methods of protein immobilisation in regard to their ability to isolate DARPins capable of binding to them after selection with phage display.

In order to compare alternative methods of membrane protein immobilisation, three different C-terminal tags which immobilised MPSIL0294 via a biotin/streptavidin interaction were added. The first being a streptavidin binding peptide (SBP) tag, the second an avitag which was subjected to *in vivo* biotinylation and finally the inclusion of an accessible cysteine via site directed mutagenesis which allowed the protein to be biotinylated chemically at a specific site using biotin maleimide.

AcrB was used as a second model membrane protein, designed to compare alternative methods of membrane protein immobilisation. It was chosen because topological analysis of its crystal structure shows a large extracellular loop, most likely a part of the TolC docking domain (Murakami, Nakashima et al. 2002), is present within the proteins natural conformation. This thereby gives it a higher chance of isolating binding partners than MPSIL0294 which doesn't have any particularly large extracellular components according to topological predictions. AcrB was also chosen due to its natural affinity to bind to IMAC resin due to a cluster of histidine residues in its C-terminus which make extensive tagging unnecessary and its purification simpler (Veesler, Blangy et al. 2008).

3.2 Cloning of MSIL0294 variants

Three MPSIL0294 variants were designed to allow for a consistent method of expression and purification to be utilised. Restriction sites were chosen which allowed the inclusion of a C-terminal polyhistidine tag consisting of 8 histidine residues in tandem downstream from the various newly introduced tags on all three proteins. The MPSIL0294 open reading frame was placed under the control of the lac operon thereby allowing its expression to be controlled via IPTG or auto-induction. A human rhinovirus 3C protease (HRV-3C) motif was engineered downstream of the polyhistidine and other tags. All of these features were included in plasmid construct pBPT0294-CVH designed by Dr. Cheng Ma (Ma 2013).

Two of the three MPSIL0294 constructs were derived from pBPT0294-CVH along with two commercially purchased open reading frames for AtPH01_SPX_Avitag and TaALMT1-CTD which provided the avitag and SBP tag via sub-cloning. Initially a double restriction digest using Hind III and AvrII was performed on pBPT0294-CVH, AtPH01_SPX_Avitag and TaALMT1-CTD (**see section 2.3.3**). The resultant pBPT0294-CVH fragment was then ligated with either the

AtPH01_SPX_Avitag fragment or the TaALMT1-CTD fragment thereby creating the intermediate plasmids pL50 and pL51 respectively (see **Figure 3.2**).

While the original pBPT0294-CVH plasmid has two BamHI cut sites PL50 and PL51 have a single unique BamHI site (see **Figure 3.2**). Therefore a BamHI restriction analysis was used as a method to confirm the construction of pL50 and pL51 on a 1% agarose gel (**see section 2.3.4**). When digested by BamHI, pBPT0294-CVH produced two DNA fragments while pL50 and pL51 produced only one (**see Figure 3.3**).

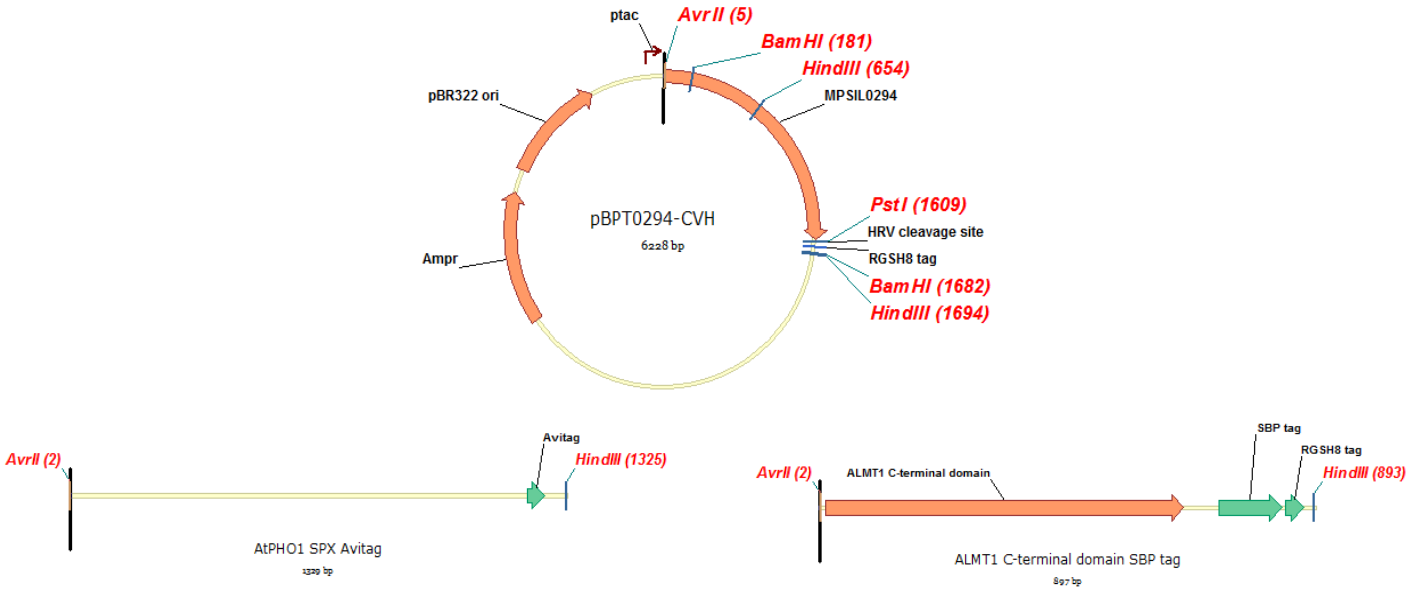


Figure 3.1: **<u>Plasmid maps of pBPT0294-CVH, AtPH01-SPX avitag and ALMT CTD SBP</u>*:* The plasmid constructs of pBPT0294-CVH developed by Dr. Cheng Ma, and the purchased open reading frames of AtPH01_SPX_Avitag and TaALMT1-CTD. In order to produce the two of the three MPSIL0294 variants, a HindIII/ AvrII double restriction digest was performed initially upon pBPT0294-CVH and a resultant fragment was ligated with those produced by AtPH01_SPX_Avitag and TaALMT1-CTD. Within the blue box are the fragments of each DNA construct which were used after digestion with HindIII and AvrII. For clarity all restriction cut sites other than HindIII, AvrII, BamHI and PstI in pBPT0294-CVH have been removed.

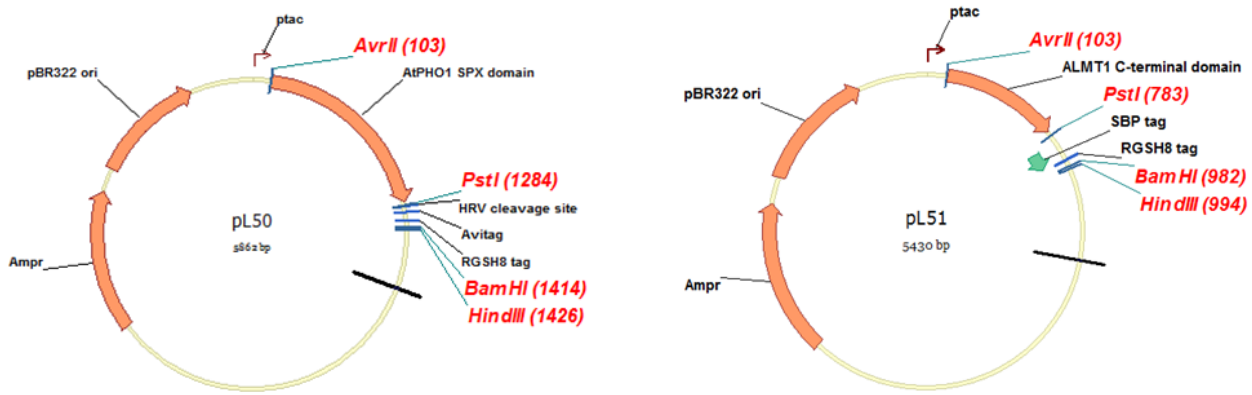


Figure 3.2 <u>Plasmid Maps of pL50 and pL51:</u> **The resultant plasmids produced as a consequence of the ligation between AvrII/HindIII fragments of pBPT0294-CVH and AtPH01_SPX_Avitag resulting in pL50 or those of pBPT0294-CVH and TaALMT1-CTD resulting in pL51. All restriction cut sites have been removed except for those relevant to the cloning strategy. These intermediate plasmids successfully place an avitag (in the case of pL50) and a SBP tag (pL51) into the expression vector but also unwanted genes such as AtPH01 SPX domain (pL50) and ALMT C-terminal domain (pL51). They also no longer contain the reading frame of MPSIL0294.**

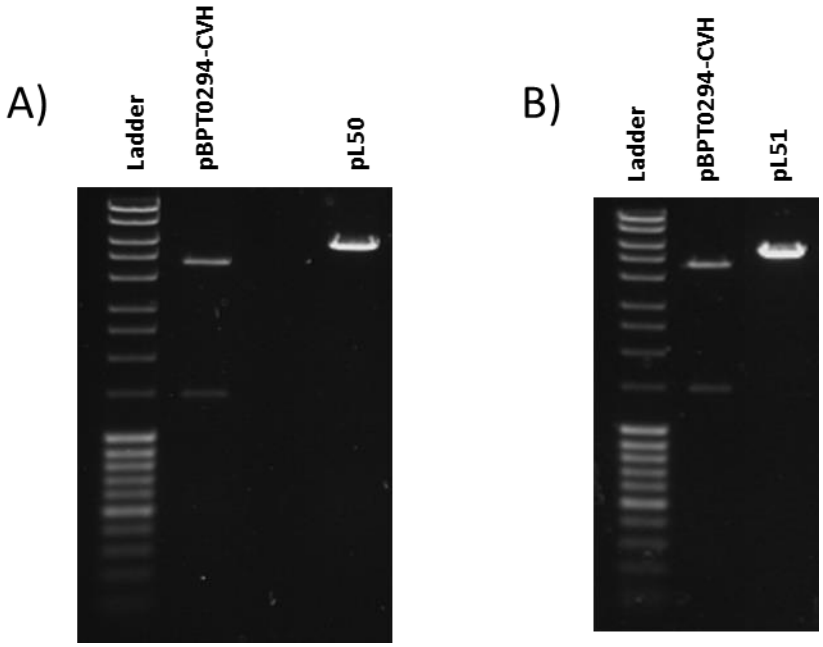


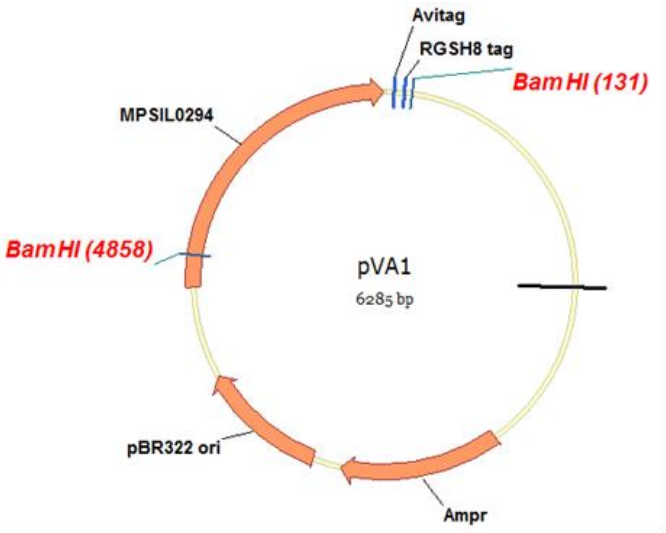
Figure 3.3 <u>BamHI analysis of pBPT0294, pL50 and pL51</u>**: A BamHI digestion was carried out on samples of the newly constructed intermediate plasmids pL50 (A) and pL51 (B). The resultant DNA fragments were analysed on a 1% agarose gel stained with SYBR®SAFE from life technologies and visualised with the G-Box from SynGene.**

pL50 contained the open reading frame for AtPH01_SPX domain which plays a role in phosphatase homeostasis in eukaryotes (Secco, Wang et al. 2012) and is not relevant to this study but introduced an avitag sequence upstream from the polyhistidine-tag. pL51 on the other hand contained the open reading frame of ALMT, a transporter in plants, which plays a

role in aluminium tolerance (Ma and Furukawa 2003). Once again this protein is not relevant to this study but provided the SBP tag (see **Figure 3.2**).

In order to construct plasmids in which the MPSIL0294 gene was upstream from the desired tags, the irrelevant genes were switched with MPSIL0294 using an AvrII and PstI double restriction digest on pBPT0294-CVH, pL50 and pL51. The enzymes removed the open reading frames of MPSIL0294, the AtPH01 SPX domain and the ALMT from the plasmids (see **Figure 3.1 and 3.2**). The MPSIL0294 fragment was then ligated with the PL50 fragment (thereby creating PVA1 MPSIL0294-avitag) and the PL51 fragment (creating PVA2 – MPSIL0294-SBP) (see **Figure 3.4)**.

The insertion of the MPSIL0294 fragment re-introduced a BamHI site thereby allowing pVA1 and pVA2 to be distinguished from pL50 and pL51 once again via BamHI restriction analysis (**see Figure 3.5**).



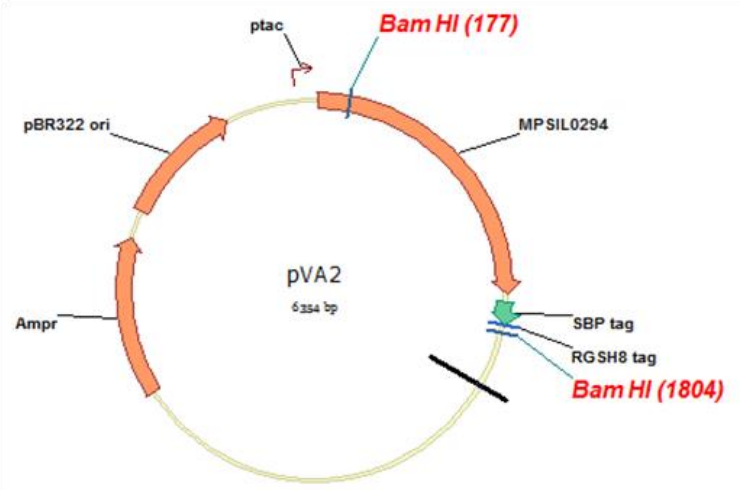


Figure 3.4 <u>Plasmid maps of pVA1 and pVA2:</u> The final plasmids which place MPSIL0294 upstream from either an avitag (pVA1) or a SBP tag (pVA2), while removing the unwanted genes from pL50 and pL51. For clarity all restriction sites other those used throughout the cloning strategy have been removed.

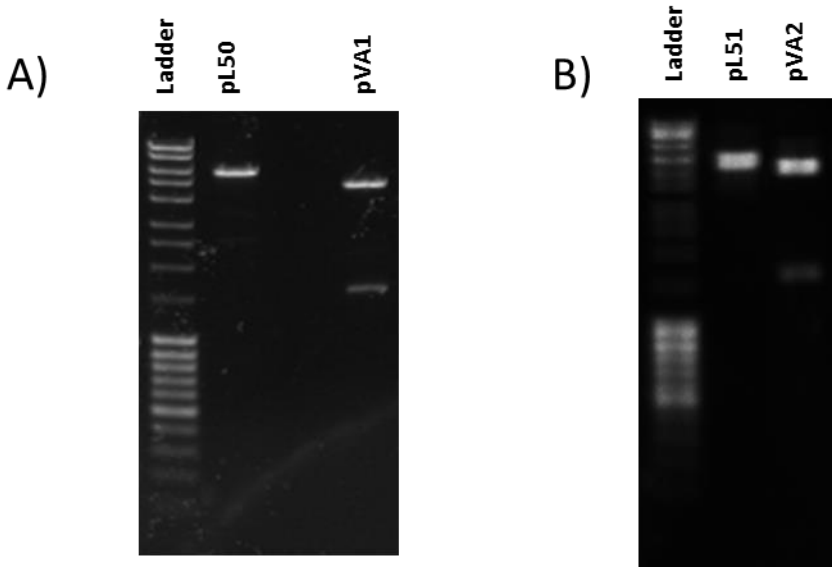


Figure 3.5 BamHI analysis of pL50, pL51, pVA1 and pVA2: Plasmids pVA1 (A) and pVA2 (B) were digested with BamHI in order to check that the two MPSIL0294 constructs had successfully been made. The plasmid fragments were analysed on a 1% agarose gel stained with SYBR®SAFE from life technologies and visualised under UV with the G-Box from SynGene.

In the case of the MPSIL0294-V532C, site directed mutagenesis was used in order to replace a C-terminal valine with an accessible cysteine residue. The original MPSIL0294 plasmid - pBPT0294-CVH was used as the template DNA (**see section 2.3.10**) and utilised the following mutagenic primers:

```
Forward: 5' GAGCAACAAGTAGAGGAGTGTAAACCTGCAGGACTAG 3'
Reverse: 5' CTAGTCCTGCAGGTTTACACTCCTCTACTTGTTGCTC 3'
```

(Mismatched nucleotides have been highlighted in red; their melting temperature was 78.3°C and had a GC content of 48.65%)

Samples from the mutagenesis experiment were analysed and found to only produce an amplification product when the mutagenesis procedure was carried out at 60 or 65°C (see **Figure 2.3.10**). Once the methylated template DNA was removed via DpnI digestion, the newly mutagenized DNA was amplified in OmniMax2 (**see section 2.3.8**).

Samples of each pVA1, pVA2 and pVA3 were then sent for sequencing via Source Bioscience using primers pTACF and pTACR (**see section 2.3.11**). While sequencing confirmed the correct construction pVA1 and pVA2, it showed that all of the pVA3 plasmids suffered from random

insertions of nucleotides into the plasmid thus rendering them unusable. Fortunately the desired codon substitution in MPSIL0294 had been successful while the arbitrary nucleotide insertions occurred upstream of a PstI/Sbf-I cut site. Therefore a double restriction digestion with AvrII and Sbf-I and subsequent ligation was performed on pVA3 and pBPT0294-CVH in order to switch the wild type MPSIL0294 gene with the mutant. The correct construction of pVA3 was confirmed via sequencing with the pTACF and pTACR sequencing primers (**see section 2.3.11**).

3.3 The purification of MPSIL0294 constructs and AcrB

3.3.1 The purification of C-terminal SBP tagged MPSIL0294

MPSIL0294-SBP membranes were prepared via auto-induction in BL21 star (DE3) transformed with pVA2 and subsequently isolated via disruption using a cell disruptor. Samples taken throughout the purification described in **section 2.5.1.1** were analysed electrophoretically via SDS-PAGE followed by staining with Coomassie blue and western blotting (**see Figure 3.6**). This analysis provides a reliable insight into the behaviour of MPSIL0294-SBP throughout the purification. The sample taken from the 'Total' provides an indication of the expression level of the protein. Analysis of the 'supernatant' sample shows the amount of protein successfully solubilised by the DDM; while samples taken from the 'pellet' represent the protein which failed to do so. The 'Flow-through' samples give an indication of how much protein fails to bind to the cobalt resin. Finally the 10 and 40 mM imidazole wash samples shows if any MPSIL0294 is prematurely stripped from the column as a result of imidazole exposure.

Firstly the noticeable lack of prominence of the protein resolved at approximately 55 kDa in the 'Total' sample indicates a poor expression level of MPSIL0294-SBP. The similarity of this protein in the 'supernatant' sample however, (as well as its disappearance in the 'Pellet') shows that the majority of the MPSIL0294-SBP was successfully solubilised. Minimal amounts of protein appear to have been lost in the 10 and 40 mM imidazole washes however they successfully remove part of the 25 kDa impurity. The protein was successfully eluted in 1 mL 200 mM fractions of elution buffer. The majority of MPSIL0294-SBP came off the column after exposure to 10x bed volume of elution buffer (i.e. 10 mL). The final eluate sample produced after the peak fractions were pooled, dialysed and concentrated is relatively pure, except for the presence of some low molecular weight proteins which appear at 20 kDa and below. These observations are given further credence by the western blot of the same samples probed with an anti-polyhistidine tag antibody conjugated with HRP (**see Figure 3.6**). A single His-tagged

protein is shown in the total and supernatant samples, therefore confirming that the protein is almost completely solubilised in 1.5% DDM. The lack of His-tagged protein in the 'Flow-through' confirms that the protein was completely bound to cobalt resin. As before samples of the 40 mM wash shows that some protein is lost however this sample represents 0.1% of the total 40 mM wash while the total and supernatant represents 0.01% therefore a direct comparison would not be accurate. Aside from this the same sample is undetectable when analysed by SDS-PAGE stained with Coomassie.

Several protein bands also appear on the blot once the elution is initiated. These bands occur between 25 and 30 kDa as well as below 10 kDa. After concentration in a 100 kDa cut off vivaspin column, bands also appear at approximately 170 kDa. As can be seen in **Figure 3.6** the majority of these extra bands are not easily detectable by Coomassie stain however. The fact these bands have a lower mass suggest that they are not aggregates and they do not appear in every sample therefore are most likely not artefacts of the SDS-PAGE. Instead they are likely minor degradation products from a protease resistant to inhibitor cocktail used during the purification procedure. Analysis via micro-spectrophotometer showed that 0.76 mg of MPSIL0294-SBP was purified from 200 mg of mixed membrane. Samples of MPSIL0294-SBP were then used to form protein filled nanodiscs and SMALPs as shown in the next chapter.

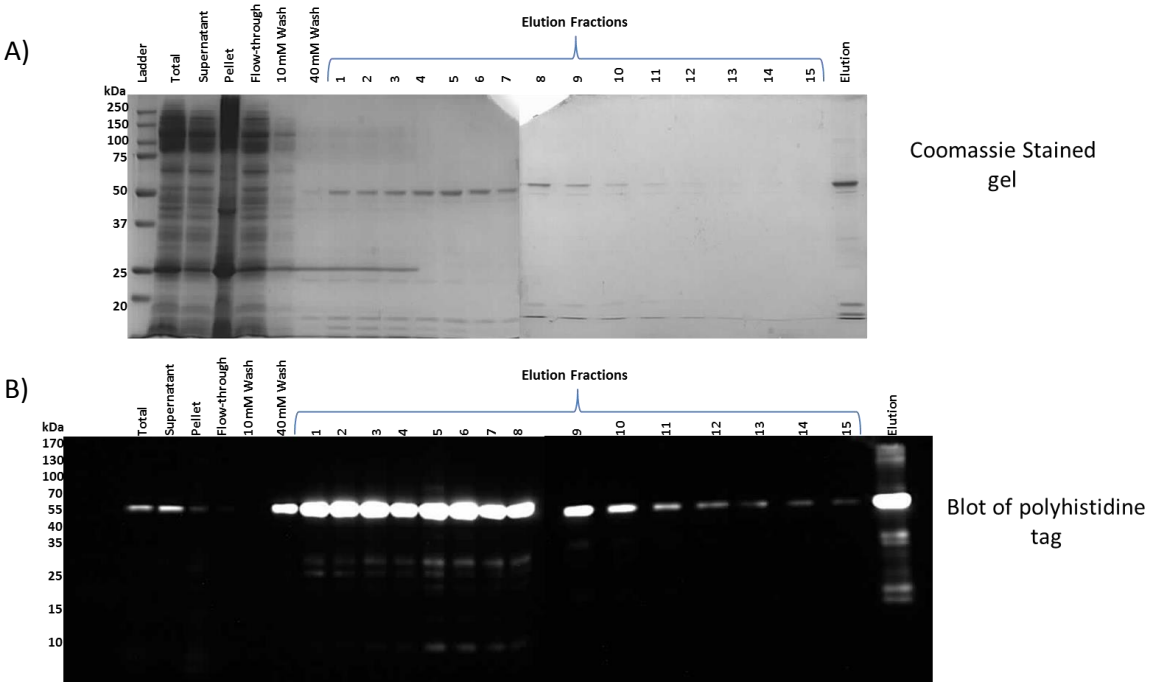


Figure 3.6: <u>IMAC Purification of MPSIL0294-SBP:</u> Samples taken throughout the purification of MPSIL0294-SBP were analysed via 12% SDS-PAGE gel subjected to staining with Coomassie blue and western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP. Each sample was normalised to a volume of 10 µL and represent different relative amounts of their total samples. The total, supernatant and flow-through represent 0.01% of their respective total samples. The pellet as well as both the 10 and 40 mM washes represents 0.1%, while all of the elution fractions as well as the final concentrated protein represent 1%.

3.3.2 The purification of C-terminal Avitag MPSIL0294 after *in vivo biotinylation*

Membranes were prepared in an identical manner than those used to purify MPSIL0294-SBP except the *E.coli* strain AVB101 was used thereby allowing *in vivo* biotinylation of MPSIL0294-avitag. Once again samples were analysed by SDS-PAGE and western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP (**see Figure 3.7**). The fact that no other protein's had a polyhistidine tag means that the band which has been produced on the blot is most likely MPSIL0294-avitag.

It is once again clear that MPSIL0294-avitag does not express to a particularly high level based on the Coomassie stained 'Total' sample. Nonetheless it can be concluded that the protein is successfully solubilised and subsequently bound to the resin, based on the protein bands shown in the blot in conjunction with those resolved on the Coomassie stained SDS-PAGE in **Figure 3.7**. A small amount of MPSIL0294 is visible in the lane loaded with material that did not solubilise in DDM (Lane 'pellet', **Figure 3.7**). However, 10 times as much of the total material

was loaded on this lane compared to the other lanes, indicating that only 10% of MPSIL0294 did not solubilise

Like the MPSIL0294-SBP purification shown previously, some of the His-tagged protein has been lost after washing with the 40 mM imidazole buffer, in this instance however none of the 25 kDa impurity was removed, therefore this wash could have been omitted from the purification procedure. The protein is of a reasonably high quality in regard to its purity after elution with 200 mM imidazole and was subjected to concentration in a 100 kDa vivaspin concentrator. Spectrophotometry was used to estimate the final amount of protein at 0.35 mg from 200 mg of mixed membranes after the final eluate was subjected to dialysis and concentration.

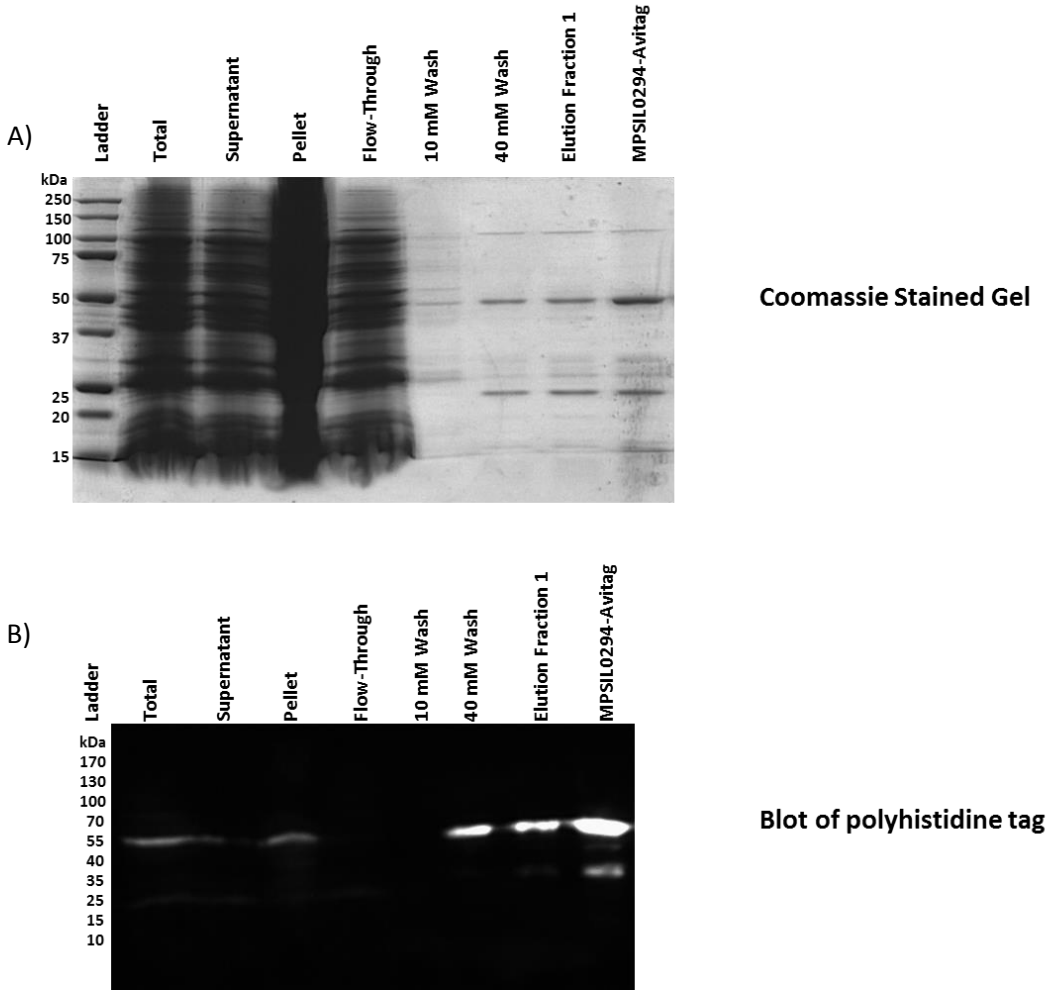


Figure 3.7: <u>IMAC Purification of MPSIL0294-avitag:</u> Samples taken throughout the MPSIL0294-avitag purification were analysed on a 12% SDS-PAGE gel with subsequent staining in Coomassie blue and by Western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP. Each sample represents a different amount relative to the original sample they were taken from; the total, supernatant and flow-through represents 0.01%.

The pellet, 10 and 40 mM washes all represent 0.1% and the elution fraction and final elution represent 1% of the total samples size.

3.3.3 The purification and subsequent labelling of MPSIL0294-V532C

Membranes of MPSIL0294-V532C were prepared by David Sharples using a 30L fermenter via auto-induction. The protein was produced in BL21 star (DE3) and the membranes were isolated using the cell disruptor (**see section 2.4.3**). Once again samples taken throughout the purification protocol provide a clear profile of MPSIL0294-V532Cs behaviour throughout the procedure. These samples were analysed by SDS-PAGE and western blotting probed with anti-histidine antibody conjugated with HRP (in an identical manner as MPSIL0294-SBP and MPSIL0294-avitag) resulting in the data shown in **Figure 3.8**.

Once again the majority of MPSIL0294-V532C was solubilised by 1.5% DDM. As described for the two other MPSIL0294 variants, relative to the total sample, ten times more was loaded for the non-solubilised protein fraction ('pellet') and the second wash step of the affinity column ('40 mM', **Figure 3.8**) and relative small amounts of the protein yield was lost in these steps. However as is clear from **Figure 3.8** a large fraction of the protein was lost due to its failure to bind to the resin and instead was found in the 'flow-through'. The protein produced through elution with 200 mM imidazole has a highly pure.

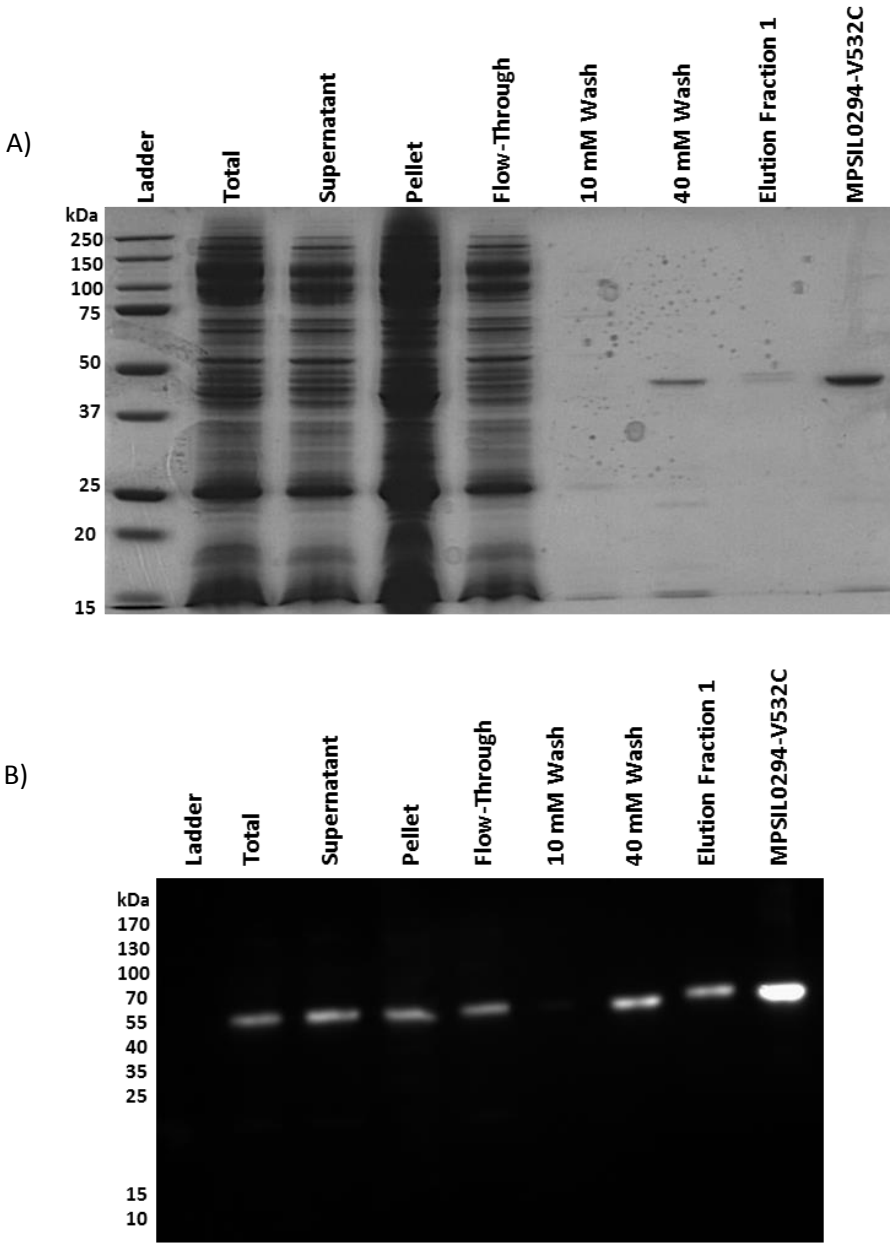


Figure 3.8: <u>IMAC purification of MPSIL0294-V532C:</u> Samples from the purification of MPSIL0294-V532C were analysed by SDS-PAGE and western blotting with anti-polyhistidine tag antibody conjugated with HRP. As with the previous two purifications, samples from the total, supernatant and flow-through represent 0.01% of their respective fractions. The pellet, 10 and 40 mM washes represent 0.1% while the eluted protein fractions and final dialyzed protein represents 1%.

Once purified the MPSIL0294-V532C was biotinylated with EZ-Link® Maleimide-PEG_{11}-Biotin. The labelling efficiency (estimated via the HABA method) (**see section 2.4.6**) was shown to be 0.46 biotin molecules per MPSIL0294 molecule. After spectrophotometric analysis, the final protein yield was determined to be 0.28 mg of MPSIL0294-V532C from 200 mg of mixed membrane.

3.3.4 The purification of AcrB

AcrB was purified according to the procedure described in **section 2.5.2** and the samples taken were analysed in a similar manner as described for MPSIL0294 resulting in the data presented in **Figure 3.9**. AcrB has a high level of expression which is typified by the large protein band which has resolved at approximately 100 kDa in the 'Total' sample (**see Figure 3.9A)**. The vast majority of this protein was successfully solubilised in 1% DDM hence no bands are present in the 'Pellet' sample when it is analysed by western blotting (**see Figure 3.9B**). A faint protein band has been produced on the blot in the 'Flow-through' sample indicating that some AcrB failed to bind to the cobalt resin and was lost. Despite this the protein is very pure once eluted via 200 mM imidazole. Spectrophotometric analysis showed that 0.5 mg of AcrB was purified from 100 mg of mixed membrane. While the final purified protein purer, based upon the blots shown in **figures 3.6B and 3.9B,** the purification of AcrB shows several similarities to that of MPSIL0294-SBP.

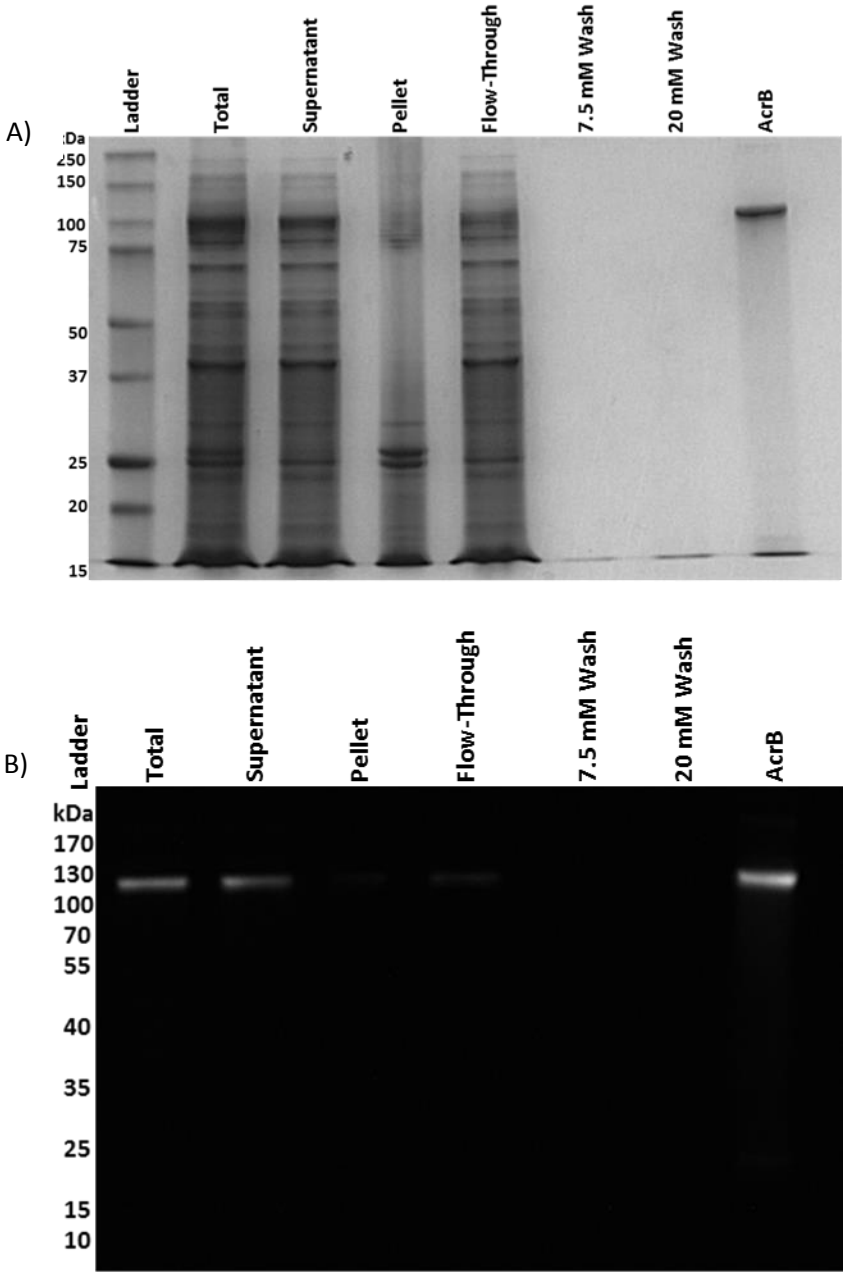


Figure 3.9: <u>IMAC Purification of AcrB:</u> Samples taken throughout the AcrB purification were analysed via 12% SDS-PAGE and western blotting. The blot was treated with an anti-polyhistidine tag antibody conjugated with HRP while the SDS-PAGE gel was stained with Coomassie blue. The samples were normalized by volume to 10 µL therefore the total, supernatant and flow-through represent 0.01% of their original fractions while the pellet and washes represent 0.1%.

3.6 Transport of Zinc by MPSIL0294

The transport of metal ions by MPSIL0294 was tested by reconstituting purified samples into liposomes loaded with the zinc sensitive dye fluozin-1 and measuring the intake of Zn ions via stopped flow analysis as described in **section 2.7.2** (Chao and Fu 2004). The F_{max} (i.e. the maximum amount of fluorescence which can be produced by the dye within the proteoliposomes) was determined with the use of 20 μM of the zinc ionophore pyrithione. The background fluorescence, F_0, was measured by mixing proteoliposomes with assay buffer without any zinc, and subtracted from F (the fluorescence produced by the test sample) and F_{max} to give ΔF and ΔF_{max} of which the ratio ($\Delta F/\Delta F_{max}$) is plotted as a function of time. However the transport assay was limited, no validation of the protein reconstitution was performed and neither the extent with which protein successfully integrated into the liposomes or the orientation of the successfully reconstituted protein were tested.

3.6.1 Transport of Zn^{2+} ions in MPSIL0294-SBP

Transport assays of MPSIL0294-SBP (and other mutants, see below) were also performed in order to prove functionality of this mutant and, by implication, to provide evidence that the protein is properly folded. Assays were carried out using two concentrations of Zn^{2+} ions (31.25 and 3.905 μM). **Figure 3.10** shows that MPSIL0294-SBP is functional as indicated by the rise of fluorescence in the presence of Zn (shown by the red and blue traces). The time trace shows a rapid increase in fluorescence for approximately 5 seconds, after which the signal plateaus and remains relatively constant. The sample mixed with 20 μM pyrithione produced more than twice the amount of fluorescence compared to MPSIL0294-SBP proteoliposomes. The pyrithione is expected to equilibrate the intravesicular compartments of the MPSIL0294-SBP proteoliposomes with zinc. However it is not clear why the samples analysed in the presence of pyrithione produce such a significant increase in fluorescence relative to those treated with zinc alone. This increase may be a consequence of empty liposomes in the proteoliposomal fraction which are able to produce fluorescence in the presence of the ionophore.

The second zinc concentration (3.905 μM) which was tested produced a peak fluorescence reading of 0.290 almost instantaneously and is approximately 1.3 times smaller than the peak fluorescence produced in the presence of 31.25 μM Zinc. Finally F_0 (proteoliposomes mixed with assay buffer in which no zinc was included) produces approximately the same level of fluorescence as the negative control as expected.

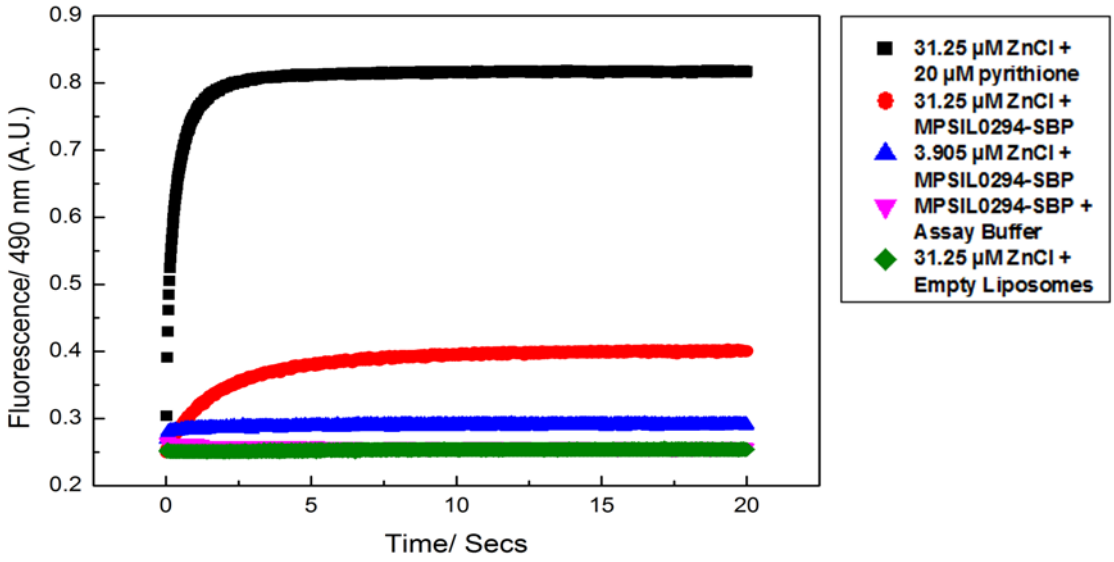


Figure 3.10 Zinc uptake assay with MPSIL0294-SBP proteoliposomes: Samples of purified MPSIL0294-SBP were reconstituted in liposomes loaded with the Zn^{2+} sensitive dye fluozin-1 and tested against two concentrations of ZnCl in order to assay the transport activity of MPSIL0294 using stopped flow analysis. The readings which have been plotted as result of mixing proteoliposomes with assay buffer containing 31.25 µM (red) and 3.905 µM (blue) Zinc are a result of dividing $\Delta F/\Delta F_{max}$. F_{max} of MPSIL0294-SBP was identified by the use of 20 µM of the zinc ionophore, pyrithione. The resultant F_{max} reading for MPSIL0294-SBP is shown in black and represents the maximum possible fluorescence produced by the proteoliposomes. F_0 is also shown in pink and was identified by mixing proteoliposomes with an assay buffer lacking Zn. A negative control was also carried out using empty liposomes mixed with 31.25 µM ZnCl (green), once again the readings on the graph represent the $\Delta F/\Delta F_{max}$ for the empty liposomes.

3.6.2 Transport of Zn^{2+} ions in MPSIL0294-Avitag

The transport assay was performed in an identical manner as described for MPSIL0294-avitag which undergone *in vivo* biotinylation in the *E.coli* strain AVB101 (see **section 2.5.1.2**). Analysis of the fluorescence produced by fluozin-1 loaded liposomes containing MPSIL0294-Avitag were analysed via stopped flow as before and were plotted as a function of time. The resultant graph shown in **Figure 3.11** shows that the MPSIL0294-Avitag protein is not able to transport Zinc ions. It is unlikely that this is due to misfolding as the protein exhibited similar behaviour as the other two MPSIL0294 constructs during their purification (see **section 3.3**). While the SBP and avitag are quite different from one another, in regard to the method by which they bind streptavidin and their length (the avitag is 15 residues while the SBP tag is 38) as discussed below, the reason for this constructs lack of function is not understood at this moment.

There also appears to be an initial 'spike' in fluorescence in the proteoliposomes, the reason for this is not clear and experimental data has shown that it occurs in an unpredictable manner. Therefore I hypothesis that it is indicative of particularly leaky vesicles which either leak Zn^{2+} ions (once transported) or the dye itself causing an instant peak of fluorescence followed by immediate quenching.

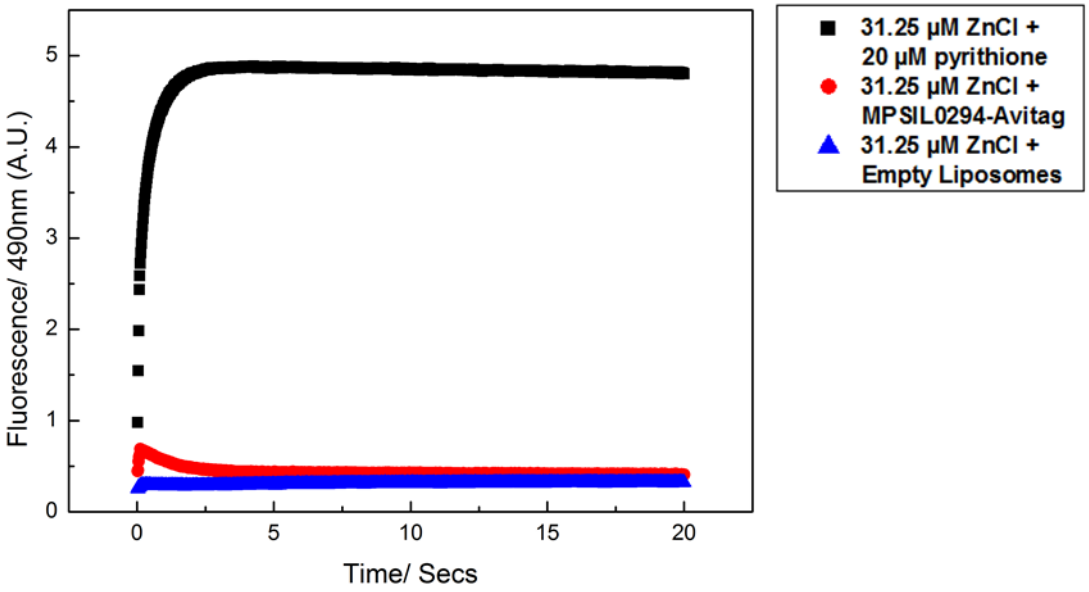


Figure 3.11 Zinc uptake assay with MPSIL0294-avitag proteoliposomes: The stopped flow analysis of samples of MPSIL0294-Avitag which had been subjected to in vivo biotinylation prior to the assay. The protein was reconstituted into liposomes loaded with the zinc sensitive dye flouzin-1 and mixed with an assay buffer supplemented with 31.25 µM ZnCl. The readings shown on the graph represent $\Delta F/\Delta F_{max}$ (red), with the maximum amount of fluorescence(black) analysed with the use of pyrithione(a zinc ionophore). The same transport assay was carried out on empty liposomes which acted as a negative control (blue).

3.6.3 Transport of Zn^{2+} ion in MPSIL0294-V532C

The zinc uptake assay described above was also carried out on samples of MPSIL0294-V532C which had been biotinylated chemically with biotin maleimide. The amount of fluorescence produced by the proteoliposomes loaded with fluozin-1 (see **Figure 3.12**), were recorded by stopped flow and plotted as a function of time. As was the case with the MPSIL0294-avitag plot, an initial sharp 'spike' of fluorescence is observed in the sample when mixed with 31.25 µM ZnCl. The MPSIL0294-V532C trace shows an initial surge similar to that displayed by MPSIL0294-avitag. The major difference between the two is in their respective F_{max} traces. While MPSIL0294-V532C peaks at approximately the same point as the $F_{max,}$ MPSIL0294-avitag produces a very large trace which plateaus just below 5 A.U. While it is not clear what is

causing this initial spike, it could be due to a rapid influx of Zn ions which saturate the encapsulated dye, therefore producing fluorescence, which is then rapidly quenched, possibly due to leaky vesicles allowing Zn^{2+} to leave once again, until the florescence remains constant. The fact that the level of florescence remains approximately 0.75 times higher than the negative control, suggests that the MPSIL0294-V532C is active, however a more reliable transport assay would be required to confirm this.

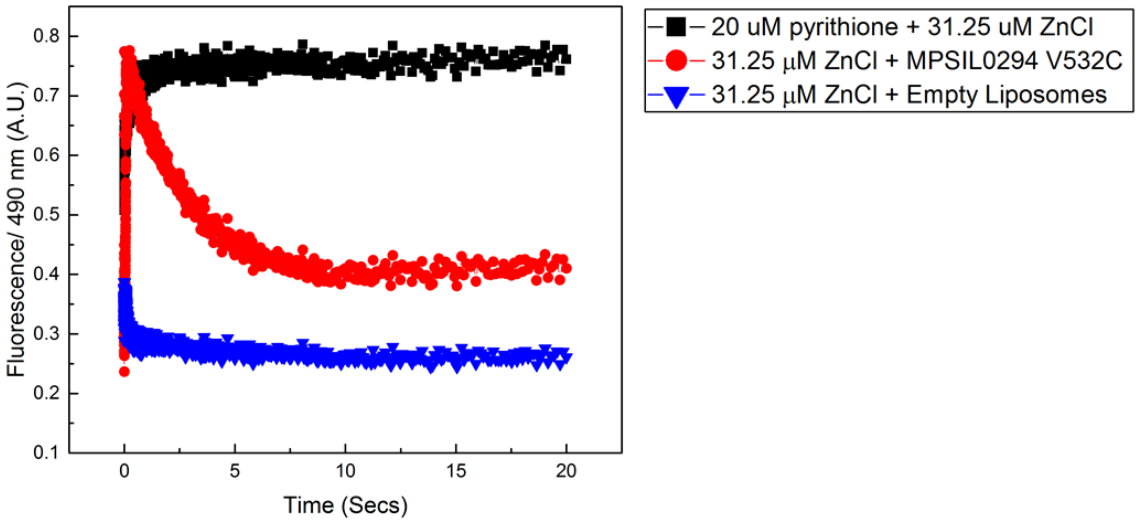


Figure 3.12 Zinc uptake assay with MPSIL0294-V532C proteoliposomes: Stopped flow analysis of samples of MPSIL0294-V532C which had undergone chemical biotinylation of its accessible cysteine with biotin maleimide prior to the assay. The protein was reconstituted into liposomes loaded with flouzin-1 and mixed with an assay buffer supplemented with 31.25 µM ZnCl. The readings shown in red on the graph represent $\Delta F/\Delta F_{max}$. The maximum amount of fluorescence is shown in black and was analysed with the use of 20 µM pyrithione (a zinc ionophore). The same transport assay was carried out on empty liposomes which acted as a negative control and are shown in blue.

3.6 Quality control of phage display samples

3.6.1 Indirect ELISA of MPSIL0294 variants on neutravidin coated plates

In order to show that the various MPSIL0294 variants could bind a solid streptavidin coated surface no matter the method of immobilisation employed, the various forms of MPSIL0294 were subjected to ELISA probed with an anti-polyhistidine tag antibody as described in **section 2.6.2.** These ELISAs involved immobilising the various proteins onto the surface of manually produced neutravidin coated plates thereby providing a useful analogue for the subsequent phage display procedure, except here we use neutravidin instead of streptavidin. Due to their structural similarities, neutravidin is often used interchangeably with streptavidin (the difference between the two being the removal of streptavidin glycosylation sites). Especially as they both have very similar affinities for biotin and isoelectric points. While it was highly probable that MPSIL0294-avitag and MPSIL0294-V532C would successfully bind to the

neutravidin surface (as they are both biotinylated). It was unclear if MPSIL0294-SBP would also be able to do so, however due to neutravidin and streptavidin's extreme functional similarity the SBP/neutravidin interaction was presumed.

The ELISA results for each of the three MPSIL0294 formats are represented graphically in **Figure 3.13** and show that while each of the protein formats follow a similar trend when subjected to this test they have different magnitudes. MPSIL0294-V532C produces the highest A_{450nm} trace and the highest reading in the presence of 50 µg/mL antigen. Intriguingly both the chemically biotinylated MPSIL0294-V532C and the *in vivo* biotinylated MPSIL0294-avitag appear to suffer from a hook affect at the highest concentration (100 µg/mL) which would suggest a very high concentration of analyte on the surface. While this was deemed sufficient for this study due to time constraints, the success of these two proteins could be tested by repeating this ELISA with non-biotinylated samples. This effect doesn't appear in the trace for MPSIL0294-SBP, possibly due to the fact that neutravidin was used as opposed to streptavidin. Nevertheless it appears as if each of the three detergent solubilised variants saturate the neutravidin coated plate at 50 µg/mL and is therefore suitable for phage display screening.

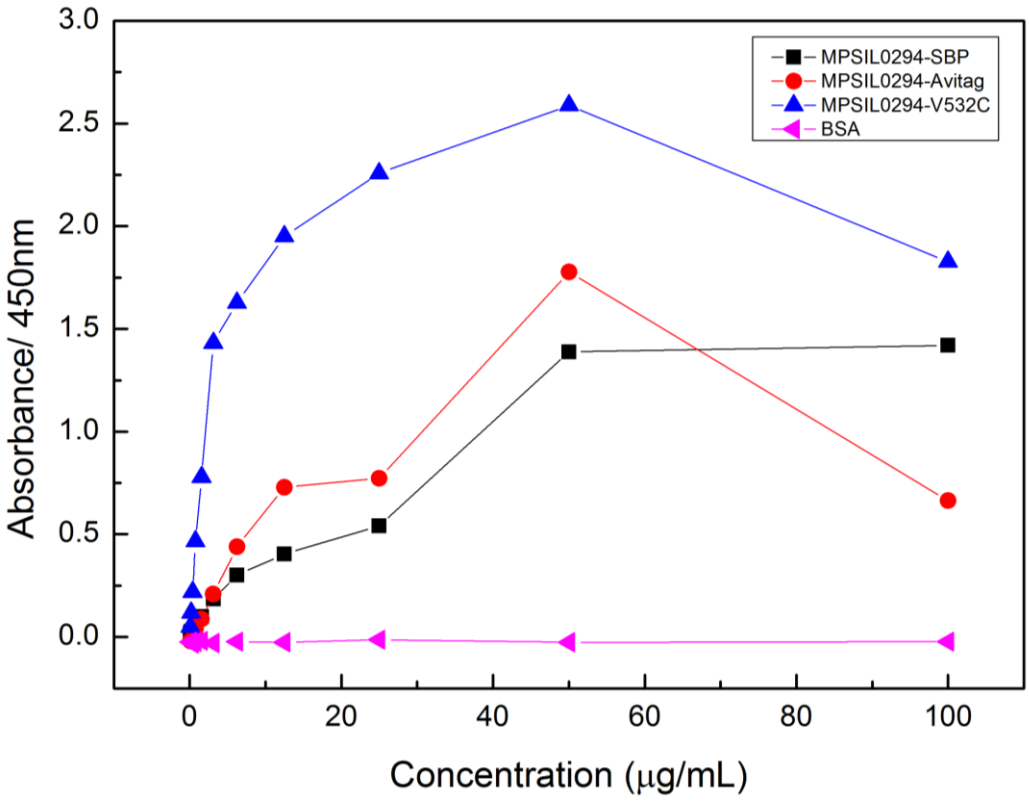


Figure 3.13 Quality control ELISA with detergent solubilised samples of the three MPSIL0294 variants: Each different MPSIL0294 variant was immobilised on a neutravidin coated plate in a serial dilution from 100 to 0.098

µg/mL and subjected to ELISA probed with an anti-polyhistidine tag antibody conjugated with HRP. Each concentration was performed in duplicate and the absorbance readings were measured spectrophotometrically after treatment with TMB stop solution followed by 0.5M H_2SO_4.

3.6.2 Analysis of detergent solubilised MPSIL0294 variants by Fluorescence-activated cell sorting

As a secondary method of assessing the three different MPSIL0294 variants ability to immobilize on a streptavidin coated surface they were each subjected to FACs analysis as described in **section 2.6.3** using streptavidin coated magnetic Dynabeads® (Thermo scientific). A mouse anti-polyhistidine tag antibody was used as a primary while a goat anti-mouse antibody conjugated with DyLight 488 was used as a secondary for this analysis. The overlay histograms shown in **Figure 3.14** show that there is distinctively more intense fluorescence on the Dynabeads® incubated with the MPSIL0294 variants compared with the negative control (which were blank beads). This increase in fluorescence represent an increase in secondary antibody and thus an increase in MPSIL0294 material immobilised on the Dynabeads®.

It is therefore clear that all three of the detergent solubilised MPSIL0294 are successfully able to immobilise on the surface of the streptavidin coated beads, while maintaining the accessibility of the polyhistidine tag to the antibody as they all show an increase in fluorescence when compared with the negative control (the blank beads). Two of the detergent solubilised MPSIL0294 variants performed particularly well in this assay (in regard to their ability to immobilise on beads), namely MPSIL0294-SBP and MPSIL0294-avitag as they produce an identical level of fluorescence that almost matches the positive control (a biotinylated, his-tagged protein extensively worked on by MedImmune which shall be referred to as 'protein x' from here on). 'Protein x' was bound to the beads at the same molar concentration as the MPSIL0294 variants, however it is considerably smaller, therefore it is possible that the increase in fluorescence which 'protein x' exhibits in the FACs analysis is due to an increased amount of packing on the surface of the beads.

The overlay histogram in **Figure 3.14** shows that the MPSIL0294-SBP peak is higher than the MPSIL0294-avitag, thus suggesting that it was better able to pack on the beads; although it appears that both are almost able to saturate the beads in a similar manner as the positive control. The Dynabeads modified with this variant exhibit a more homogeneous fluorescence profile with less of the Dynabeads exhibiting a lower fluorescence compared to MPSIL0294-avitag. MSPIL0294-V532C shows the least binding to streptavidin in this particular assay, but still shows that the protein is successfully able to immobilise onto the beads.

The three MPSIL0294 variants were immobilised onto the beads in the presence of 0.05% DDM as properly folded protein immobilised on the surface of the plate was required in order to isolate confirmation specific DARPins during phage display. However there is no evidence provided in this study that the protein, once immobilised, is in its native confirmation.

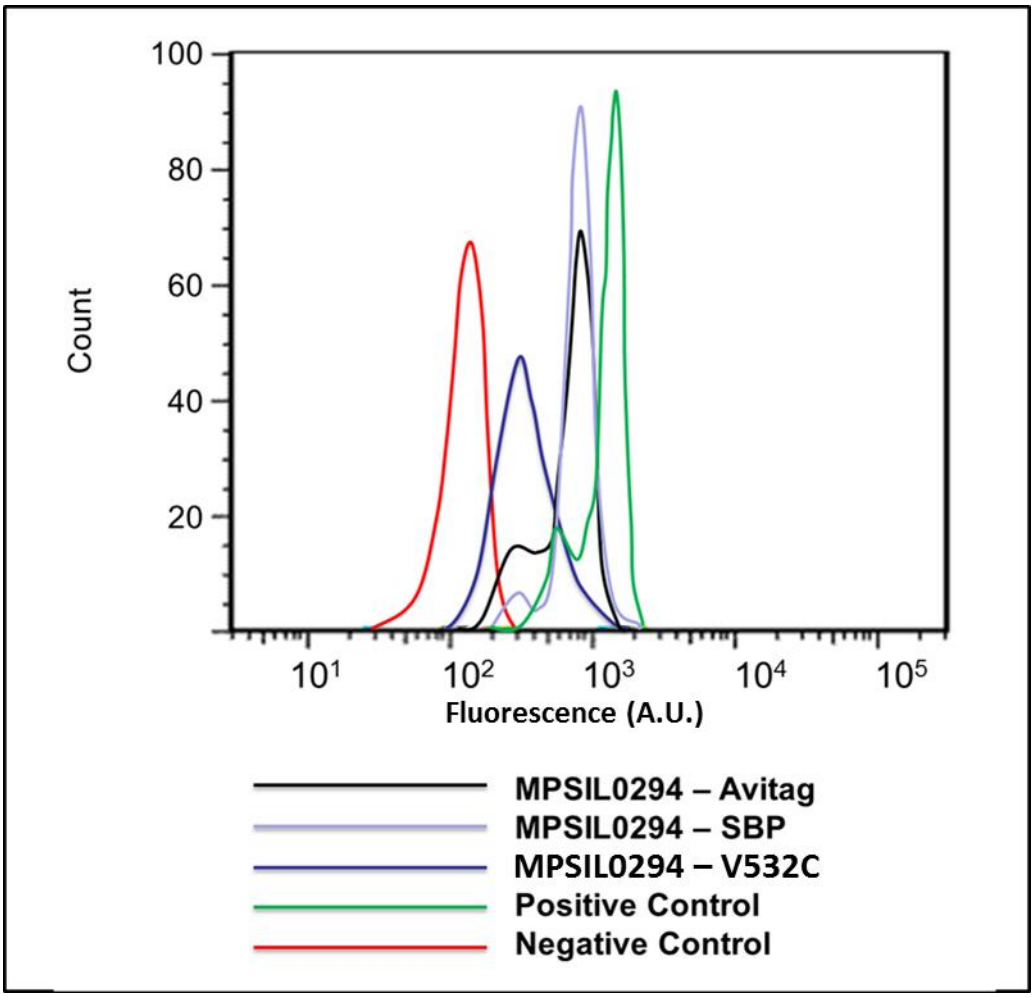


Figure 3.14: <u>FACs analysis of detergent solubilised samples of the three MPSIL0294 variants with streptavidin coated beads:</u> FACs analysis was performed on each of the detergent solubilised variants of MPSIL0294 mixed with magnetic streptavidin coated DynaBeads® (Thermo scientific). The primary antibody was against the polyhistidine tag while a goat anti mouse antibody conjugated with DyLight 488 was used as the secondary. A biotinylated, his-tagged 'protein x' was the positive control. Blank streptavidin coated magnetic beads were used as the negative control.

3.7 A Comparison of different methods of membrane protein immobilisation using phage display

Each of the samples was subjected to two rounds of phage display with a functional DARPin library of 7.5 x 10^9 (data provided by MedImmune) as described in **section 2.8.** In an attempt to retain the tertiary structure of the target protein during the phage display panning, selection was performed in the presence of 0.05% DDM throughout (**see section 2.8**) instead of the more commonly used Tween20. In chapter 4, experiments are described where different methods of membrane protein 'solubilisation' are compared. The experiments described here compared alternative methods of membrane protein immobilisation (and to obtain a high affinity DARPin binder). As such, ideally a minimum of three rounds of phage display would typically be employed, however due to time constraints only two rounds of panning could be performed, thereby increasing the risk of background noise in the selection output.

As previously mentioned (**see section 2.8.3**) the input titre for a phage selection typically should fall between 5 x 10^{11} – 10^{13} while the output should be 10^3 – 10^6. **Table 3.1** shows that all of the titres calculated for the detergent solubilised samples fall well within this remit, with MPSIL0294-avitag having the highest input titre (5 x 10^{12}) while AcrB has the lowest output titre with 6.5 x 10^5.

The selection diversity was calculated by aligning the DARPin amino acid sequences after the selection was complete and the 88 DARPins were sequenced. Any repeats within the selection output were removed and the remaining number of unique sequences divided by 88 to give the diversity in percentages). The sequence diversity can be used as an early indicator of the quality of the selection output. It is generally desirable for the diversity to be moderately high, as a very low percentage may be an indicator of sequence bias within the output, a common sign that the majority of DARPins are binding directly to the streptavidin. Conversely, a sequence diversity which is too high may be an indicator of a lack of selectivity in the output. The highest sequence diversity of the detergent solubilised samples has been attained by MPSIL0294-SBP, while the output with the lowest diversity is the *in vivo* biotinylated MPSIL0294-Avitag (**see Table 3.1**). Despite this, all of the samples fall within an acceptable range.

Sample	Method of immobilisation	Input Titre (cfu/mL)	Output Titre (cfu/mL)	Sequence Diversity
MPSIL0294-SBP	SBP binding	3.6×10^{12}	1.315×10^{6}	86%
MPSIL0294-Avitag	*In vivo* biotinylation	5×10^{12}	8.5×10^{5}	72%
MPSIL0294-V532C	Biotinylation via Cys residue	2.6×10^{12}	1.145×10^{6}	81%
AcrB	Lysine biotinylation	4×10^{12}	6.5×10^{5}	77%

Table 3.1 An overview of the titres produced after phage display carried out on the detergent solubilised members of the membrane protein panel: All samples were solubilised in DDM and were immobilised within the resultant micelles. They subsequently subjected to two rounds of phage display against a DARPin library. All of the selections were also carried out in the presence of DDM and at 4°C. The input titre was calculated after round one of phage display while the output titre was calculated after round two.

3.7.1 Removing streptavidin binding DARPins via Phage ELISAs on blank streptavidin plates

3.7.1.1 The Three Detergent solubilized MPSIL0294 variants

During the selection the phage were deselected against streptavidin in a single deselection step using streptavidin coated beads, however as previously stated two rounds of selection carries a high risk of background noise in selection outputs. Therefore phage ELISAs against streptavidin were carried out on the selection outputs of the detergent solubilised MPSIL0294 as described in **section 2.9.2** in order to exclude any DARPins that bind directly to streptavidin. A threshold A_{450nm} value of 0.1 was used in order to reduce the large amount of data to a more manageable level. Any DARPin which crossed this threshold was deemed capable of binding to directly to streptavidin to some extent and was no longer studied. **Table 3.2** shows that between the three methods of immobilising MPSIL0294 there is a vast difference in the number of DARPins which appear to cross the threshold. The SBP tagged MPSIL0294 only has 19 sequences which did not produce a signal in the presence of streptavidin thereby making it the worst of the three as the avitagged protein produced 85 potential binders while V532C produced 67.

Interestingly within the DARPins capable of binding to streptavidin, V532C was the only version of MPSIL0294 which produced DARPins capable of binding streptavidin strongly (i.e. DARPins which produced an A_{450nm} value above 1.0). In total 14 of the 21 DARPins which were rejected at this point showed a strong ability to bind, this is a characteristic unique to the V532C output as neither the SBP or the avitagged protein produced any strong streptavidin binders. In regard to the other two detergent solubilised MPSIL0294 samples, the 3 rejected DARPins from the avitagged protein are actually relatively close to the threshold and thus are considered weak streptavidin binders. The SBP tagged protein on the other hand mainly produced DARPins with a moderate streptavidin binding ability with A_{450nm} values ranging between 0.5 and 0.7 (**see Table 3.2**).

MPSIL0294-SBP												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.188	0.189	0.381	0.438	0.071	0.059	0.081	0.728	0.394	0.669	0.354	0.075
B	0.554	0.222	0.367	0.694	0.454	0.406	0.338	0.244	0.214	0.876	0.255	0.051
C	0.046	0.13	0.728	0.16	0.087	0.191	0.514	0.118	0.477	0.684	0.432	0.025
D	0.1	0.081	0.22	0.29	0.637	0.325	0.433	0.278	0.439	0.587	0.462	0.057
E	0.22	0.147	0.08	0.133	0.109	0.447	0.36	0.408	0.152	0.598	0.478	0.051
F	0.059	0.268	0.107	0.227	0.148	0.304	0.535	0.455	0.149	0.74	0.53	0.041
G	0.047	0.166	0.071	0.161	0.198	0.089	0.095	0.054	0.232	0.142	0.411	0.048
H	0.057	0.094	0.085	0.086	0.274	0.305	0.135	0.129	0.439	0.065	0.076	0.062

MPSIL0294-Avitag												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.08	0.073	0.082	0.064	0.055	0.082	0.064	0.055	0.067	0.05	0.063	0.082
B	0.072	0.053	0.054	0.09	0.047	0.051	0.073	0.054	0.046	0.053	0.053	0.049
C	0.057	0.06	0.053	0.048	0.068	0.047	0.064	0.05	0.058	0.055	0.051	0.056
D	0.052	0.059	0.08	0.057	0.052	0.061	0.061	0.07	0.053	0.05	0.057	0.044
E	0.044	0.062	0.058	0.065	0.059	0.062	0.062	0.172	0.287	0.049	0.068	0.056
F	0.057	0.078	0.061	0.06	0.061	0.065	0.063	0.048	0.053	0.056	0.048	0.044
G	0.058	0.068	0.292	0.06	0.062	0.061	0.055	0.066	0.054	0.051	0.054	0.054
H	0.06	0.051	0.057	0.051	0.054	0.049	0.054	0.051	0.07	0.054	0.05	0.042

MPSIL0294-V532C												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.044	0.068	0.087	0.076	0.072	0.081	0.061	0.071	0.062	0.075	1.11	0.056
B	0.055	0.06	0.31	0.08	0.088	0.09	0.073	0.078	0.063	0.071	1.415	0.046
C	0.041	0.055	0.096	1.12	0.07	0.077	1.411	0.226	0.077	1.487	0.098	0.052
D	0.047	0.068	0.08	0.095	0.073	0.059	0.083	0.067	0.083	0.094	0.088	0.061
E	0.075	1.281	0.104	0.094	0.441	0.069	0.064	0.067	0.066	1.322	0.075	0.057
F	1.025	0.086	1.235	1.484	1.279	0.061	0.055	0.069	0.067	0.089	0.073	0.044
G	0.062	0.057	1.141	0.111	0.113	0.065	0.52	0.06	0.058	0.076	0.072	0.05
H	0.052	0.052	0.309	1.174	1.355	0.048	0.047	0.049	0.048	0.049	0.048	0.041

Table 3.2 Quality control of the selection outputs generated against the three MPSIL0294-variants performed on blank streptavidin coated plates: The data produced after the selection outputs from each detergent solubilised form of MPSIL0294 is analysed via phage ELISA on blank streptavidin plates with an M13 antibody conjugated with HRP followed by treatment with TMB solution and 0.5M H_2SO_4. All DARPins which do not show an ability to bind directly to streptavidin are highlighted in green, weak streptavidin binders are highlighted in yellow while DARPins which bind streptavidin moderately are highlighted in orange. Finally, DARPins which strongly bind to streptavidin are highlighted in red and the negative control values are written in blue text.

3.7.1.2 Detergent solubilized AcrB

Detergent solubilised AcrB protein was immobilised via chemical biotinylation with NHS-Biotin (**see section 2.4.4**) and subjected to two rounds of phage display as previously described. The selection output was then analysed so that any DARPins with a capability to bind directly to streptavidin could be identified via phage ELISA carried out in an identical manner as described previously (**see section 2.9.2**). **Table 3.3** shows the results of the AcrB phage ELISA and resulted in 33 DARPins which did not produce an A_{450nm} reading above 0.1 in the presence of a blank streptavidin plate and were therefore taken for further study. Of the 55 DARPins which did show an ability to bind, one produced a signal over 1.0 while the rest were moderate in a similar vein as MSIL0294-SBP (**see Table 3.2**).

Detergent solubilised AcrB												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	0.096	0.059	0.379	0.097	0.071	0.087	0.063	0.062	0.28	1.985	0.171	0.058
B	0.076	0.067	0.334	0.13	0.709	0.272	0.081	0.372	0.076	0.109	0.073	0.045
C	0.121	0.097	0.269	0.501	0.273	0.182	0.433	0.181	0.058	0.161	0.207	0.052
D	0.11	0.137	0.241	0.054	0.119	0.12	0.49	0.09	0.104	0.431	0.148	0.096
E	0.048	0.061	0.123	0.183	0.35	0.937	0.395	0.071	0.149	0.201	0.537	0.065
F	0.054	0.519	0.052	0.098	0.114	0.74	0.476	0.097	0.142	0.093	0.641	0.047
G	0.056	0.391	0.064	0.075	0.357	0.132	0.28	0.33	0.073	0.13	0.064	0.056
H	0.063	0.052	0.073	0.07	0.319	0.079	0.231	0.053	0.135	0.27	0.073	0.046

Table 3.3 Quality control phage ELISA of selection output generated against detergent solubilised AcrB: The data produced after a phage ELISA on a blank streptavidin plate using the selection output isolated against AcrB solubilised and presented within DDM micelles. The plates were probed with an anti M13-HRP antibody conjugate and visualised with TMB solution followed incubation with 0.5 M H_2SO_4. The absorbance of the plate was then analysed at 450 nm and all DARPins with a reading over 0.1 were no longer studied. DARPins highlighted in green were taken on for further validation; those which bind streptavidin weakly are highlighted in yellow. DARPins which bind streptavidin moderately are highlighted in orange while those that bind strongly are highlighted in res. Finally, A_{450nm} readings in blue were the negative control in which no phage was added.

3.7.2 Refining selection outputs via Phage ELISA in the presence of target protein

3.7.2.1 The three detergent solubilized MPSIL0294 variants

A phage ELISA was performed the phage identified in **Section 3.7.1**. These ELISAs were done against the same MPSIL0294 variant used during the screening (panning) procedure in order to

validate the selection outputs against each target and confirm the DARPins ability to bind them. In this repeat 100 μg of target protein was first immobilised onto the streptavidin coated surface and tested against the DARPins which did not produce a signal in the previous phage ELISAs (**see Table 3.2**) as described in **section 2.9.3**. In order to attain a more manageable amount of data, a threshold A_{450nm} value of 1.0 was selected for these phage ELISAs. The resultant 96 well plates shown in **Table 3.4** shows that the vast majority of DARPins failed to produce a strong signal in the presence of the various detergent solubilised MPSIL0294 variants. The majority of the phage isolated against MPSIL0294-avitag (66 out of the 85 DARPins of interest) produced a signal between 0.1 and 0.5 which could be considered very weak binders. This pattern is unique to this variant as MPSIL0294-V532C and MPSIL0294-SBP only produced 6 and 4 DARPins respectively which were capable of producing a similar signal. However MPSIL0294-V532C was able to produce two DARPins which appeared to strongly bind the target at positions E6 and G6 respectively (see **Table 3.4,** MPSIL0294-V532C).

Detergent solubilised MPSIL0294-SBP												
	1	2	3	4	5	6	7	8	9	10	11	12
A					0.083	0.091	0.096					0.067
B												0.054
C	0.089				0.082							0.064
D		0.107										0.059
E			0.078									0.067
F	0.125											0.061
G	0.072		0.076			0.069	0.081	0.071				0.076
H	0.1	0.094	0.095	0.1						0.109	0.095	0.058

Detergent solubilised MPSIL0294-Avitag												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.089	0.108	0.159	0.144	0.103	0.112	0.163	0.13	0.092	0.126	0.115	0.099
B	0.119	0.153	0.217	0.151	0.183	0.136	0.199	0.218	0.135	0.184	0.148	0.093
C	0.089	0.106	0.125	0.11	0.099	0.096	0.116	0.157	0.152	0.125	0.159	0.079
D	0.095	0.133	0.145	0.111	0.145	0.131	0.166	0.175	0.149	0.143	0.145	0.061
E	0.067	0.1	0.124	0.092	0.114	0.136	0.108			0.125	0.112	0.074
F	0.103	0.114	0.131	0.144	0.139	0.133	0.143	0.138	0.152	0.153	0.147	0.069
G	0.066	0.096		0.104	0.078	0.083	0.098	0.094	0.075	0.078	0.093	0.072
H	0.107	0.107	0.101	0.105	0.1	0.101	0.105	0.106	0.106	0.114	0.098	0.062

Detergent solubilised MPSIL0294-V532C												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.058	0.087	0.068	0.059	0.062	0.065	0.064	0.068	0.045	0.063		0.065
B	0.089	0.095		0.082	0.079	0.084	0.082	0.072	0.055	0.086		0.069
C	0.059	0.083	0.065		0.059	0.07			0.043		0.207	0.068
D	0.089	0.104	0.106	0.091	0.088	0.095	0.085	0.092	0.054	0.09	0.097	0.071
E	0.062			0.052		1.233	0.177	0.062	0.042		0.081	0.065
F		0.091				0.089	0.087	0.08	0.05	0.08	0.114	0.066
G	0.054	0.068			0.063	1.31		0.06	0.045	0.115	0.071	0.065
H	0.092	0.079				0.091	0.079	0.075	0.052	0.074	0.084	0.056

Table 3.4 <u>The Secondary quality control phage ELISA against the three detergent solubilised MPSIL0294 variants selection outputs:</u> Phage ELISAs were performed on streptavidin coated plates on which 100 µg of the three detergent solubilised MPSIL0294 variants had been immobilised. The ELISAs were probed with anti- M13 antibody conjugated with HRP, they were visualised by incubation with TMB followed by treatment with 0.5 M H_2SO_4 and analysis of it absorbance at 450 nm. Values which have been redacted in black denote the DARPins which showed an ability to bind streptavidin directly while green values are DARPins which showed an ability to bind to MPSIL0294. Values highlighted in red represent DARPins which failed to bind to the target while those in yellow represent DARPins considered very weak binders.

3.7.2.2 Detergent solubilized AcrB

The phage ELISAs were also carried out on samples of detergent solubilised AcrB which had been chemically biotinylated. These ELISAs were performed in an identical manner as described above for the detergent solubilised MPSIL0294 variants and were once again designed to further validate the selection output. As shown in **Table 3.5** no DARPins were able to produce a reading which surpassed the 1.0 A_{450nm} threshold. Unlike MPSIL0294 none of the tested DARPins (of the 39 DARPins which did not show any activity on blank streptavidin coated plates) were able to cross 0.1, with the highest being DARPin B1 with an A_{450nm} of 0.096.

Detergent solubilised AcrB												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	0.061	0.053		0.048	0.045	0.043	0.044			0.036		0.049
B	0.096	0.088					0.071		0.067	0.049	0.07	0.04
C		0.053							0.045	0.037		0.051
D								0.069		0.05		0.044
E	0.061	0.062						0.045		0.037		0.054
F	0.083		0.082					0.069				0.047
G	0.049		0.051	0.051					0.052	0.038	0.048	0.048
H	0.075	0.068	0.075	0.078		0.075		0.074		0.056	0.074	0.042

Table 3.5 The secondary quality control phage ELISA of the detergent solubilised AcrB selection output: The data produced by the phage ELISA performed on streptavidin coated plates on which 100 µg AcrB in DDM micelles which had been immobilised via chemical biotinylation. The plates were probed with anti-M13 antibody conjugated with HRP and visualised using TMB. Incubation with H_2SO_4 stopped this reaction and allowed for its absorbance at 450 nm to be analysed. Black squares have been used to represent the DARPins which were removed due to their ability to bind directly to streptavidin in the previous phage ELISAs.

3.7.3 An overview of phage display on different methods of membrane protein immobilisation

Each of the immobilised target proteins was subjected to the identical regimen of procedures in order to isolate DARPins capable of binding and subsequently validate them. These procedures included the initial two rounds of phage display which were performed on a 96 well plate and resulted in the isolation of 88 DARPins (stage 1). Followed by the initial phage ELISAs aimed at identifying DARPins capable of binding directly to streptavidin (stage 2). Thirdly, the phage ELISAs were repeated in the presence of immobilised target protein in order to validate their ability to bind (stage 3) and finally the DARPins with A_{450nm} readings above the

threshold were taken on for further study (stage 4). The outcome of this procedure is outlined in **Table 3.6** and shows that neither the C-terminal avitag which had undergone *in vivo* biotinylation, the C-terminal SBP tag or the aspecifically biotinylated AcrB were able to isolate any DARPins after stage 3. MPSIL0294-V532 on other hand managed to isolate two DARPins (E6 and G6) which are seemingly able to bind the protein and were taken for further study (see **Chapter 5**).

Target	Method of Immobilisation	Method of Solubilisation	Stage 1	Stage 2	Stage 3
MPSIL0294	Avitag	DDM	88	85	0
	SBP	DDM	88	19	0
	Cys/Biotin maleimide	DDM	88	67	2
AcrB	Lysine Biotinylation	DDM	88	33	0

Table 3.6 An Overview of the entire phage display and subsequent validation steps performed on all detergent solubilised target proteins: The detergent solubilised protein samples aimed to compare alternative methods of protein immobilisation via either different tags or aspecific biotinylation. The data has been compiled through the three stages of DARPin isolation: Stage 1 represents the initial two rounds of phage display carried out in 96 well plates, in each instance 88 arbitrary DARPins were isolated prior to sequencing. Stage 2 was the phage ELISAs performed on blank streptavidin plates while Stage 3 was the second round of phage ELISAs performed in the presence of immobilised target protein. The two DARPins which seemingly bound to MPSIL0294-V532C were further characterised in chapter 5. The numbers in each cell of denote the number of DARPins available for the subsequent stage of validation.

3.8 Discussion and conclusion

In this chapter two model membrane proteins were prepared in a variety of different formats for phage display, in order to compare different methods of membrane protein immobilisation and presentation and generate successfully binding DARPins. The different immobilisation methods under test are all based around the interaction between streptavidin and biotin and involve the use of different C-terminal tags on a model membrane protein – MPSIL0294. The tags were placed on its C-terminal because Dr Cheng Ma has previously shown that a C-terminal His tag does not affect its expression or ability to transport metal ions. However when the expression levels of MPSIL0294-CVH produced by Dr. Cheng Ma are compared with those

of MSPIL0294-SBP, avitag and V532C (shown by the 'Total' samples of Coomassie stained SDS-PAGEs in **Figure 3.5, Figure 3.6 and Figure 3.7**) it is clear that there is a sharp decrease, therefore resulting in lower yields of protein (Ma 2013). Also zinc uptake assays appear to suggest that MPSIL0294-avitag is not active. This lack of function is not caused by the polyhistidine tag as the other constructs have this tag and are functional, although a more consistent transport assay would help to clarify this (see **section 3.6**). Despite this, enough protein was obtained for phage display as demonstrated by the fact that each of the MPSIL0294 variants successfully saturated a neutravidin coated plate probed with an anti-polyhistidine tag antibody conjugated with HRP in **Figure 3.13**. This ELISA result alongside the FACs analysis shown in **Figure 3.14** provides evidence that the proteins can be immobilised while retaining their accessibility to a binding probe, which in both of these instances was an antibody against the C-terminal polyhistidine tag.

The purifications for two of the differently tagged forms of MSIL0294 (SBP and the biotinylated avitag) exhibit what appear to be degradation products on the blots shown in **Figure 3.5 and Figure 3.6**. These superfluous bands cannot definitively be identified as MPSIL0294, but the fact that they appear on the blot shows that they must have a polyhistidine tag therefore it is very likely that they are. The fact that the MPSIL0294-V532C does not exhibit the same behaviour suggests that this is not being caused by a protease, as all three MPSIL0294 variants were purified in an identical manner with the exception of 0.5 mM TCEP in all buffers for the MPSIL0294-V532C purification (**see section 2.5.1**). The addition of TCEP prevented the undesired formation of disulphide bridges with endogenous *E.coli* proteins and subsequent co-purification. Potentially degraded protein relative to the full length MPSIL0294 is barely detectable on Coomassie stained SDS-PAGEs.

AcrB was included as a second model membrane protein in the comparison of alternative immobilisation strategies. Like MPSIL0294 it was solubilised in DDM and presented to the DARPin library in the presence of 0.05% DDM (**see section 3.3.4**). As previously mentioned AcrB was chosen due to its large extracellular loop which is thought to form a TolC docking domain after the protein forms the functional trimer. This soluble loop could make the isolation of binding proteins against AcrB easier. A second reason for its inclusion is the fact that there is precedence for the isolation of DARPins against AcrB which were subsequently found to act as an inhibitor to the protein and were used as crystallisation partners (Sennhauser, Amstutz et al. 2007, Eicher, Cha et al. 2012, Du, Wang et al. 2014). Sennhauser et al, isolated DARPins using four rounds ribosome display on chemically biotinylated detergent

solubilised samples presented to the library in 0.03% DDM. The study presented here followed a very similar outline however despite this, no DARPins showed an ability to bind to the detergent solubilised AcrB. This inability to isolate DARPins is probably due to the fact that only two cycles of phage display were utilised during the selection and only 88 DARPins were selected arbitrarily as a result of it. Two rounds of phage display will suffer from a high level of background in the results, however the selection was designed in this manner because this study aims to compare immobilisation strategies rather than specifically isolate DARPins to particular targets. Therefore an increase in the number of cycles and the number of phage particles selected would most likely alleviate this problem. It must also be noted that due to time constraints, no validation was performed that looked at the ability of biotinylated AcrB to immobilise on a streptavidin coated surface however. Therefore, the lack of successfully binding DARPins to detergent solubilised AcrB due to its poor immobilisation on the surface cannot be ruled out. ELISAs should be performed in which streptavidin plates are coated with a titration of serially diluted biotinylated AcrB probed with an anti-His-HRP antibody conjugate.

The FACs data in **Figure 3.14** suggests that when compared with the other two, less MPSIL0294-V532C is able to immobilise to a streptavidin coated surface. However, the ELISA data shown in **Figure 3.13** does not support this notion as it appears as though the MPSIL0294-V532C is most able to saturate the plate. . It may be that while the MPSIL0294-avitag and MPSIL0294-SBP are able to saturate a streptavidin coated surface. On the other hand in the case of MPSIL0294-V532C it may be that its inability to do so may result in a population of DynaBeads® which do not produce fluorescence thus explaining the overlay histogram. However, the saturation of a streptavidin coated surface by MPSIL0294-avitag and MPSIL0294-SBP may decrease the accessibility of the polyhistidine tag antibody conjugated with HRP leading to the decrease in magnitude of the fluorescent traces in the ELISA (see **Figure 3.13**). As a subsequent consequence of MPSIL0294-V532C's decrease in immobilisation, it is possible that more bare streptavidin was exposed to the DARPin library during the phage display selections thus resulting in the increase in streptavidin binders observed in **Table 3.2**. While this is theoretically possible the fact that the ELISA was performed on neutravidin while the FACs and phage display were performed on streptavidin may also account for the differences between the ELISA and the FACs which are observed, particularly with MPSIL0294-SBP.

The validation experiments highlighted several differences between all three MPSIL0294-variants. Firstly MPSIL0294-SBP isolated far more DARPins capable of moderate binding to streptavidin (see **Table 3.1**) when compared with its two counterparts. While MPSIL0294-

V532C was the only variant that isolated DARPins capable of binding to the target after the two validation tests (see **Table 3.2**). This unfortunately is not enough data to draw any meaningful conclusions. Although fundamental differences in the nature of the interaction which results in the immobilisation of the protein may be a factor. For instance, while MPSIL0294-avitag and MPSIL0294-V532C both rely on the biotin/streptavidin interaction, the SBP tag binds directly to streptavidin. Binding is dependent upon a HVV motif in the SBP N-terminal and a HPQ motif in its C-terminal, thus occupying two of the four biotin binding pockets within streptavidin, while several of the central residues form an α-helix (SBP is comprised of 35 residues in total) which do not play a role (Barrette-Ng, Wu et al. 2013). While two SBPs can bind to a single streptavidin molecule, four biotins are able to bind with a single molecule occupying on of the aforementioned biotin binding pockets. This means that the MPSIL0294 variants differ from one another in regard to the density with which they can pack on a solid surface. A decrease in which may in turn be beneficial to the isolation of binding partners via phage display. However it is not clear if all four of these binding pockets are accessible once streptavidin has been immobilised onto a solid surface. Also, it is unlikely that a single streptavidin would be able to accommodate either two SBP tagged, or four biotinylated MPSIL0294 molecules due to steric hindrance, especially due to the large size of the DDM micelle. Nevertheless, if the density with which MPSIL0294 is able to pack on the streptavidin coated surface has an influence on the number of binding partners isolated after phage display, it could be tested by comparing different surface coverages in a comparative phage display screen in which the number of binding proteins isolated are compared.

Alternatively the isolation of the two DARPins by MPSL0294-V532C may be due to the flexibility the protein is afforded by the presence of a PEG_{11} linker separating the maleimide group from the biotin in the biotinylation reagent. This seems unlikely however as the PEG linker separating the protein from the tag is similar in length to the amino acid linker which separates the protein from both the SBP and avitag (PEG_{11} is approximately the same length as an amino acid chain with 11 residues). As mentioned earlier the SBP tag is 35 residues long but immobilises the protein by binding to the streptavidin directly via its terminals (Barrette-Ng, Wu et al. 2013), the linker which separates the SBP from MPSIL0294 is 15 residues. Similarly the avitag is composed of 15 residues and biotinylation occurs at the lysine residue in the 10^{th} position. Once again the linker which separates MPSIL0294 from the avitag is 15 residues long. It is not clear if increasing the amount of flexibility that the linker provides to the target protein will improve the number of DARPins which are isolated by phage display. A potential method to test this would be to label MPSIL0294-V532C with a panel of biotinylation

reagents with alternative length PEG linkers separating the maleimide and biotin, while the amino acid linker separating the MPSIL0294 from the SBP or the avitag is increased. All of these altered variants could then be subsequently subjected to two rounds of phage display as described here.

The amino acid sequences of the two DARPins (E6 and G6) which showed an ability to bind to MPSIL0294 are markedly similar to one another as can be seen in **Figure 3.15.** In several positions within the variable regions for example at positions 33, 64 and 77 of the DARPin either the same amino acid is present, amino acids with a similar side chain or residues with the same physical characteristics. When aligned with a third DARPin (DARPin H2 which produced an A_{450nm} value of 0.079) from the same selection output it appears to confirm the proposition that the chemical nature of the amino acids at position 33, 64 and 77 are important for binding to MSPIL0294.

Figure 3.15 MPSIL0294-V532C DARPin binders sequence alignment: An alignment of the two DARPins which showed an ability to bind detergent solubilised MPSIL0294-V532C when presented in DDM micelles, E6 in the central row and G6 on the bottom. These DARPins have been aligned with a third DARPin H2 from the same selection output which did not show any evidence of binding activity in the presence of MPSIL0294-V532C. The variable regions of the DARPins have been highlighted, residues in red differ from the reference but are not

related to one another, those in yellow are amino acids in the same group in regard to their side chains or their physical characteristics and those in green differ from the reference DARPin but are the same in both E6 and G6. Residues in H2 are highlighted in blue to emphasis the fact that they are not related to the residues in the corresponding positions in the two protein binding DARPins. Residues which are not highlighted are in the conserved region and have been removed in both E6 and G6.

Chapter 4

Alternative methods of membrane protein solubilisation compared using Phage Display against a DARPin library

4.1 Introduction

Membrane proteins are traditionally solubilised using strong surfactants as described in the previous chapter. However detergent solubilisation can result in the adoption of inactive conformations by the membrane protein under test due to its removal from the lipid bilayer (Bowie 2001). This thereby limits the applicability of any successful binding protein, both as a real world therapeutic and as a research tool for the elucidation of the membrane protein structure (see **section 1.4.1**). Furthermore, micelles can hide epitopes, affecting the success of phage display screening on membranous targets because the micelle, depending on the carbon chain of the detergent (longer chain detergents result in larger micelles) can be bigger than the width of a lipid bilayer (Oliver, Lipfert et al. 2013). Larger micelles will alter the accessibility of surface epitopes to potential binding partners due to the hydrophobic mismatch between the protein and the detergent (i.e. the micelle does not properly align with the hydrophobic portion of the protein) thereby lowering the chances of a successful screen and potentially reducing the stability of the target protein (Kunji, Harding et al. 2008).

In this chapter different methods of detergent free solubilisation, namely nanodiscs and styrene maleic acid lipid particles (SMALPs), were compared in regard to their ability to isolate DARPins screened using phage display. A panel of membrane proteins composed of MPSIL0294-SBP and AcrB (both of which were introduced in chapter 3) solubilised in nanodiscs and SMALPs were subjected to two rounds of phage display. Furthermore, two concentrative nucleoside transporters (CNT) were also included the first being NupC – from *E.coli,* and the second being VcCNT - NupCs homologue from *Vibrio cholerae* both of which were prepared in SMALPs. NupC and VcCNT were chosen due to the extensive body of work which has already been done with the aim of characterising them (see Chapter 1). This panel was chosen with the aim of isolating DARPins capable of binding.

4.2 The preparation of membrane protein Nanodiscs

As in chapter 1 nanodiscs are a novel method of membrane protein solubilisation which rely on a membrane scaffold protein (MSP) derived from apolipoprotein 1. There are several different MSPs that are commonly used, each of which differs from one another in their amino acid sequences and length which in turn alters the size of nanodiscs they can produce. All of them, however, are derived from apolipoprotein 1 and form nanodiscs in an identical manner (Bayburt and Sligar 2010). They are typically expressed in *E.coli* and purified in a similar

manner as described in **section 2.5.4** with the help of a C-terminal His tag. The MSP are eluted in a buffer consisting of 40 mM Tris, 0.3 M NaCl, 0.4 M Imidazole, pH 8 and stored at -80°C. For this study the MSP was provided by Emily Caseley from the University of Leeds.

In the case of MPSIL0294-SBP nanodiscs, samples of detergent solubilised MPSIL0294-SBP protein were mixed with the MSP1D1 (the membrane scaffold protein) and POPC (Avanti) in a ratio of 1:2:35, in a manner described in **section 2.4.5.** The pertinent information when deciding upon a molar ratio was the fact that a nanodisc composed of MSP1D1 and POPC is thought to contain 61 - 65 lipid molecules and have an occupancy of 69$Å^2$ (Bayburt and Sligar 2010). Secondly, the area of MPSIL0294 was determined to be 3216.99 $Å^2$; this was estimated by working out the radius of the protein (which itself was estimated using a model built by Jocelyn Baldwin of the university of Leeds of MPSIL0294 based on the crystal structure of the amino acid exchanger ApcT using the software package Pymol from Schrödinger). By dividing MSIL0294s area by the occupancy of a single nanodisc it can be estimated that the proteins insertion will displace 46.6 lipid molecules. Because there are two lipid envelops within a nanodisc this is actually 93.2 lipid molecules; an empty nanodisc has a total of 130 lipids (1 MSP1D1 to 65 POPC). Therefore in a MSIL0294 nanodisc it can be concluded that there should be an estimated 36.8 lipids (130 – 93.2), which was rounded down to 35.

Once all detergent was removed (see **section 2.5.5**) nanodisc samples were analysed via ÄKTA explorer in a Superdex 200 10/300 GL column at a flow rate of 0.5 mL/min. The resultant chromatogram shown in **Figure 4.1** shows a large peak in the void volume of the column, the cause of which is not clear. Several smaller absorbance peaks (280 nm) were visible at an elution volume of 20-25 mL, which suggested the presence of folded protein. Further analyses of these samples however by western blotting or SDS-PAGE followed by silver staining using the Pierce® Silver stain kit (Thermo scientific), showed no protein present (**Figure 4.2**). Samples from the void volume showed a very faint band at approximately 55 kDa which corresponds to an immunogenic band on the western blot which is most likely MPSIL0294. Conversely, on the SDS-PAGE, a clear 25 kDa band can be seen which corresponds to a very faint band of the same size on the western blot most likely the MSP (**Figure 4.2**). While the MPSIL0294 may fail to show up on the SDS-PAGE due to a lack of sensitivity (despite the fact that silver staining was used as opposed to conventional staining with Coomassie), it is not clear why detection of the MSP appears to have failed on the blot. The presence of both bands at their expected masses suggests the presence of protein filled nanodiscs, however nanodiscs typically have two MSPs

per disc therefore the intensity of the bands does not support this. It is therefore likely that the ratio used for the nanodiscs is not ideal.

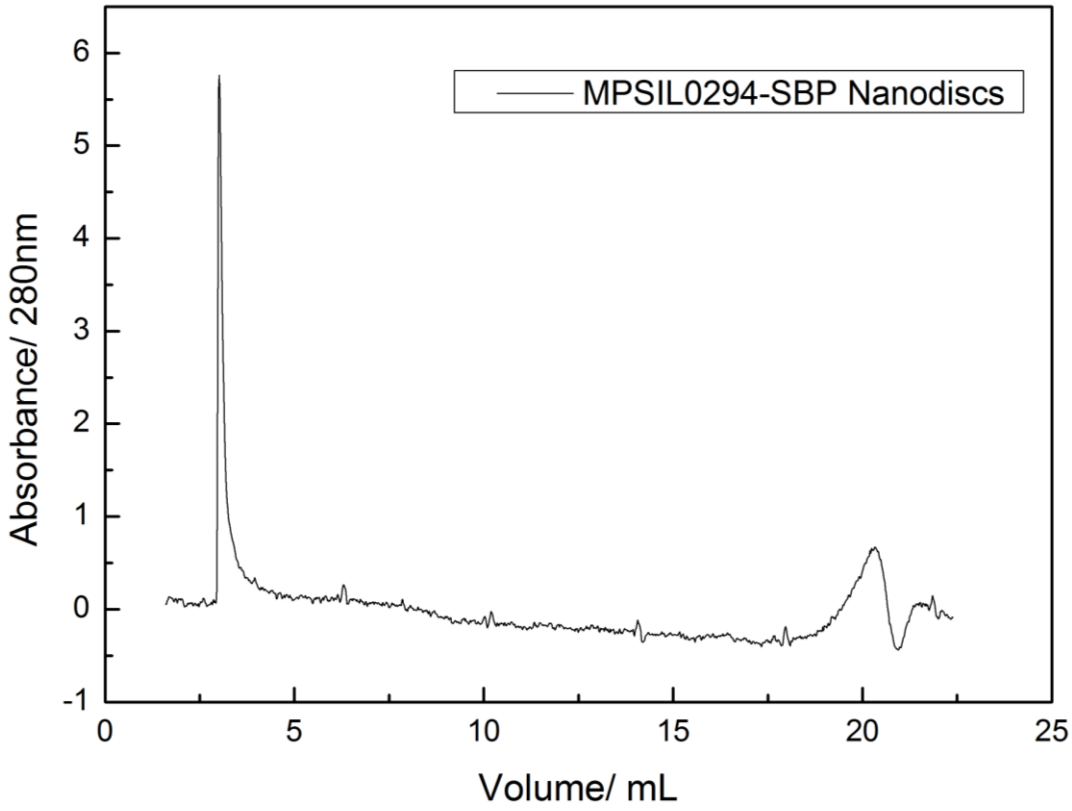


Figure 4.1: <u>Gel filtration chromatogram of MPSIL0294-SBP nanodiscs:</u> Samples of Nanodiscs produced with MPSIL0294-SBP were analysed using a Superdex 200 10/300 GL column at a flow rate of 0.5 mL/min in an AKTA explorer.

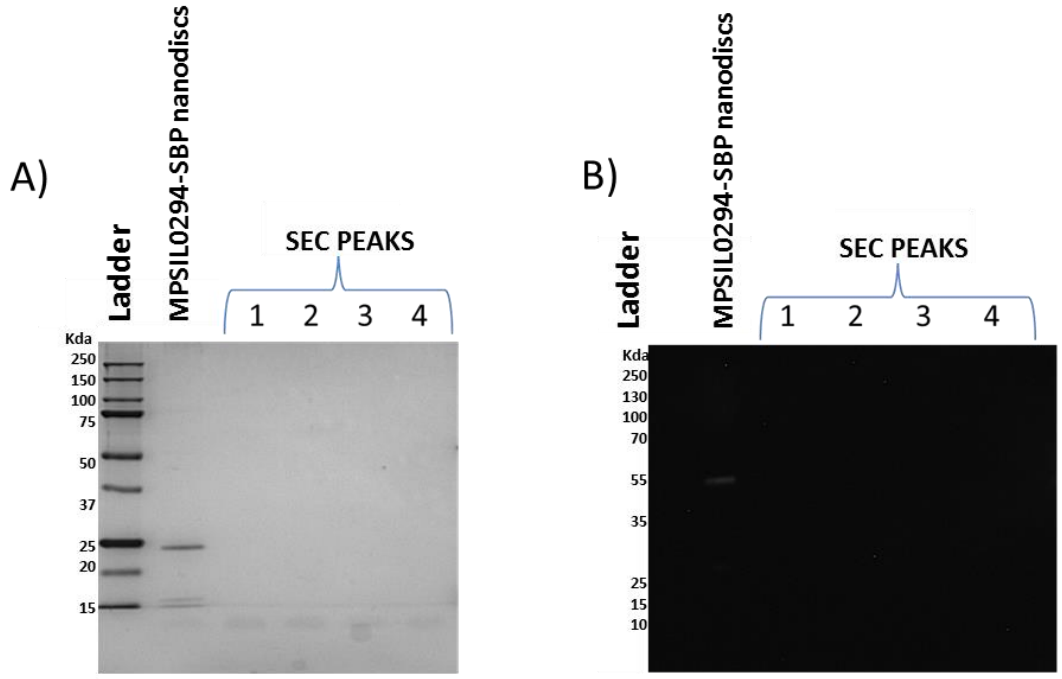


Figure 4.2: <u>Analysis of MPSIl0294-SBP nanodisc samples recovered from the gel filtration analysis:</u> Samples collected from the size exclusion chromatography of MPSIL0294-SBP filled nanodiscs run on a 12% SDS-PAGE before A) silver staining using the Thermo scientific silver staining kit on which a single protein band is resolved with a mass of 25 kDa. Or B) Western blotting probed with an anti polyhistidine tag antibody conjugated with HRP on which a single protein band has been resolved at 55 kDa. The numbers denote each peak from the chromatogram shown in figure 4.1 in order from left to right.

Fresh samples of the nanodiscs were also analysed by sucrose density gradient (see **Section 2.5.5**). Each fraction was analysed by western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP. The blot shown in **Figure 4.3** shows two bands with a mass of approximately 55 and 25 kDa, indicative of MPSIL0294 and the MSP respectively. Densitometric analysis via the software imageJ (National institutes of Health) shows that the 25 kDa band is approximately 1.7 times brighter than the 55 kDa band. This relative difference between MPSIL0294 and the MSP is as expected for a homogenous population of nanodiscs due to the fact that a typical nanodisc has two MSPs for each disc.

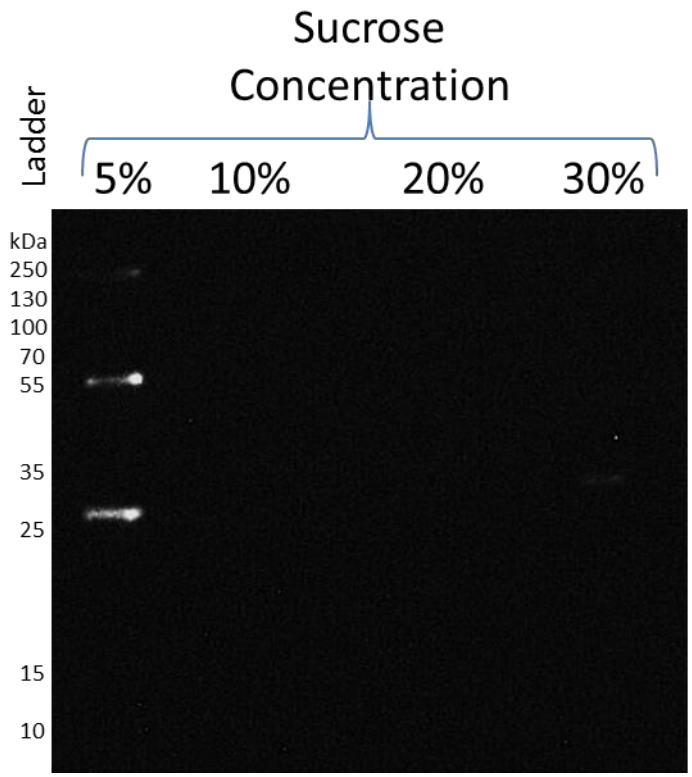


Figure 4.3 Sucrose density gradient analysis of MPSIL0294-SBP nanodiscs: Samples of freshly prepared nanodiscs were analysed via sucrose density gradient and subsequently analysed by western blotting probed with anti-polyhistidine tag antibody conjugated with HRP. The resultant blot shows the presence of two distinct protein bands. Densitometric analysis of these bands shows that the 25 kDa protein band is 1.5 times brighter than the band at 55 kDa.

Purified samples of AcrB were also used to form nanodiscs with POPC and MSP1E3 in a molar ratio of 2:3:101 in an identical manner as described above and in **section 2.4.5**. This formulation was performed by Dr. Vincent Postis of the University of Leeds. However the membrane scaffold protein was biotinylated using EZ-Link® NHS-SS-Biotin (Thermo scientific) as described in **section 2.4.6**.

4.3 The preparation of styrene maleic acid lipid particles (SMALPs)

SMALPs are a novel alternative to the solubilisation of membrane proteins in detergent micelles, instead holding membrane proteins within discoidal structures which are conceptually similar to nanodiscs. SMALPs differentiate themselves, however, in several ways; firstly, they do not require the initial purification of the target membrane protein from its lipid environment. Instead, the styrene maleic acid co-polymer can be added directly to membranes as described in **section 2.4.4.** Therefore SMALPs are able to capture membrane proteins in

their native environments, surrounded by their native lipids. Secondly, SMALPs do not require the use of any other proteins, unlike nanodiscs which require MSPs to encircle the phospholipid core. Inclusion of exogenous proteins such as MSP into phage display screens can be very problematic and requires extra precautions to be taken to ensure that isolated binding proteins are not binding directly to the exogenous protein. Thirdly, they can be purified in an identical manner as detergent solubilised membrane protein and can withstand multiple freeze-thaw cycles. A major drawback of SMALPs is that the maleic acid moiety on the polymer is very negatively charged which has negative implications for future assays and their purification as described in the next section.

4.3.1 The production of MPSIL294-SBP SMALPs

In order to maintain consistency between the different forms of protein presentation MSPIL0294-SBP was chosen to be placed within a SMALP. The procedure was performed as described in **section 2.4.4** on inner membrane of *E.coli* strains expressing MPSIL0294-SBP produced via auto-induction in 4L SB-5052 media (**see Table 2.6**) and isolated using a cell disruptor as described in **section 2.4.3.** The SMALPs were purified in an identical manner to MPSIL0294-SBP in detergent except in a buffer composed of 50 mM Tris, 500 mM NaCl and 10% glycerol. The particularly high concentration of NaCl was used to compensate for the negative charge introduced by the SMA co-polymer.

Samples were taken throughout the SMALP purification process in order to assess the success of each stage. These samples were then analysed electrophoretically on a 12% SDS-PAGE gel followed by extensive staining in Coomassie blue (see **Figure 4.4**). Interestingly there appears to be an abundance of an unidentified protein produced with a weight of approximately 25 kDa in the inner membranes isolated from an *E.coli* strain expressing MPSIL0294. This endogenous *E.coli* protein appears to be insoluble in the presence of the SMA co-polymer as it does not appear in the adjacent 'supernatant' sample but it was present when MPSIL0294-SBP was purified in detergent (see chapter 3). Several proteins were subsequently eluted from the resin as bands appear at approximately 100, 55 and 25 and 20 kDa.

In order to confirm the presence of MPSIL0294-SBP SMALPs the elution fractions were pooled, dialysed in order to remove the 250 mM imidazole in the elution buffer and concentrated prior to analysis via western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP as well as a Coomassie stained 12% SDS-PAGE. The gel and western blot shown in **Figure 4.5** provide confirms the presence of MPSIL0294 in **Figure 4.4**. The Coomassie stained SDS-PAGE gel resolves several bands which do not appear on the western blot of the same sample.

This therefore indicates a large degree of co-elution of endogenous *E.coli* proteins. Protein bands occur at approximately 100, 50, just above 25 and just below 20 kDa. The 50 kDa protein has a polyhistidine tag as proven by its appearance on the western blot and is therefore MPSIL0294. The identities of the other proteins cannot be determined definitively, but due to its propensity to spontaneously bind to IMAC resin it is quite likely that the 100 kDa protein band is AcrB. This is a common contaminant of IMAC on *E.coli* produced proteins due to its histidine rich C-terminal (Veesler, Blangy et al. 2008). There are several different endogenous proteins which could be the identity of the other protein bands on the gel which have been shown to commonly contaminant protein samples purified via IMAC with *E.coli* membranes, for example the cadmium response protein YodA (22.3 kDa) or chloramphenicol acetyltransferase (CAT- 25.5kDa) (Bolanos-Garcia and Davies 2006) such as . It could be argued that these endogenous proteins should not interfere with downstream attempts to screen MPSIL0294-SBP SMALPs as they have no way of binding a streptavidin coated surface and therefore should be washed away following immobilisation.

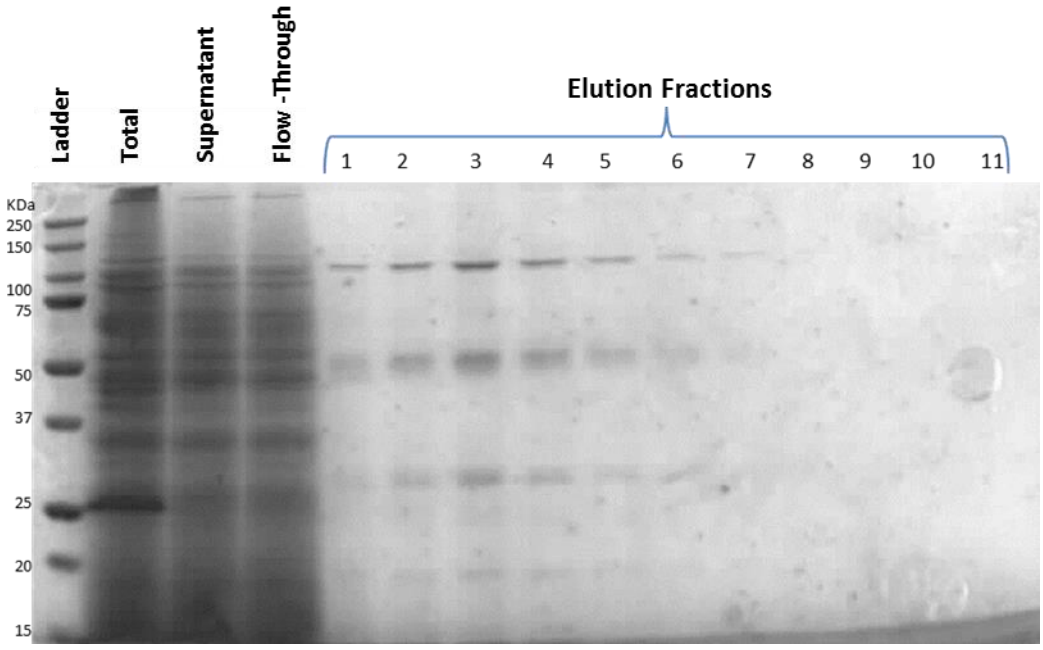


Figure 4.4: <u>Purification of MPSII0294-SBP SMALPs:</u> Samples taken throughout the solubilisation and purification of MPSIL0294-SBP with SMA co-polymer analysed on a 12% SDS-PAGE gel stained with Coomassie blue. 'Total' is a sample taken from the membrane after solubilisation with SMA, 'Supernatant' is a sample of the solubilised material alone, 'Flow-Through' is a sample of the material which passed through the column.

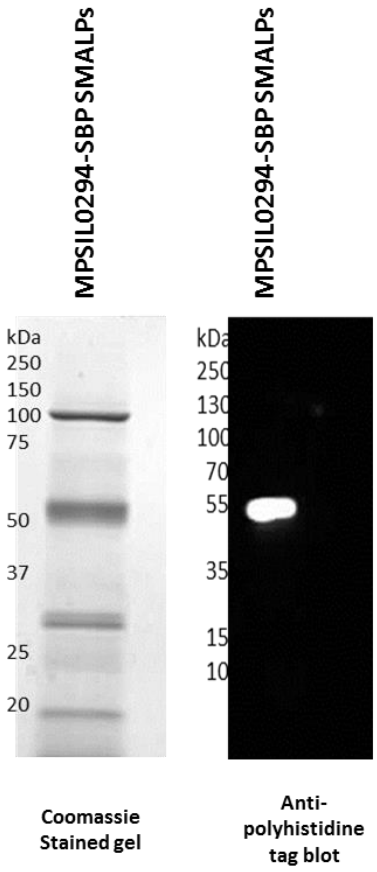


Figure 4.5: <u>Analysis of final MPSIL0294-SMALPs samples:</u> The elution fractions of MPSIl0294-SMALPs were pooled, dialysed and concentrated before analysis via Coomassie stained 12% SDS-PAGE gel and western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP.

4.3.2 The biotinylation of various membrane protein SMALPs

Samples of AcrB, VcCNT and NupC SMALPs were prepared by Dr. Vincent Postis and Zhenyu Hao in the University of Leeds and were produced in an identical manner as described in **section 2.4.4**. In order to bind these samples to streptavidin coated plates as part of the phage display process each of these samples were biotinylated using EZ-Link® NHS-SS-Biotin (Thermo scientific) in an identical manner as described in **section 2.4.6** on site in the MedImmune facility located in Cambridge. These samples were first subjected to extensive dialysis in order to exchange the buffer in which they were stored in from the 50 mM Tris, 500 mM NaCl and 10% glycerol buffer to PBS, 500 mM NaCl and 10% glycerol. Samples were each then biotinylated with a 5 molar excess of biotin targeting primary amine groups such as lysine residues within the protein SMALPs. The samples were then analysed by SDS-PAGE stained with Coomassie blue and Western blotting probed with an anti-biotin antibody.

The gel and blot shown in **Figure 4.6** indicate that the VcCNT and NupC SMALP samples (and to a much lesser extent both the AcrB SMALPs and detergent solubilised samples) contain various impurities, which might limit the specificity if hits acquired by the phage display procedure.

As expected, after biotinylation, all target membrane proteins are detected with Western blotting with biotin antibodies (**Figure 4.6**). Especially those at approximately 30 kDa in NupC and VcCNT, 50 kDa in the detergent solubilised AcrB and 30 kDa in AcrB SMALPs. The samples were then immobilised on a streptavidin coated surface and subjected to two rounds of phage display.

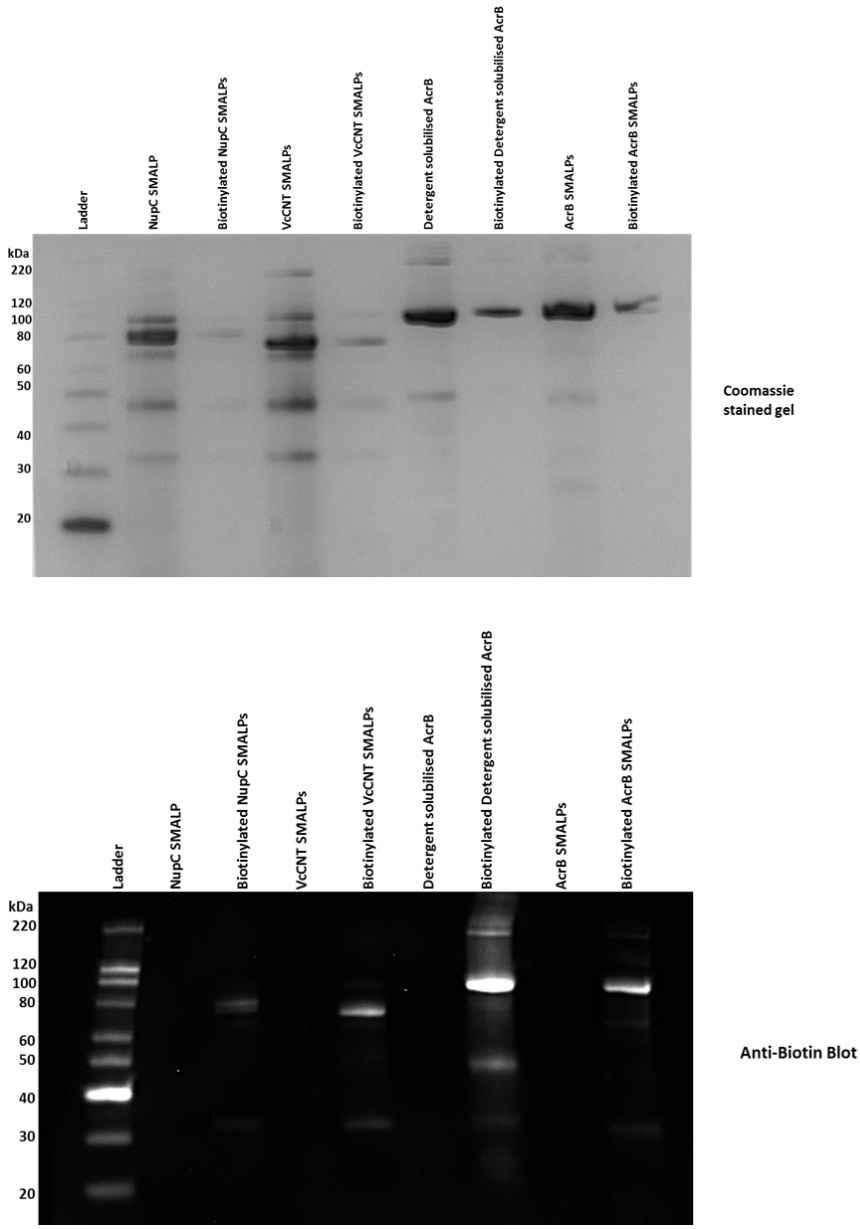


Figure 4.6 <u>Biotinylation of samples of the various members of the membrane protein panel:</u> Samples of NupC, VcCNT and AcrB SMALPs along with detergent solubilised AcrB were biotinylated using a 5 molar excess of EZ-Link® NHS-SS-Biotin (Thermo scientific) and analysed via Coomassie stained SDS-PAGE and western blotting probed with anti-biotin antibody conjugated with HRP.

4.4 Quality control of alternative detergent free methods of protein solubilisation with MPSIL0294 via ELISA

Samples of MPSIL0294-SBP SMALPs and nanodiscs were subjected to indirect ELISA on neutravidin coated plates, in the same way its detergent solubilised counterparts were (**see section 3.6.1**). This ELISA was performed in order to prove that the nanodiscs and SMALPs could successfully be immobilised onto a surface via the SBP tag. The resultant graph in **Figure**

4.7 shows that despite the fact that the MPSIL0294-SBP nanodiscs seem to saturate the plate, they produce a trace which is 3.5 times smaller than the detergent solubilised MPSIL0294-SBP (**see Figure 3.9**). Its highest data point is observed in the presence of 100 µg/mL with an A_{450nm} reading of 0.414, the detergent solubilised MPSIL0294-SBP on the other hand achieved an A_{450nm} trace of 1.4 at the same concentration. The SMALPs trace on the other hand (which used the same MPSIL0294 variant) failed to produce any signal and overlays that of the BSA negative control. This could either be due to the failure of the antibody in gaining access to the polyhistidine tag or due to the failure of the protein to bind to the plate, either of which could be a consequence of the substantial negative charge associated with the maleic acid moiety.

In order to prove that the MPSIL0294-SBP SMALPs are capable of immobilisation via its SBP tag, they were re-bound to strep-tactin resin pre-equilibrated in SMALP buffer (50 mM Tris, 10% glycerol and 500 mM NaCl, pH 8) in a disposable column (Thermo scientific). The bound protein was then eluted in the aforementioned SMALP buffer supplemented with 2.5, 5 and 10 mM desthiobiotin followed by the same buffer supplemented with 5 and 10 mM biotin. Samples were taken throughout this procedure, including an aliquot of the resin itself after protein binding and elution (**Figure 4.8**). These samples were analysed by western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP. The resultant blot shown in **Figure 4.8** shows that while the majority of the SMALPs fail to bind to the resin and are therefore detected in the 'Flow Through'; a proportion of the SMALPs were successfully eluted primarily by 2.5 mM desthiobiotin. Very faint bands can still be seen after treatment with 5 mM of both desthiobiotin and biotin.

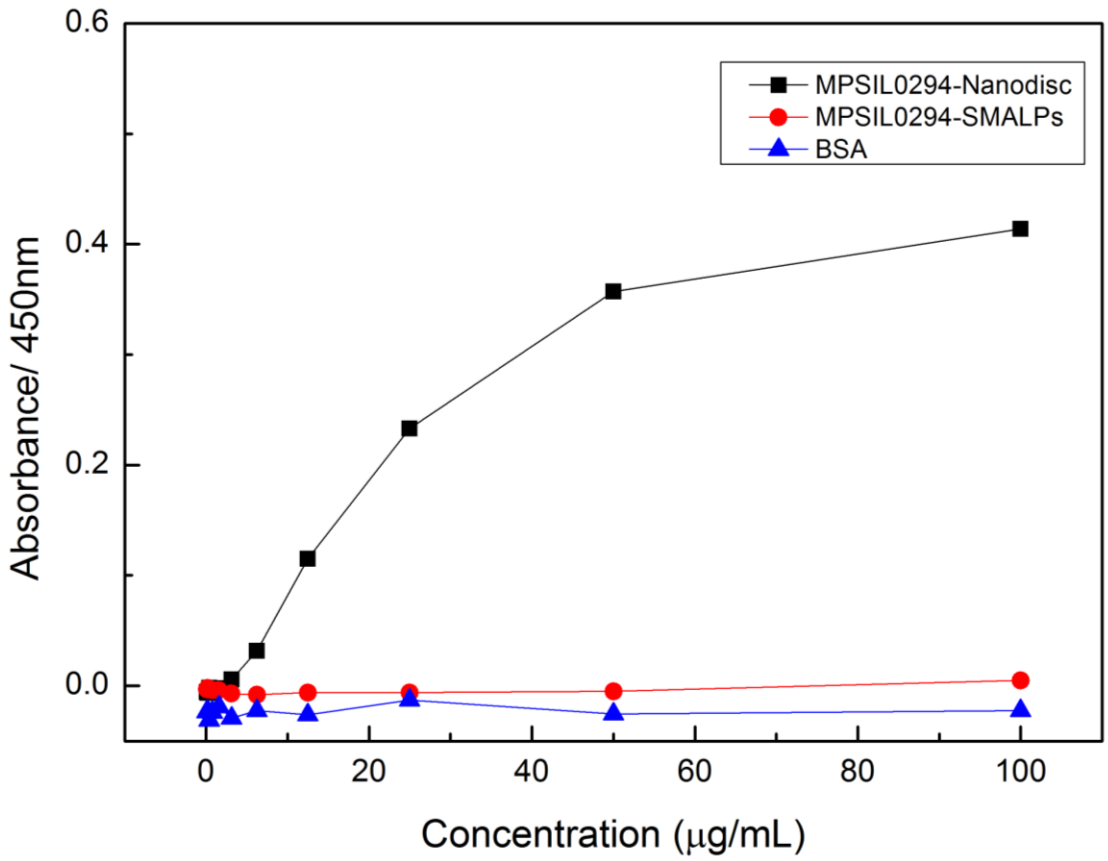


Figure 4.7: <u>Quality control ELISA of detergent free MPSIL0294 samples:</u> The absorbance data at 450 $_{nm}$ were recorded after samples of MPSIL0294-SBP SMALPs and nanodiscs were immobilised on a neutravidin coated surface and probed with an anti-polyhistidine tag antibody conjugated with HRP. Readings were then recorded after incubation with TMB followed by 0.5 M H_2SO_4.

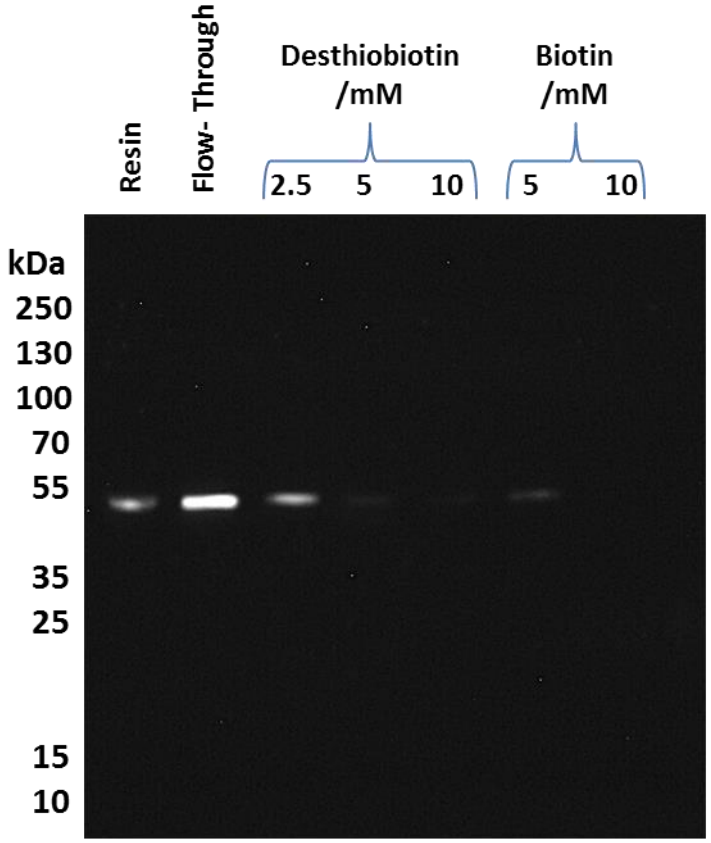


Figure 4.8 Immobilisation of MPSIL0294-SBP SMALPs on strep-tactin resin: Samples of MPSIL0294-SBP SMALPs were bound to strep-tactin resin and samples were taken throughout the procedure in order to assess whether or not the protein could be immobilised via it C-terminal SBP tag. These samples were analysed by western blotting probed with an anti polyhistidine antibody conjugated with HRP.

4.5 A Comparison of different methods of membrane protein solubilisation

Two rounds of phage display were carried out on the various membrane protein samples comparing alternative methods of protein solubilisation. These rounds of phage display (and the subsequent validation tests performed on the various selection outputs) were carried out in an identical manner as previously described for the comparison of membrane protein immobilisation strategies in chapter 3. In order to carry out the phage display, samples of AcrB SMALPs (which had been biotinylated as discussed earlier) and nanodiscs (the MSPs of which had undergone chemical biotinylation) were immobilised on streptavidin coated plates. MPSIL0294-SBP SMALPs and nanodiscs along with samples of chemically biotinylated NupC and VcCNT SMALPs were also prepared in a similar fashion; after which all were subjected to phage display as before with a functional library of DARPins (**see section 2.8**) the results of which are in **Table 4.1** below.

The input titres, calculated before the second round of phage display, are well within expected parameters at 10^{12} as are the output titres at 10^5 and 10^6. The same is true for all of the sequence diversities except for the VcCNT SMALPs which had a diversity measure of 93%, thereby suggesting a large number of unique sequences in its selection output.

Sample	Format	Input titre (cfu/mL)	Output titre (cfu/mL)	Diversity
MPSIL0294-SBP	Nanodisc	2.8×10^{12}	4.3×10^5	82%
	SMALP	3.5×10^{12}	3.2×10^5	66%
AcrB	Biotinylated Nanodisc	3×10^{12}	5×10^5	70%
	SMALP	1.1×10^{12}	7.13×10^6	79%
NupC	SMALP	$2.7x\ 10^{12}$	2.65×10^5	73%
VcCNT	SMALP	3.5×10^{12}	9.5×10^5	93%

Table 4.1: Overview of the input and output titres of the phage used in phage display against the detergent free protein samples: The input titre calculated before the second round of phage display and the subsequent output titres obtained after the second round with immobilised samples of membrane proteins in different presentation formats. The level of uniqueness in each selection has also been calculated after the removal of all repeating sequences.

4.5.1 Removing streptavidin binding DARPins via Phage ELISAs on blank streptavidin plates

In order to validate the selection outputs, an identical series of tests as those performed on the alternatively tagged proteins (**see Chapter 3**) were performed on the various nanodisc and SMALP samples. The first of which was the identification and removal of any DARPins capable of binding streptavidin directly via phage ELISA on blank streptavidin plates. The same A_{450nm} threshold of 0.1 was imposed upon these ELISAs as it was with the validation tests performed on the previous samples and just like those ELISAs any DARPins which crossed this threshold were deemed to have the ability to bind streptavidin and were consequently removed from further characterisation.

4.5.1.1 AcrB Biotinylated Nanodiscs and SMALPs

Phage ELISAs on the selection output of AcrB biotinylated nanodiscs successfully identified 67 DARPins which produced an A_{450nm} reading above 0.1.This is the highest number of the streptavidin binding DARPins identified for all AcrB formats. The SMALPs identified 56 and the

detergent solubilised identified 55. In terms of the actual A_{450nm} readings produce by the nanodisc DARPins, many are quite elevated relative to detergent solubilised MPSIL0294, with the highest value being 0.556 from DARPin C6. However none of them produce a signal above 1.0 like both the detergent solubilised AcrB and the AcrB SMALPs. A few of these DARPins do appear to bind streptavidin with a moderately high affinity producing A_{450nm} values of 0.441 (DARPin F3), 0.484 (G7) and 0.394 (E9) (**see Table 4.2**).

Of the 58 streptavidin binding DARPins identified by the phage ELISA on the AcrB SMALP selection output, four bound particularly tightly to the streptavidin. These DARPins produced A_{450nm} readings at 1.596 (DARPin G7), 1.833 (D5), 1.955 (D8) and the highest at 1.966 (E2). Generally, many of the other streptavidin binding DARPins also produce signals higher than those identified by the nanodiscs. The DARPins which were identified as able to bind to streptavidin were not used in any further validation tests.

AcrB Biotinylated Nanodiscs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.072	0.133	0.143	0.065	0.119	0.147	0.072	0.153	0.123	0.159	0.197	0.041
B	0.068	0.06	0.109	0.085	0.175	0.113	0.079	0.13	0.226	0.129	0.199	0.078
C	0.058	0.062	0.221	0.135	0.137	0.556	0.099	0.37	0.115	0.149	0.195	0.055
D	0.081	0.06	0.123	0.109	0.242	0.368	0.161	0.242	0.189	0.166	0.229	0.055
E	0.075	0.118	0.086	0.095	0.161	0.346	0.315	0.246	0.394	0.218	0.179	0.063
F	0.067	0.08	0.441	0.121	0.121	0.148	0.367	0.198	0.34	0.245	0.166	0.045
G	0.061	0.071	0.18	0.257	0.151	0.158	0.484	0.172	0.178	0.227	0.229	0.049
H	0.059	0.073	0.176	0.115	0.147	0.151	0.283	0.339	0.207	0.24	0.334	0.072

AcrB SMALPs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.065	0.063	0.098	0.082	0.254	0.06	0.083	0.074	0.199	0.172	0.124	0.073
B	0.068	0.332	0.184	0.062	0.236	0.225	0.076	0.21	0.344	0.216	0.33	0.042
C	0.068	0.255	0.171	0.076	0.3	0.115	0.238	0.257	0.667	0.278	0.216	0.057
D	0.105	0.159	0.219	0.607	1.833	0.143	0.058	1.955	0.301	0.093	0.266	0.059
E	0.082	1.966	0.094	0.122	0.39	0.098	0.086	0.262	0.597	0.058	0.635	0.065
F	0.139	0.061	0.206	0.214	0.088	0.147	0.067	0.196	0.685	0.106	0.333	0.061
G	0.209	0.077	0.085	0.127	0.192	0.1	1.596	0.093	0.19	0.302	0.606	0.072
H	0.094	0.075	0.071	0.063	0.279	0.06	0.165	0.249	0.06	0.075	0.419	0.078

Table 4.2 Primary phage ELISA validation of the selection outputs selected against AcrB solubilised via detergent free methods: The data produced after a phage ELISA was performed on the selection outputs of both the AcrB nanodiscs and the AcrB SMALPs on blank streptavidin plates with no immobilised target protein. Both of these protein formats were immobilised via a biotin - streptavidin interaction during the initial phage display. The nanodiscs were constructed using an MSP which had undergone chemical biotinylation, while the SMALPs were chemically biotinylated prior to phage display. The ELISAs were probed with anti-M13 antibody conjugated with HRP and visualised by incubation with TMB prior to treatment with 0.5 M H_2SO_4 and the analysis of its absorbance at 450nm. DARPins which did not pass the 0.1 threshold are highlighted in green and were taken on for further validation while DARPins which showed weak streptavidin binding (0.1 – 0.49) are highlighted in yellow, DARPins which showed a moderate ability to bind streptavidin (0.5 – 0.99) are highlighted in orange and DARPins which strongly bound streptavidin (above 1.0) are marked in red. The negative control was composed of wells with no phage added and are written in blue text.

4.5.1.2 MPSIL0294-SBP Nanodiscs and SMALPs

The same initial validation test was carried out on the selection outputs of MPSIL0294-SBP nanodiscs and SMALPs. The nanodiscs highlighted 17 DARPins which passed the 0.1 A_{450nm} threshold, while the SMALPs identified 11. Thereby making the detergent solubilised MPSIL0294-SBP the highest with 69 DARPins. Of the 17 nanodisc streptavidin binders 9 DARPins produce a signal above 1.0, the highest of which is DARPin E2 and A6 which both produce a signal at 1.666. Of the remaining 8 DARPins, 2 produced signals above 0.9 namely DARPins F5 (0.938) and G9 (0.991).The SMALPS on the other hand isolated 2 DARPins with an A_{450nm} above 1.0 out of the 11 DARPins which bound to streptavidin. These 2 DARPins produced signals of 1.359 (C4) and the highest of the DARPins, 1.572 (E6). The majority of the remaining DARPins A_{450nm} readings are very close to the threshold (**see Table 4.3**)

MPSIL0294-SBP Nanodiscs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.041	0.052	0.049	0.049	0.057	1.66	0.072	0.087	0.095	1.318	0.062	0.049
B	0.045	0.048	0.042	1.282	0.049	0.045	1.418	1.652	0.075	0.056	0.061	0.04
C	0.033	0.051	0.06	0.044	0.058	0.198	0.049	0.065	0.064	0.217	0.064	0.052
D	0.047	0.047	0.05	0.094	0.055	0.075	0.055	0.101	0.075	0.212	0.057	0.133
E	0.038	1.666	0.053	0.046	0.052	0.051	0.057	0.056	0.066	1.565	0.286	0.049
F	0.041	0.044	0.051	0.039	0.938	0.064	0.077	0.053	0.059	0.218	0.048	0.038
G	0.036	0.048	0.041	0.041	0.056	0.056	0.055	0.068	0.991	0.052	0.052	0.053
H	0.048	1.329	0.059	0.047	0.047	0.047	1.226	0.049	0.046	0.049	0.046	0.04

MPSIL0294-SBP SMALPs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.07	0.063	0.063	0.068	0.057	0.061	0.088	0.056	0.074	0.064	0.058	0.048
B	0.302	0.053	0.067	0.126	0.064	0.102	0.051	0.054	0.08	0.066	0.05	0.042
C	0.064	0.067	0.071	1.359	0.077	0.057	0.053	0.067	0.07	0.061	0.055	0.051
D	0.052	0.053	0.06	0.069	0.064	0.166	0.055	0.124	0.063	0.053	0.059	0.057
E	0.267	0.054	0.055	0.062	0.065	1.572	0.104	0.054	0.055	0.05	0.056	0.056
F	0.176	0.048	0.051	0.053	0.065	0.404	0.051	0.054	0.056	0.045	0.047	0.044
G	0.046	0.051	0.054	0.052	0.062	0.057	0.052	0.059	0.057	0.056	0.064	0.047
H	0.051	0.053	0.054	0.049	0.049	0.048	0.051	0.051	0.05	0.052	0.046	0.038

Table 4.3: Primary phage ELISA validation of detergent free MPSIL0294 samples: The data from the phage ELISAs carried out on blank streptavidin plates using the selection output from MPSIL0294-SBP Nanodiscs and SMALPs. The ELISAs were probed with anti-M13 antibody conjugated to HRP; analysis of the A_{450nm} was achieved via incubation with TMB followed by the addition of 0.5M H_2SO_4. DARPins which crossed the threshold of 0.1 were deemed to be binding directly to streptavidin and were therefore removed from further validation experiments. DARPins which bound streptavidin weakly (0.1 – 0.49) are highlighted in yellow, those which bound to streptavidin moderately (0.5 – 0.99) are highlighted in orange and those that bound streptavidin strongly (above 1.0) are highlighted in red. The DARPins which did not bind to streptavidin and could therefore be subjected to further validation tests are highlighted in green. The negative control was composed of wells which were treated with phage and are highlighted in blue text.

4.5.1.3 NupC and VcCNT SMALPs

The selection outputs produced as a result of phage display performed against NupC and its homologue VcCNT were also validated via phage ELISA on blank streptavidin plates. These ELISAs identified 54 and 10 DARPins with the ability to bind directly to streptavidin respectively. Despite this large difference in streptavidin binders between the two homologues none of them produced an A_{450nm} signal above 1.0 as was the case with both MPSIL0294 and AcrB SMALPs. Instead the highest signal produced by these ELISAs was DARPin F1 from the NupC selection output and D7 from VcCNT (**see Table 4.4**).

NupC SMALPs												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	0.079	0.19	0.072	0.26	0.061	0.202	0.06	0.238	0.134	0.166	0.313	0.055
B	0.065	0.111	0.28	0.066	0.344	0.059	0.09	0.455	0.273	0.448	0.073	0.055
C	0.049	0.094	0.225	0.1	0.34	0.078	0.085	0.281	0.203	0.335	0.102	0.041
D	0.113	0.133	0.104	0.128	0.354	0.053	0.104	0.257	0.468	0.434	0.487	0.053
E	0.183	0.284	0.096	0.108	0.097	0.099	0.092	0.19	0.098	0.248	0.365	0.032
F	0.661	0.1	0.093	0.063	0.463	0.119	0.194	0.334	0.221	0.187	0.274	0.06
G	0.17	0.059	0.256	0.082	0.085	0.074	0.136	0.273	0.487	0.114	0.449	0.23
H	0.083	0.081	0.065	0.089	0.438	0.061	0.054	0.087	0.119	0.09	0.574	0.046

VcCNT SMALPs												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	0.063	0.068	0.056	0.249	0.055	0.054	0.296	0.056	0.177	0.058	0.058	0.051
B	0.049	0.054	0.063	0.053	0.051	0.053	0.05	0.051	0.055	0.057	0.05	0.063
C	0.046	0.057	0.058	0.056	0.058	0.055	0.295	0.058	0.059	0.059	0.057	0.058
D	0.052	0.053	0.062	0.057	0.049	0.053	0.443	0.056	0.05	0.211	0.051	0.043
E	0.046	0.078	0.057	0.053	0.056	0.137	0.285	0.052	0.055	0.053	0.054	0.077
F	0.061	0.051	0.067	0.055	0.276	0.053	0.069	0.061	0.054	0.054	0.047	0.043
G	0.044	0.051	0.055	0.055	0.059	0.056	0.315	0.058	0.051	0.051	0.054	0.05
H	0.08	0.053	0.056	0.066	0.056	0.058	0.058	0.074	0.059	0.059	0.056	0.043

Table 4.4 Primary Phage ELISA validation of selection outputs generated against NupC and VcCNT SMALPs: The data produced after phage ELISAs were performed using the selection outputs from the VcCNT and NupC SMALPs on blank streptavidin plates. The ELISA plates were probed with an anti-M13-HRP antibody conjugate and their A_{450nm} were measured after incubation with TMB and subsequently 0.5M H_2SO_4. DARPins which did not exceed the pre-determined A_{450nm} threshold of 0.1 are highlighted in red and were taken on for further validation.

4.5.2 Refining selection outputs via Phage ELISA in the presence of target protein

Once all DARPins with a capability to bind streptavidin alone were identified and removed, the phage ELISAs were repeated in the presence of 100 μg of immobilised target protein and the second A_{450nm} threshold of 1.0 was once again used to confirm if a DARPin binds to its target. Due to the low expression and purification yields of MPSIL0294-SBP, further characterisation by phage ELISA was not feasible for this target protein and no further action was taken in this study with MSPIL0294-SBP

4.5.2.1 AcrB Biotinylated Nanodiscs and SMALPs

The AcrB nanodiscs and SMALPs were immobilised to the surface of streptavidin coated plates via chemical biotinylation (of the MSP in the nanodisc and of surface lysine residues aspecifically in the SMALP) and treated with the DARPins which did not produce a signal in the previous phage ELISAs. The resultant A_{450nm} readings show that 4 of the remaining 21 DARPins produced a signal above the 1.0 threshold and may binding to AcrB (a success rate of 19%). These DARPins are D1 with a reading of 1.581, F1 at 1.921, E3 at 2.125 and H1 at 2.136 which is the highest reading on the plate. Other than these four DARPins no others are able to produce a reading above 0.15, therefore the cut off value of 1.0 is appropriate for this selection.

In contrast to this, all of the 30 remaining DARPins from the selection output against AcrB SMALPs bind tightly to AcrB when tested in a phage ELISA. The lowest A_{450nm} value was produced by DARPin D10 and had a value of 2.421 thereby suggesting that each DARPin is capable of saturating their individual wells. All of the DARPins which showed an ability to bind on this produced A_{450nm} readings above 2.4 (**see Table 4.5**).

AcrB Biotinylated Nanodiscs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.064			0.097			0.076					0.06
B	0.085	0.137		0.117			0.098					0.056
C	0.066	0.07					0.084					0.068
D	1.581	0.044										0.058
E	0.079		2.125	0.113								0.063
F	1.921	0.121										0.06
G	0.061	0.077										0.064
H	2.136	0.101										0.056

AcrB SMALPs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	2.903	2.829	2.785	2.64		2.637	2.661	2.674				0.043
B	2.542			2.579			2.611					0.038
C	2.718			2.582								0.043
D							2.486			2.421		0.053
E	2.743		2.655				2.451			2.453		0.042
F		2.679					2.514					0.037
G		2.664	2.575					2.422				0.043
H	2.767	2.76	2.642	2.674		2.562			2.521	2.649		0.037

Table 4.5 Secondary phage ELISA validation of the selection outputs generated against the detergent free methods of AcrB solubilisation: The data produced after the second phage ELISAs were performed on streptavidin coated plates with 100 µg of immobilised AcrB SMALPs and Biotinylated nanodiscs on their surface. The ELISAs were probed with an anti-M13-HRP antibody conjugate and the A_{450nm} values were recorded after incubation with TMB and subsequent colorimetric change by treatment with 0.5M H_2SO_4. DARPins which were removed after the previous phage ELISAs have been represented with black squares while all DARPins which produced a signal in the presence of target protein are highlighted in green. DARPins which did not show any binding activity in the presence of target protein (below 0.09) are highlighted in red and those which only showed a weak ability to bind to the target (0.1 – 0.49) are highlighted in yellow. The negative control consisted of wells which were treated with any phage and are highlighted in blue text.

4.5.2.2 NupC and VcCNT SMALPs

Samples of NupC and VcCNT SMALPs were treated in an identical manner as discussed above with 100 µg of each SMALP immobilised onto the surface of a streptavidin coated plate. The

phage ELISA results in **Table 4.6** shows that the vast majority of both selection outputs produced a signal above the threshold. In the case of NupC 23 out of the 34 DARPins (67%) bound to NupC. The plate shows (as was the case of AcrB in **Table 4.5**) a very clear distinction between DARPins which bind the target and those that fail to do so with the exception of H1 which produced a moderate reading. Once again however, the 1.0 A_{450nm} threshold appears to be appropriate for this selection as the highest value was C1 with a value of 1.839.

The VcCNTs phage ELISA plate shown in **Table 4.6** shows that 60 out of 78 (77%) DARPins successfully produced a signal above 1.0 in the presence of immobilised target protein presented in SMALPs. The highest value was obtained from DARPin C1 with a value of 1.858. While the 1.0 threshold was appropriate for the selections against AcrB and NupC, it is more arbitrary for VcCNT as many of the DARPins produced readings above 0.9 such as H1, G2 and B3. In general both phage ELISAs for NupC and VcCNT have resulted in a cohort of DARPins which produce distinctively lower A_{450nm} values than AcrB did when it was placed in a SMALP. The cause of this decrease is not clear, although, considering the large extracellular domain of AcrB it is possible that the DARPins had greater accessibility to more binding sites. The difference in A_{450nm} values could alternatively be caused due to an increased amount of AcrB adsorbed onto the plate.

NupC SMALPs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	1.551		1.823		1.705		1.712					0.042
B	0.101			0.065		1.68	1.668				1.409	0.036
C	1.839	1.738				0.045	0.051					0.043
D						1.523						0.032
E			0.049		1.397	0.066	1.183		1.635			0.04
F			1.518	0.071								0.034
G	1.578	1.617		1.688	1.631	1.603						0.039
H	0.602	0.074	1.025	0.076		1.64	0.145	1.759		1.705		0.035

VcCNT SMALPs												
	1	2	3	4	5	6	7	8	9	10	11	12
A	1.581	1.422	0.075		1.117	0.922		1.032		0.469	1.421	0.05
B	1.523	1.346	0.949	1.36	1.594	0.394	1.349	1.482	1.359	1.623	1.558	0.04
C	1.858	1.328	1.507	1.366	1.471	1.395		1.413	1.301	1.493	1.455	0.046
D	1.299	0.079	1.285	1.411	0.91	1.297		1.203	1.439		1.46	0.086
E	1.213	1.473	1.327	1.486	1.325			1.183	1.61	0.06	0.882	0.082
F	1.58	1.393	1.635	1.448		1.39	1.488	0.559	1.509	1.514	1.388	0.056
G	0.213	0.967	0.115	1.471	1.124	1.25		1.266	1.607	1.709	1.525	0.13
H	0.947	0.186	0.6	0.081	1.632	1.565	0.142	1.481	1.581	1.25	1.765	0.037

Table 4.6 Secondary phage ELISA validation of the selection outputs generated against NupC and VcCNT SMALPs: The absorbance readings recorded after phage ELISAs were performed on streptavidin coated plates onto which 100 µg of NupC and VcCNT SMALPs were immobilised. The ELISAs were carried out the selection outputs of the aforementioned proteins and were probed with an anti-M13-HRP antibody conjugate. The initial colorimetric change was initiated by the incubation of the plate with TMB. The final change and absorbance measurements were taken after the addition of 0.5M H_2SO_4. Wells in which DARPins previously showed an affinity toward streptavidin alone in the previous phage ELISAs have been blacked out, DARPins which produce a signal higher the 1.0 threshold after treatment in this phage ELISA have highlighted in red.

4.5.3 An overview of phage display on different methods of membrane protein solubilisation

The alternative protein solubilisation samples were subjected to two rounds of phage display identical to those performed in chapter 3, on membrane protein samples solubilised in DDM. The phage display was followed by the same series of validation tests which aimed to refine the various selection outputs by removing DARPins which bound streptavidin directly and then subsequently confirm their ability to bind to the target used in the phage display.

Table 4.7 shows an overview of the results produced by the phage display and subsequent validation tests with regard to the number of DARPins that were available for each subsequent stage of testing. When compared with the same overview shown in chapter 3 (see **Table 3.6)**, the data in **Table 4.7** shows that in the case of MPSIL0294, there is profound increase in the number of DARPins that did not produce a signal in the presence of streptavidin alone. AcrB on the other hand, shows fairly consistent numbers which may be the result of AcrB packing onto the surface of the plate in a consistent manner regardless of the solubilisation strategy employed.

The DARPins were then removed from the second phage ELISA performed in the presence of immobilised target protein (stage 2). As previously mentioned the nanodisc and SMALP samples of MPSIL0294 could not be further validated due to restrictions in the availability of sample. The AcrB samples show an increase in the number of DARPins that appear to bind the target when detergent free methods of solubilisation are used. This is particularly true for the SMALPs, which also show a high number of DARPins binding to the selection outputs of NupC and VcCNT were tested in the same manner (see **Table 4.7**).

Target	Method of Immobilisation	Method of Solubilisation	Stage 1	Stage 2	Stage 3
MPSIL0294	SBP/Streptavidin	SMALP	88	77	N.D.
		Nanodisc	88	71	N.D.
		DDM	88	19	0
VcCNT	Lysine Biotinylation	SMALP	88	78	65
NupC	Lysine Biotinylation	SMALP	88	34	23
AcrB	Lysine Biotinylation	SMALP	88	30	30
	MSP Biotinylation	Biotinylated Nanodisc	88	21	4
	Lysine Biotinylation	DDM	88	33	0

Table 4.7: An overview of the entire phage display and subsequent validation process carried out on the various membrane protein samples solubilised via detergent free methods: The detergent free methods of membrane protein solubilisation: aimed to compare nanodiscs with SMA solubilisation in regard to their ability to isolate DARPins capable of binding the target proteins. The numbers with each cell represent the number of DARPins available for the subsequent stage. The entire process included: the initial two rounds of phage display (stage 1), followed by the first phage ELISA which identified all DARPins which bound streptavidin directly (stage 2), the second phage ELISA which provided initial confirmation of the DARPins ability to bind the various targets (stage 3) and finally the DARPins selected for expression, purification and further analysis. Note: N.D. - No data.

4. 6 Discussion and Conclusion

In order to compare different methods of membrane protein presentation, different methods of detergent free solubilisation were compared, namely the production of SMALPs and nanodiscs. Initial attempts to analyse the nanodiscs by SEC were not conclusive based upon the data generated by SDS-PAGE and western blotting, due to the lack of corroboration between the two (**see Figure 4.1 and 4.2**). Typically, SEC is employed in the analysis of nanodiscs, however when an incorrect ratio of MSP: lipid: protein is used, the excess lipid often results in a broad elution profile in the void volume due to aggregates (Bao, Duong et al. 2012, Pavlidou, Hanel et al. 2013). The peak shown in **Figure 4.1** however, is relatively sharp

thus suggesting the correct formation of MPSIL0294-SBP nanodiscs. The inconclusive nature of the subsequent SDS-PAGE and western blot which were used to confirm the presence of nanodiscs are the reason that analysis via sucrose density gradient was used as an alternative. The resultant blot shown in **Figure 4.3** was performed on a subsequent nanodisc preparation. The fact that the MSP protein band at 25 kDa is 1.7 brighter than the MPSIL0294 is expected due to the fact that a single nanodisc particle is composed of two MSP proteins encircling a phospholipid core. Confirmation that the MPSIL0294-SBP nanodiscs were properly formed is also provided by the ELISA shown in **Figure 4.7** which confirms the presence of a polyhistidine tag in the nanodiscs, although it cannot be established whether the his-tag of MPSIL0294 or that of MSP is detected. Although the presence of two his-tags should lead to an increase in fluorescence, as such it is unclear why a drastic decrease in adsorption value is observed instead. Possible explanations for this however may be due to the decreased accessibility of the polyhistidine tag to the anti his-HRP antibody conjugate. Nanodiscs composed of MPS1D1 and POPC which are estimated to have a diameter of 98 Å (9.8 nm) (Bayburt and Sligar 2010) while a full IgG antibody has a molecular weight of approximately 150 kDa. These two large molecules may sterically hinder one another thereby obstructing the antibody probe. Alternatively due to their large size, fewer nanodiscs may be able to pack onto the surface of the plate.

The production of MPSIL0294-SBP SMALPs as described in **section 2.4.4** produced samples which contain many contaminants. These protein bands which resolved on the Coomassie stained SDS-PAGE gel shown in **Figure 4.5** most likely co-purified as a consequence of either endogenous positively charged *E.coli* proteins strongly interacting with the negatively charged maleic acid moiety within the SMA co-polymer and co-purifying as a result. Alternatively, these protein bands may co-purify as a result of the electrostatic interaction between the negatively charged maleic acid moiety and the positively charged polyhistidine tag resulting in a lower affinity between the his tag and the cobalt resin. It has been observed that the presence of the SMA polymer reduces the affinity of the polyhistidine tag for the resin. In order to overcome this, an increased concentration of NaCl is used along with an overwhelming excess of resin. This excess of resin when taken into account alongside the poor expression level of MPSIL0294-SBP may explain why many endogenous *E.coli* proteins have co-purified alongside the MSPIL0294-SBP SMALPs. Unfortunately the effect of these co-purified proteins on the phage display outcome is not known. This electrostatic interaction is also thought to reduce the accessibility of the tag, thereby offering a possible explanation for the SMALPs failure to produce a signal in the ELISA shown in **Figure 4.7**. This might suggest that that the availability

of epitopes for phage display is diminished in SMALPs. The blots in **Figure 4.8** prove that the SBP tag on MPSIL0294 is accessible however and that this tag is capable of immobilising the protein onto a streptavidin surface despite the fact that a lot of SMALPs failed to do so.

Due to the fact that such a direct comparison of different methods of immobilisation and solubilisation on the same protein has not been done before, an interesting continuation of this work would be to increase the scope of presentation methods beyond detergent solubilised, nanodisc and SMALP. This would involve including molecules such as amphipols which are conceptually similar to nanodiscs in so far as they require membrane proteins to be initially solubilised in detergent. The key difference is that amphipols are comprised of an amphipathic polymer backbone similar to SMALPs (Popot, Berry et al. 2003, Zoonens, Zito et al. 2014). Amphipols are particularly interesting because biotinylated amphipols are already available and have been used to entrap an *E.coli* outer membrane protein, subsequently immobilise it onto a streptavidin coated surface and probe it with a bacteriophage receptor binding protein (Basit, Shivaji Sharma et al. 2012). These could therefore be used as a direct comparison with the AcrB nanodiscs which were constructed with biotinylated MSPs.

In this chapter AcrB was used as a model membrane protein in the comparison of alternative methods of detergent free membrane protein solubilisation in regard to their ability to isolate DARPins after two rounds of phage display. The solubilisation methods chosen were nanodiscs, which had been immobilised via an aspecifically biotinylated MSP, and SMALPs which had been immobilised by chemical biotinylation directly. SMALPs which had been constructed with NupC and VcCNT were also tested. The phage display and subsequent validation experiments show a consistent pattern in which screening with nanodiscs results in a higher proportion of the selection output binding directly to streptavidin when compared to SMALPs, however definitive conclusions cannot be made due to the fact that initial quality control experiments such as ELISAs on streptavidin coated plates were not carried out on these samples due to time constraints. Therefore confirmation of the SMALPs ability to bind to streptavidin coated surfaces would be the first experiment required in any continuation of this work.

AcrB and MPSIL0294 are the only members of the target protein panel which have samples subjected to all methods of solubilisation under test in this study (DDM, nanodisc and SMALP). Therefore, they may provide insights into the relationship between solubilisation strategies and the likelihood of isolating DARPins which bind streptavidin directly. In total, 78% of the DARPins isolated against MPSIL0294-SBP solubilised in DDM showed an ability to bind streptavidin (see **Table 3.2**). When MPSIL0294 is solubilised with SMA the on the other hand,

12.5% of the DARPins selected against the resultant SMALPs showed this activity (see **Table 4.3**). Finally when encapsulated in a nanodisc 19% of the DARPins selected against MPSIL0294 were able to bind streptavidin. This would appear to suggest that detergent free methods of solubilisation are superior to solubilisation via DDM when isolating DARPins in phage display (either as a result of improved packing on the streptavidin coated surface or by increased accessibility to surface epitopes).

The conclusion that detergent solubilisation is inferior in regard to the number of 'streptavidin binders' is also not supported by AcrB which shows a contradictory trend to MPSIL0294. In this instance, 59% of the DARPins selected against AcrB solubilised in DDM appeared to bind streptavidin (see **Table 3.3**), while 64% showed the same ability when selected against AcrB solubilised with SMA and finally 76% when selected against AcrB nanodiscs. The inclusion of NupC and VcCNT SMALPs also help to show that there is no clear relationship between the solubilisation method and the number of DARPins which bind streptavidin. As 61% and 11% of the DARPins selected against NupC and VcCNT respectively were streptavidin binders. The fact that these two proteins are homologues suggests that the solubilisation method and the number of streptavidin binding DARPins are independent factors. However, a prudent test to fully determine if a relationship exists between solubilisation method and the number of streptavidin binding DARPins, would be to tag AcrB with SBP and subject the subsequent samples of AcrB-SBP (solubilised in DDM, nanodiscs and SMALPs) to two cycles of phage display with the same DARPin library. A decrease in the number of 'streptavidin binders' selected against the nanodiscs and SMALPs, would confirm that detergent free methods of solubilisation are superior for SBP tagged proteins, in so far as reducing the number of streptavidin binding DARPins during a selection.

When the number of DARPins that bind to the target protein (as determined with phage ELISA) are compared, SMALPs consistently show a significantly higher number of 'hits' than both nanodiscs and DDM solubilisation as can be seen in **Table 3.6** and **Table 4.7**. While the AcrB nanodisc sample was only able to identify 4.5% of the total selection output as DARPin binders, the detergent solubilised samples failed to produce any DARPins capable of binding to the target, except for MPSIL0294-V532C which identified 2.3%. SMALPs on the other hand, identified 34% (AcrB), 26% (NupC) and 74% (VcCNT) of the total number of DARPins in their respective selection outputs, as 'hits'. This therefore suggests that in the comparison of different methods of solubilisation (DDM, SMALP or nanodisc), SMALPs are the best method for presenting proteins to a library of potential binding partners. Evidence for this conclusion

can be seen no matter the protein which has been encapsulated within the SMALP as both NupC and VcCNT follow this trend with a high number of DARPins which apparently bind to these targets being identified. The drastic difference in the number of DARPins identified between NupC and VcCNT also suggests that either VcCNT is better able to pack on the surface than NupC or that VcCNT has surface epitopes which are more accessible to the DARPins than NupC.

The amino acid sequence of the DARPins isolated against AcrB SMALPs have been aligned with one another and compared to a third DARPin, B10 from the detergent solubilised AcrB plate which failed to bind (**see Figure 4.9**). A complete alignment of all of the binding DARPins from the AcrB nanodisc selection is presented in chapter 5. The alignments highlights several positions within the DARPin variable region which show a degree of preference toward certain amino acid residues, for example in position 57 where the majority of the DARPins have an asparagine residue and those that don't have a tyrosine or histidine. The most prominent example of this conservation however is the 69th residue, which in the overwhelming majority of the DARPins aligned is an isoleucine and the majority which do not have this residue have a valine in its place, with the exception of A8 which has an alanine instead. Thereby suggesting that an aliphatic residue is essential for binding in this position, with the exception of H2, which has a serine residue.

Some degree of DARPin cross-reactivity was tested as shown in chapter 5, as DARPins selected against NupC, VcCNT and AcrB SMALPs (as well as AcrB nanodiscs) were tested in ELISAs performed on these proteins detergent solubilised counterparts and still showed an ability to bind to the various proteins (see **Chapter 5**).

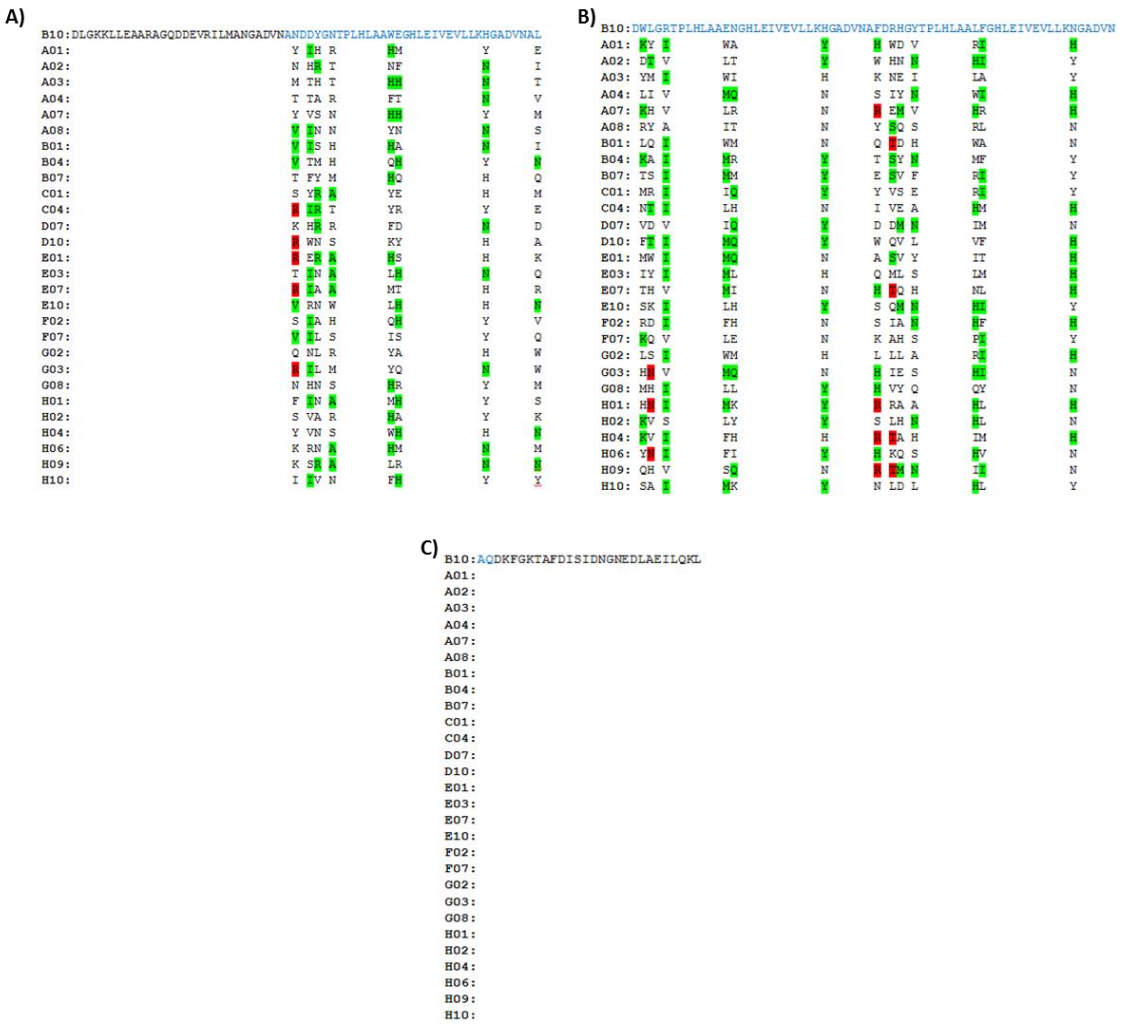


Figure 4.9: <u>AcrB SMALP DARPin binder's amino acid sequence alignment</u>: The amino acid sequence alignment of DARPins that showed an ability to bind to AcrB SMALPs aligned against B10 which was selected against detergent solubilised AcrB. Only residues in the variable regions of the DARPins have been shown with the exception of B10. The N and C terminal caps are shown in black text of B10 while the variable region is in blue. The most common residues in each position across all of the DARPins are highlighted in green. In an instance in which two residues have the same frequency the second has been highlighted in red.

Very recently the use of crystallisation chaperones has yielded a high profile crystal structure from the same protein family as MPSIL0294. The structure was of one of its prokaryotic homologues from *Staphylococcus capitis* and was determined using co-crystallisation with a nanobody which was isolated against it. However in order to obtain the crystals several residues were removed from the proteins N-terminal and the nanobodies were generated via the immunisation of llamas (Ehrnstorfer, Geertsma et al. 2014). The use of DARPins as crystallisation chaperones has some precedence and have been used in order to solve the structures of a wide range of proteins included some membrane embedded (Sennhauser and Grutter 2008).

In regard to solving crystal structures the work presented here has several benefits. Firstly, SMALPs are thought to entrap full length membrane protein within a lipid environment more reminiscent of its natural surroundings (Postis, Rawson et al. 2015). Secondly, the DARPins are easy to produce without the need of animals making it more widely accessible unlike antibody based binding partners such as nanobodies (Ehrnstorfer, Geertsma et al. 2014).

Overall it tentatively appears as though SMALPs are superior to nanodiscs, however further validation is required to confirm this. Outside of phage display the physical characteristics of SMALPs (i.e. their stability and storage) make them a more attractive prospect to work with in the future. Experimental and observational evidence has shown that nanodiscs have a tendency to destabilise when stored at -80°C and when stored at 4°C for an excess of 2 weeks. SMALPs on the other hand have shown a high degree of stability after long term storage at -80°C and several freeze-thaw cycles. Indeed even the procedure by which both of these protein formats are made show the same pattern, as SMALPs are made in an identical manner as detergent micelles while nanodiscs require a minimum of 1 week in order to remove detergent. Although a major drawback of the use of SMALPs as shown here is that when constructed with a protein which neither expresses nor purifies to a high yield, many impurities can co-purify. As such it is unclear whether or not these impurities are able to immobilise on the streptavidin surface particularly for samples which underwent aspecific biotinylation such as AcrB, VcCNT and NupC. Therefore the DARPin selection outputs from the SMALPs will require further validation to ensure that they only bind to the target instead of impurities or the SMA polymer itself. This validation should use a third phage ELISA in the presence of immobilised detergent solubilised protein. Although it is also possible that some DARPins will bind to the MSP of the nanodiscs, therefore the outputs from the nanodisc selections should be subjected to a third phage ELISA with immobilised ‘empty’ nanodiscs.

Chapter 5

A Characterisation of the DARPins identified after phage display

5.1 Introduction

There are several examples in literature in which DARPins have been successfully selected against membrane proteins (Huber, Steiner et al. 2007, Steiner, Forrer et al. 2008). Many have been identified with a capability to bind therapeutically relevant receptors such as the epidermal growth factor receptor (EGFR) which is overexpressed in many cancers and plays a role in cell proliferation. The high affinity of DARPins lead to the creation of bispecific derivatives capable of reducing the receptors presence on the cancer cell surface by inhibiting receptor recycling. Bispecific DARPins were constructed using two DARPins which target different EGFR epitopes, with either a leucine zipper or a flexible linker (Boersma, Chao et al. 2011). Similarly Eggel et al. also formed bivalent DARPins in an identical manner however in this instance two identical DARPins were used (Eggel, Baumann et al. 2009, Boersma, Chao et al. 2011). These bivalent DARPins have been shown to target multiple epitopes while retaining all of biophysical benefits which monovalent DARPins exhibit (Eggel, Baumann et al. 2009). For EGFR and IgE receptors, these DARPin derivatives have shown activities which rival (and in some cases surpass) therapeutic antibodies (Boersma, Chao et al. 2011) (Eggel, Baumann et al. 2009). Although bivalent DARPins would only be therapeutically relevant when multimeric proteins are targeted. DARPins have also been conjugated with small molecule cytotoxins thereby increasing their pharmacokinetics and uptake. DARPin against Epithelial Cell Adhesion Molecule (EpCAM) were conjugated with mouse serum albumin to increase their half-life and to the antimitotic agent monomethyl auristatin F (MMAF) to achieve the targeted inhibition of cancer cell proliferation (Simon, Frey et al. 2013).

Typically after a screen phage ELISAs are often used in order to confirm the ability of the selected DARPins to bind to the target of interest. These DARPins are selected arbitrarily after several rounds of display and are sequenced after this point. Purification of DARPins tends to utilise IMAC followed by further characterisation using techniques such as surface plasmon resonance (SPR), analytical HPLC or flow cytometry (Steiner, Forrer et al. 2008, Eggel, Baumann et al. 2009, Stefan, Martin-Killias et al. 2011). Their expression often requires subcloning into an expression vector either to avoid the production of phage coat proteins or to maximise the yield of DARPin (Steiner, Forrer et al. 2008, Boersma, Chao et al. 2011, Stefan, Martin-Killias et al. 2011). A minimum of three cycles are typically used in a selection screen thereby resulting in higher quality selection outputs (Stefan, Martin-Killias et al. 2011). Although Steiner et al. did report a monomeric DARPin against the constant domain of human IgG1 with a K_D of 140 nM (Steiner, Forrer et al. 2008). DARPins selected in this manner have been used in several

subsequent applications such as those described above and co-crystallisation trails which has shown to be very effective with DARPins (see chapter 1).

Chapter 3 and 4 described the use of two cycles of phage display with a naïve DARPin library against a panel of membrane proteins, as well as the subsequent validation tests performed on the various selection outputs. In this chapter, 21 DARPins from these validated selection outputs were selected and subjected to further study via expression, purification and the use of a DARPin binding assay. The DARPins include 2 which showed activity in the presence of detergent solubilised, labelled MPSIL0294-V532C, 6 isolated against AcrB (2 against the SMALPs and 4 against the biotinylated nanodiscs), 5 DARPins isolated against the VcCNT SMALPs and 8 DARPins which were able to bind to NupC SMALPs. These DARPins were selected based upon their A_{450nm} readings produced after the two phage ELISAs shown in **Table 3.4, Table 4.5** and **Table 4.6.** While these previous assays used entire phage particles presenting the DARPins upon their surface, the assays described in this chapter shall use purified DARPin alone. These assays shall confirm the DARPins ability to bind their target and attempt a rough estimate at the affinity of the DARPin for their target. DARPins G6 and E6 will also be introduced in stopped flow experiments with MPSIL0294-V532C in order to investigate if they have any functional effect on the protein.

5.2 Sub-cloning DARPins selected against a panel of membrane proteins

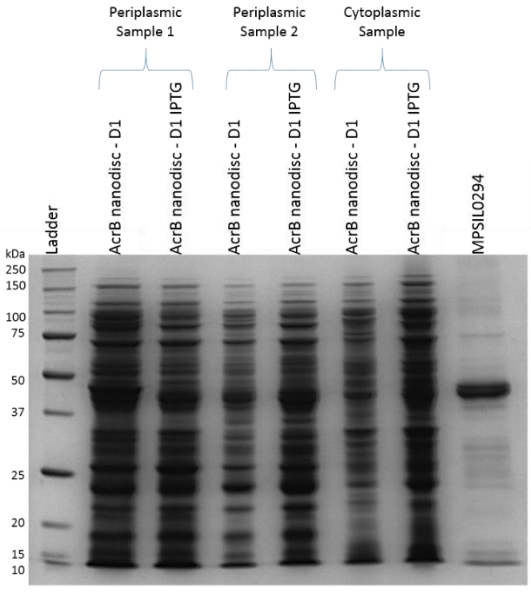
Typically DARPins produced for phage display are expressed with an N-terminal signal recognition peptide, DsbA and a C-terminal amber stop codon between the DARPin gene and the phage coat protein (see chapter 1).

Efforts to express DARPins in Medimmune have found that that high aeration of TG1 cells results in high levels of DARPin expressed and released into the growth medium. This was attempted initially, however the procedure failed to produce any extracellular DARPin. Therefore expression was subsequently attempted in BL21 gold cells instead as they are a non-suppressor strain and therefore should only be able to produce soluble DARPin protein not fused to the M13 major coat protein - P3. The amber stop codon in the pC6 plasmid which houses the DARPin open reading frame (**see section 2.5.6**) is responsible for this. This concept was tested by focussing attempts at expression on a single arbitrarily selected (D1 selected against AcrB biotinylated nanodiscs) DARPin. The plasmid DNA was first isolated from the TG1 cells (**see section 2.3.8**) and subsequently used to transform BL21 gold cells (**see section 2.3.7**).

Expression of the DARPin was performed via IPTG induction in 50 mL SB medium supplemented with 100 µg/mL carbenicillin as described in **section 2.4.2.** The cells were incubated at an elevated RPM (300 rpm) in order to once again subject the cells to the high aeration conditions that MedImmune have found to facilitate DARPin production. The transformed BL21 gold (DE3) cells were then subjected to a periplasmic extraction (see **Section 2.5.7**).

As can be seen in **Figure 5.1** this expression protocol produced an immunogenic band in the cytoplasmic sample with a molecular mass between 15 and 25 kDa when the cells have been treated with IPTG. This band appears very faintly in the non-IPTG band and not at all in the periplasmic fractions. Although the mass of this band is within the correct range in regard to the predicted mass of a DARPin (16.6 kDa), its proximity to the 25 kDa marker is somewhat unexpected. The reason for this increase in mass is not clear, it could possibly be due to the presence of the DsbA signal sequence which would increase the predicted mass of the DARPin to 18.8 kDa. The fact that the DARPin failed to translocate to the periplasm provides some evidence for this. This failure to translocate also explains why the DARPin unexpectedly is only observed in the cytoplasm. Although the expression level is also very low and is barely detectable by western blot.

A)



Coomassie Stained Gel

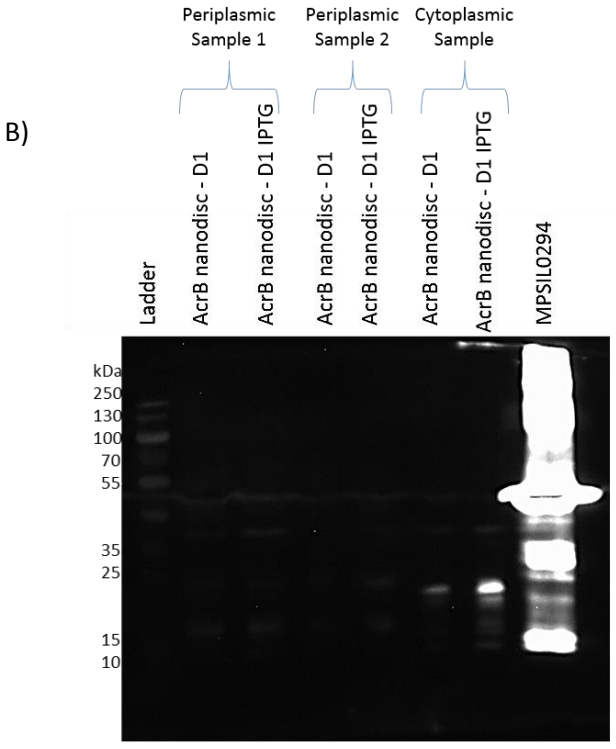


Figure 5.1: <u>Periplasmic extraction of BL21 gold (DE3) cells expressing DARPins:</u> Samples taken from the different fractions produced after a periplasmic extraction was performed on cultures of BL21 gold were analysed via Coomassie stained 12% SDS-PAGE and western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP. The Bl21 gold cells were transformed with plasmid pC6 in which DARPin D1 isolated against AcrB nanodiscs was encoded, samples treated with IPTG were compared with samples that were not. MPSIL0294 was used as a positive control of the blot.

Because attempts to express the DARPin in this way failed, it was decided to sub-clone the DARPin gene from pC6 in **Figure 2.1** into the expression vector pET16B shown in, thus removing the potential problems associated with the DsbA signal sequence and amber stop codon. In order to achieve this, an N-terminal NdeI and a C-terminal BamHI cut site were inserted into the pC6 plasmid via PCR (**see section 2.3.1**) using the following primers:

- Forward primer (introducing the **<u>NdeI</u>** site) pET16b_DP_Fwd:
  - 5' – CGAT**<u>CATATG</u>**GATCTGGGAAAAAAACTGCTGGA – 3'
- Reverse primer (introducing the **<u>BamHI</u>** site) pET16b_DP_Rev:
  - 5' – ATCG**<u>GGATCC</u>**TCACTACAGTTTCTGCAGG – 3'

The amplified DARPin was inserted into an empty pET16B vector via double restriction digest with NdeI and BamHI followed by ligation (**see section 2.3.3** and **section 2.3.6**). After transformation in Omnimax™2 in order to amplify and isolate the DNA, the correct

construction of each resultant plasmid was analysed via double restriction digest with NdeI and BamHI followed by electrophoretic analysis on a 1% agarose gel (**see Figure 5.2**). Empty pET16B was used for the subcloning therefore upon restriction analysis the presence of a band at approximately 500 bp is indicative of the insertion of the DARPin gene. **Figure 5.2** shows the agarose gel produced after 20 of the 21 selected DARPins were analysed in this manner. This 500 bp gene is present in 19 of the 20 DARPins and the subcloning of DARPin C1 was not successful. The successful subcloning of the remaining DARPin (H10 selected against NupC SMALPs) was confirmed via sequencing. Attempts to express each were then performed in BL21 (**see section 5.3**).

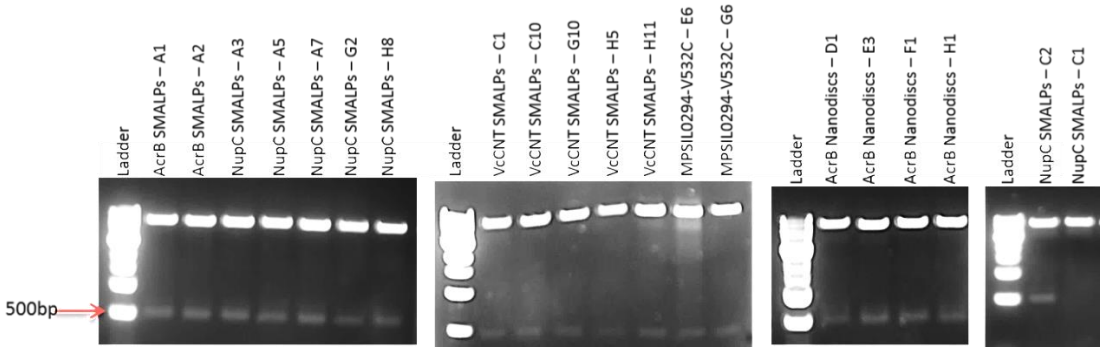


Figure 5.2: <u>NdeI/BamHI double restriction analysis of pET16B-DARPin expression vectors:</u> Following the experiments which sub-cloned all 20 DARPins into pET16B the resultant plasmids were analysed via electrophoretic separation on a 1% agarose gel after double restriction with NdeI and BamHI. Visualisation of the gels was performed in a Gbox from SynGene.

5.3 The Expression and purification of the DARPins

An initial expression test using the DARPin_pET16B constructs was performed in 5 mL cultures of BL21 cells following the same (albeit scaled down) methodology described in **section 2.5.4.** Whole cells samples of each expression were then pelleted and resuspended in 1 mL 2x PBS prior to further analysis. These samples were then diluted by mixing 1µL of them with 9 µL of 2x PBS and 2.5 µL of loading buffer (**see Table 2.8**) before incubation at 95°C for 10 minutes. The samples were then analysed via 12% Coomassie stained SDS-PAGE (**see Figure 5.3**) run in 1x Tris-glycine running buffer (**see Table 2.9**). A positive control DARPin prepared by MedImmune was run on the gel alongside the DARPins while uninduced BL21 cells were run as a negative control. As previously mentioned the predicted mass of the DARPins is approximately 16.6 kDa and as the resultant gel shows each DARPin has successfully been expressed. The majority of the relevant bands have migrated to the bottom of the SDS-PAGE

gel and have resolved at different molecular weights and different expression levels to one another. Despite this, these bands are most likely the DARPins as they are the only band not present in the uninduced BL21 negative control. The DARPin C1 selected against NupC SMALPs did not express to a significant level and was therefore no longer studied leaving only 20 DARPins remaining for further analysis. This was expected as the double restriction digest described above indicated the sub-cloning failed (**see Figure 5.2**). The sub-cloning and expression of C1 was repeated separately and results are not included here.

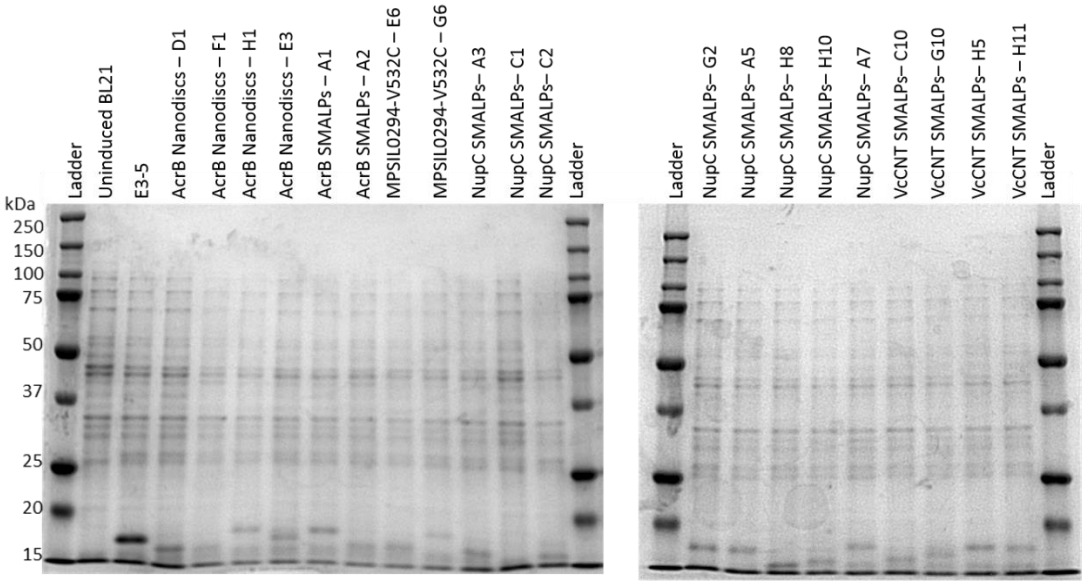


Figure 5.3 <u>Small scale expression test of 20 binding DARPins:</u> Each DARPin was expressed in the E.coli strain BL21 via IPTG induction therefore uninduced BL21 has been added as a negative control. A previously characterised DARPin was provided by MedImmune designated E3-5 and was used as a positive control

A larger scale expression and purification of a single DARPin was attempted for a 500 mL BL21 culture. The DARPin was purified via IMAC with cobalt resin as described in **section 2.5.4** with the inclusion of an initial wash with 2x PBS supplemented with 10 mM imidazole followed by a wash using 20 mM imidazole**.** These samples were also analysed by 12% SDS-PAGE gel followed by staining with Coomassie blue (**Figure 5.4).** A single protein band resolved just above the 15 KDa marker corresponding to the DARPin with a predicted mass of 16.6 kDa. No DARPin was present in the flow through confirming all expressed DARPin bound to the cobalt resin. However, some DARPin is eluted in the wash samples with 10 and 20 mM imidazole. The

washes seem ineffective and are not required considering the purity of the eluted protein. However, a single superfluous protein might have co-purified, as shown by the unidentified protein at a Mr <15 kDa in the SDS-PAGE. This could also be a degradation product or an SDS-PAGE artefact.

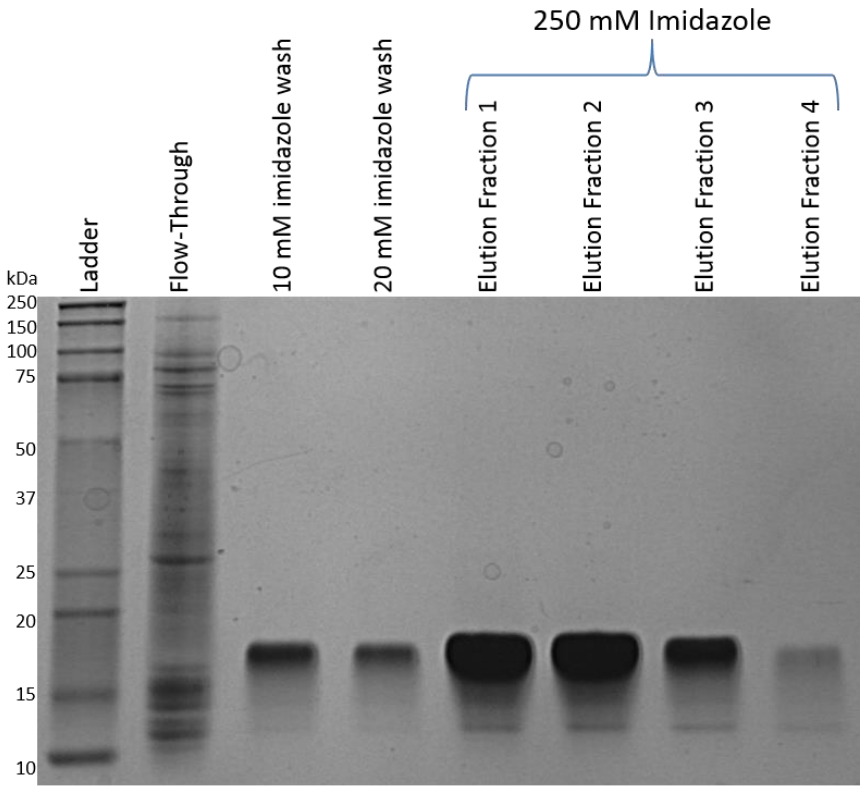


Figure 5.4 <u>Initial purification of DARPin C1 against VcCNT SMALPs:</u> An initial attempt to purify a DARPin using C1 selected against VcCNT SMALPs with samples taken throughout the purification analysed via 12% SDS-PAGE gel followed by staining with Coomassie blue. The flow through sample is a measure of the amount of protein which failed to bind to the resin and was taken from the solution realised once the resin was packed into a disposable column. All washes were performed by gravity flow and the peak elution fractions were collected and concentrated in a vivaspin concentrator with a 10 kDa molecular weight cut off.

Following the success of the VcCNT SMALP DARPin C1 purification, large scale expressions were repeated for all 20 of the successfully expressed DARPins. The peak fractions of each DARPin were pooled and concentrated in a vivaspin concentrator with a 10 kDa molecular weight cut off. Samples of each purified DARPin was analysed via 12% SDS-PAGE followed by Coomassie staining in 1x MOPs running buffer as opposed to the buffer used in **Figure 5.3** (**see Table 2.10**). Running SDS-PAGE in MOPS reduces the migration of the low molecular weight

protein, improving their separation. **Figure 5.5** shows that each DARPin is estimated to be more than 90% pure and suitable for further study. Some DARPins, such as A1 (against AcrB SMALPs), E3 (against AcrB nanodiscs), C1 (against VcCNT SMALPs) and E6 (against detergent solubilised MPSIL0294-V532C) amongst others produce a protein band at approximately 37 kDa. Some of the DARPins have also produced a second band in very close proximity to the putative DARPin band, e.g., F1 (against AcrB nanodiscs) and H8 (against NupC SMALPs) although is somewhat common in DARPins with more than two internal repeats and is often caused by incomplete denaturation by SDS (Steiner, Forrer et al. 2008). There is also evidence of co-elution with some contaminants, for instance for A5 (against NupC 5) and G10 (VcCNT SMALPS).

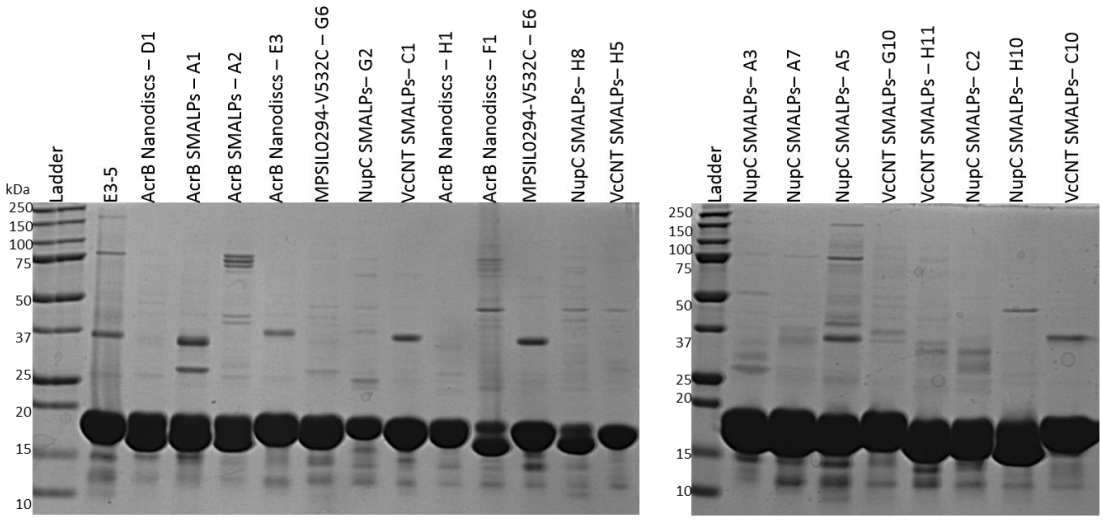


Figure 5.5: <u>Purification of 20 DARPins from the pET16B expression vector:</u> Each DARPin was purified via IMAC on cobalt resin, the peak fractions were then concentrated with a vivaspin concentrator with a 10 kDa molecular weight cut off and analysed on a 12% SDS-PAGE gel followed by extensive Coomassie staining. DARPin E3-5 was previously characterised by MedImmune and was used as a positive control for the DARPins

5.4 An Alignment of the protein binding DARPins Sequences

The sequences for the 20 DARPins selected for further study have been aligned and analysed in order to highlight any commonality between the sequences. In the sections below the alignment will be discussed separately for the different targets. Analysis was done in an identical manner as for DARPins E6 and G6 (selected against detergent solubilised MPSIL0294-V532C) in chapter 3 (see **Figure 3.12**). Once again residues in the variable regions are colour coded, while conserved residues such as the aspartic acid residue in position 32 and the glycine in position 35 have been left colourless (see **Figure 5.6**).

Generally, between the sequences of each DARPin there is slight variation observed at the end of 'GHLEIVEVLLK' repeats. The next residue is always one of three residues: a tyrosine, a histidine or an asparagine with tyrosine being the most common. The reason for this diversity is not immediately clear.

5.4.1 DARPins selected against AcrB Nanodisc

Four DARPins selected against AcrB nanodiscs were taken for further study, DARPins D1, F1, H1 and E3. These DARPins do not show a great degree of sequence commonality between them (**see Figure 5.6**). This is especially apparent between positions 97 and 100 in which there is a large cluster of alternative residues between the different sequences. In one case, even a non-variable aspartic acid residue at position 98 is mutated to a histidine in DARPin H1. Indeed the only similarity in this region exists at position 99 between D1 which has an aspartic acid residue and E3 which has a glutamic acid residue. There are however individual positions which show commonality either in the majority of the DARPins or in all four, for example the first variable region of all four DARPins begins with an arginine residue at position 31, similarly at position 66 in which three of the DARPins have a threonine residue. Despite this, it is quite clear that DARPins D1 and F1 are the most similar to one another out of all the possible DARPin combinations with a total of 9 variable residues which are either identical, have sidechains with the same R group or have similar physical characteristics.

Figure 5.6 <u>Amino acid sequence alignment of the DARPins generated against AcrB nanodiscs which were selected for further study:</u> The sequences of the four DARPins selected against AcrB nanodiscs constructed with a biotinylated membrane scaffold protein have been aligned in order to highlight any common themes between the sequences. Amino acid residues in variable regions of the DARPins have been colour coded in regard to their similarity to the residues in the corresponding positions in the other DARPins while conserved residues have been

removed. Amino acids highlighted in red show no similarity with any of the other corresponding residues, those in bright green are identical, yellow have either the same R group or similar physical characteristics, similarly those in dark green share these characteristics as those in green. Residues in grey are identical to one another but different from those in green in at the same position and finally those in pink are similar to one another but not those in yellow in the same position.

5.4.2 DARPins selected against AcrB SMALPs

Two DARPins which were identified as AcrB SMALP binders were chosen for further study and were aligned in an identical manner as above. These two DARPins showed a very low amount of similarity between one another with 14 out of the 18 variable amino acid residue positions occupied with residues with no link to one another (**see Figure 5.7**). As only two sequences are aligned, no further meaningful comparison is possible. These two DARPins are not indicative of the entire output which showed an ability to bind to this target which were aligned in chapter 4 (see **Figure 4.9**). Indeed, the majority of the variable residues in A1 are typical of the AcrB SMALP DARPin output.

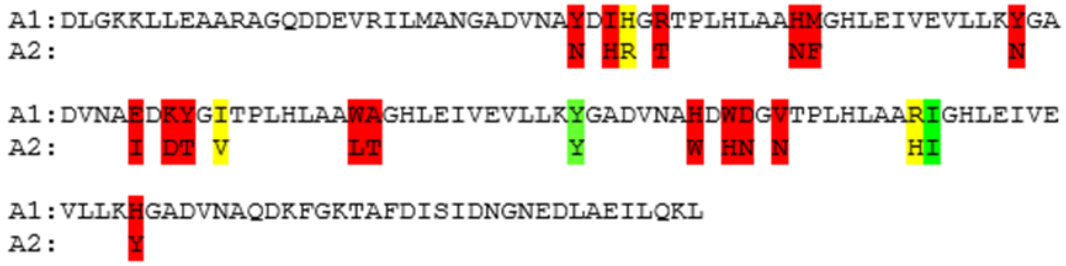


Figure 5.7: <u>Amino acid alignment of the DARPins generated against AcrB SMALPs selected for further study:</u> DARPins A1 and A2 selected against AcrB SMALPs were chosen for further study therefore their sequences were aligned so that any similarity between the DARPins could be highlighted. As before, residues in the variable regions were colour coded in regard to their similarity to one another with red being used to classify residues which show no similarity, bright green denoting identical residues and yellow denoting similar residues.

5.4.3 DARPins selected against NupC SMALPs

Due to the relatively high number of DARPins that were aligned, it is possible to make a more conclusive analysis of the DARPins in the variable regions in order to determine amino acid preference **(Figure 5.8)**. One the most conserved positions in the DARPin is position 36 in which 5 of the 8 DARPins have an alanine residue, C1 has a leucine residue (an aliphatic residue like alanine) and A7 and H10 both have serine residues. A second position with this extremely high level of sequence similarity is at position 76 at which 5 of the DARPins have a methionine residue while G2 and H8 both have aliphatic residues in isoleucine and alanine

respectively. Finally A5 has a phenylalanine in this position which, while being an aromatic amino acid is a derivative of alanine.

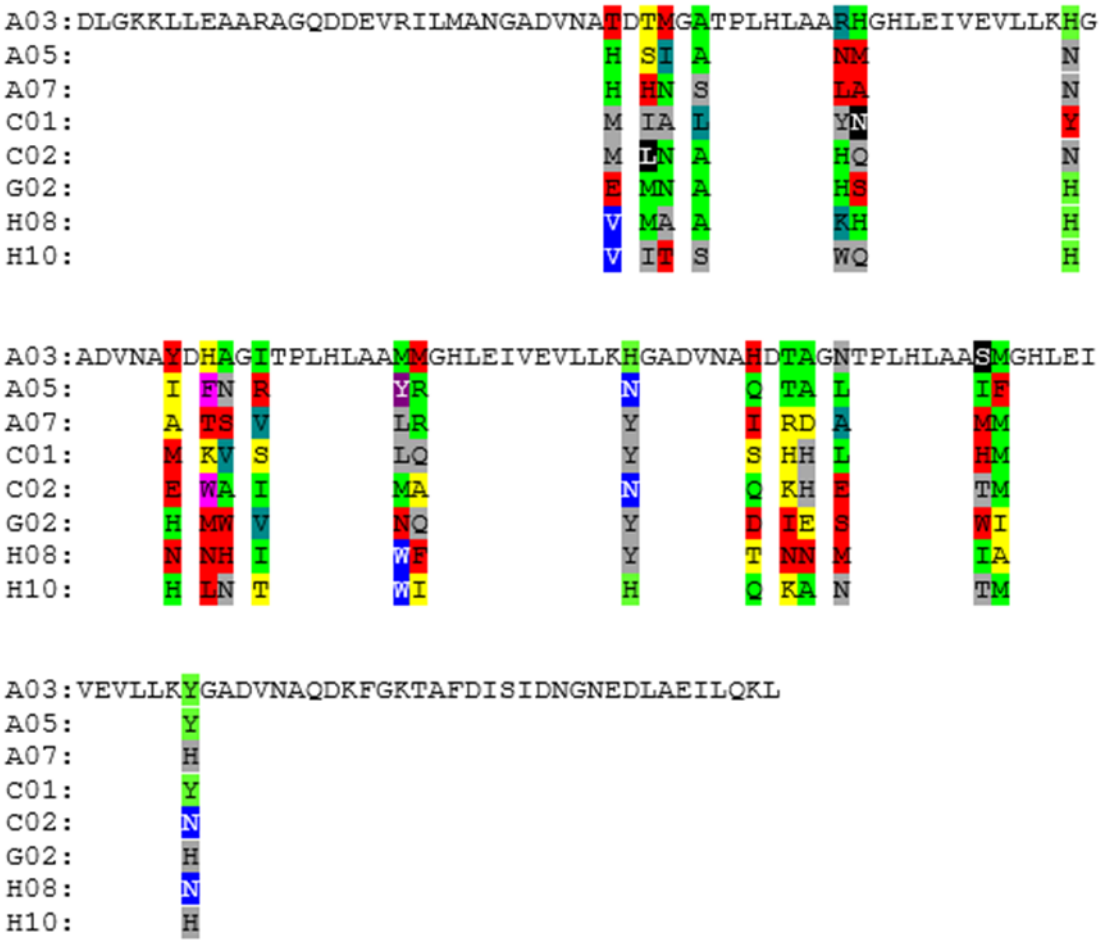


Figure 5.8: <u>Amino acid sequence alignment of the DARPins generated against NupC SMALPs selected for further study:</u> The sequences of the 8 DARPins selected against NupC SMALPs were aligned so that the level of similarity between them could be investigated. Residues in the variable regions of the DARPins have been colour coded in relation to their similarity to one another at particular positions within the sequence. Residues highlighted in red are not related to any of the corresponding amino acids in that position, residues in yellow share the same amino acid characterisation with one another, similarly residues in pink are also from the same characteristic amino acid group separate from those in yellow. Residues in bright green are identical to one another and those in dark green are related to them. Residues in grey are also identical to one another but are not related to those in green, only residues highlighted in black are related to these. Finally residues in blue are identical to one another but are unrelated to those in green and grey instead only residues in violet are related to them.

5.4.4 DARPins selected against VcCNT SMALPs

Within the five DARPins selected against the VcCNT SMALPs, the sequence of C10 was unfortunately not available and therefore could not be aligned with the other VcCNT SMALP DARPins. The other four DARPins were successfully aligned, the results of which can be seen in **Figure 5.9.** The four DARPins show several positions which have a high degree of sequence similarity between one another. For example with the exception of G10, the 69th residue in the DARPins is an isoleucine and both of these residues are aliphatic amino acids. Similarly in position 67 all the DARPins except H11 have an asparagine residue or position 111 in which all of the DARPins with the exception of G10 (which has a histidine) have a valine residue.

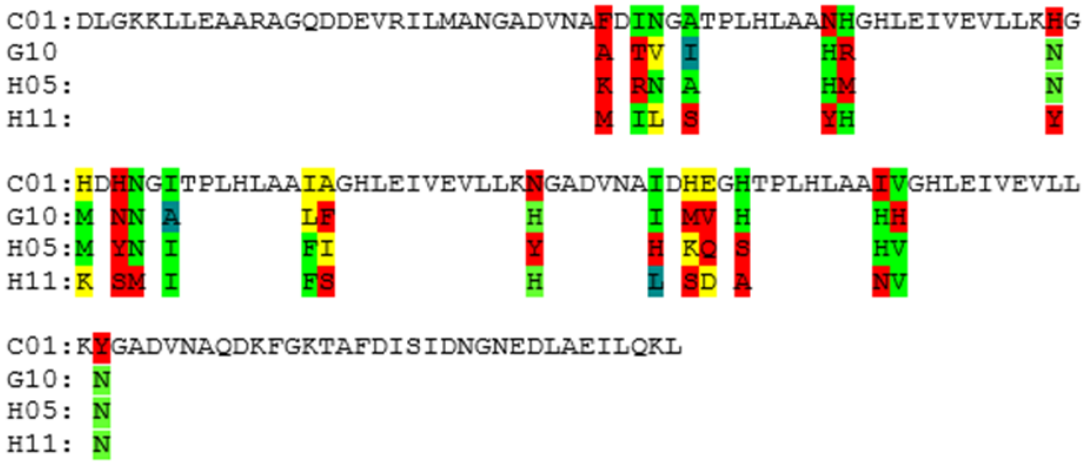


Figure 5.9: <u>Amino acid sequence alignment of the DARPins generated against VcCNT SMALPs selected for further study:</u> The sequences of the DARPins selected against VcCNT SMALPs and chosen for further study were aligned. DARPins C10 which was also further characterised did not have an available sequence and so was not included in this particular alignment. The residues in the variable positions were colour coded in relation to their similarity to one another. Residues in red were not related to the other amino acids in the same position, residues in yellow were from the same amino acid group. Residues in bright green are identical to one another and those in dark green are from the same amino acid group as these residues.

5.5 Confirming the DARPins ability to bind via ELISA

To test whether the various DARPins are able to bind to their respective targets, three ELISAs were performed in which the DARPins were tested in duplicate with a serial dilution ranging from 10 µM to 1 pM. In each case 10 µg of target membrane protein was adsorbed onto a solid surface. Adsorption was used due to the fact that a polyhistidine tag is present on all of the target membrane proteins as well as the DARPins, therefore his-tag antibodies could not be used to detect the DARPin binding and an alternative method was required. To this end, DARPins were biotinylated with NHS-biotin (Thermo Scientific) and probed with streptavidin-

HRP (**see section 2.10**) and A_{450nm} readings were recorded after treatment with TMB and 0.5 M H_2SO_4 the absorbance reading at 450 nm was plotted for each DARPin concentration. Due to the aspecific nature of the biotinylation it is not known what affect it will have upon DARPin binding. It could be that biotinylation of the DARPin reduces its affinity for its target. It is thought however, that biotinylation leaves enough of the DARPin unaffected that it should retain its stability and its ability to bind target epitopes. During the ELISA procedure both the initial protein adsorption and treatment of the immobilised target with DARPin was carried out in the presence of DDM in an attempt to retain the targets native confirmation. Although the possibility of denaturation upon adsorption on the maxisorb plate cannot be excluded, the fact that these ELISAs were performed on detergent solubilised protein does help to occlude the possibility of false positives in the selection output, in a similar manner as the use of Tween-20 in a typical ELISA.

The majority of the 20 DARPins that were tested in this manner showed a significant difference when in the presence of their immobilised targets as determined with a T-test against the negative control (negative controls consisted of the DARPin under test without any membrane protein target present). Even at the highest concentration tested (10 µM), DARPin binding failed to saturate and no 'plateau' was observed (as one expects on an IC_{50} graph). Unspecific 'background' was also monitored by performing the ELISA without DARPins. The background absorbance readings were subtracted from those of the DARPins and negative controls.

5.5.1 DARPins selected against detergent solubilised MPSIL0294-V532C

Two DARPins were selected against chemically biotinylated MPSIL0294-V532C after two rounds of phage display. They were subsequently analysed via ELISA, the resultant data of which is shown in **Figure 5.10**.

The ELISAs show that both DARPins are able to bind to MPSIL0294-V532C, where E6 gives higher absorbance values (note different axis in Figure 5.12). The data produced by G6 exhibits larger standard deviations than E6, which mean that less of the data points are significantly different from the negative control. Nonetheless, a very similar trend are found for E6 and G6. As previously mentioned the data fails to plateau which therefore makes it not possible to determine the IC_{50} for either E6 or G6. Instead a lower limit for the IC_{50} can be estimated which is 100 nM.

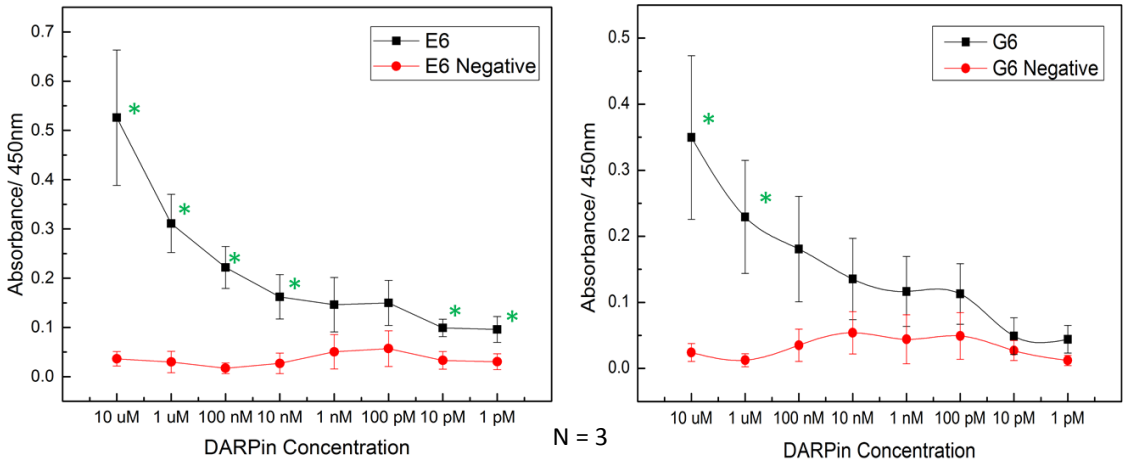


Figure 5.10: <u>DARPin binding ELISA performed against DARPins selected against DDM solubilised MPSIL0294-V532C:</u> Purified samples of DARPins E6 and G6 were subjected to aspecific biotinylation with NHS-biotin prior to their use against 10 µg samples of detergent solubilised unlabelled MPSIL0294-V532C, adsorbed onto the surface of Nunc Maxisorp® plates. Samples were probed with 10 µM to 1 pM DARPin prior to analysis with Streptavidin-HRP. Absorbance readings at 450 nm were recorded after the plates incubated in TMB solution for 2 mins followed by treatment with 0.5 M H_2SO_4. Each data point on the graph represents the average of six tests subtracted from a background reading which was measured using 10 µg BSA in the presence of no DARPin. The negative control was DARPin tested in blank wells with no target protein. The standard deviation of each data point was also calculated and the statistical significance was determined by a one tailed T-test. Each significantly different data point has been highlighted with a green asterisk.

5.5.2 DARPins selected against AcrB SMALPs and Nanodiscs

As for NupC-V532C, samples of detergent solubilised AcrB were aspecifically adsorbed onto the surface of a Nunc Maxisorp® plate and probed with DARPins. Two of these DARPins (A1 and A2) were selected against AcrB SMALPs and the ELISAs (**see Figure 5.11**) show that despite the lack of commonality in the sequences of these two DARPins (**see Figure 5.7**) both produced similar binding curves. However neither DARPin saturated binding at 10 µM but both have similar adsorption values at this concentration. As for MPSIL0294-V532C, negative controls were measured without target protein and the data sets were analysed for a significant difference using a two tailed T-test. Despite the similarities between the two traces. In the A2 trace the data points might start to level off typical of an IC_{50} plot. However, the low absorbance values and standard deviations in the data points, makes an unambiguous assessment impossible and for both DARPins only a lower limit can be determined for the IC_{50} of 100 nM.

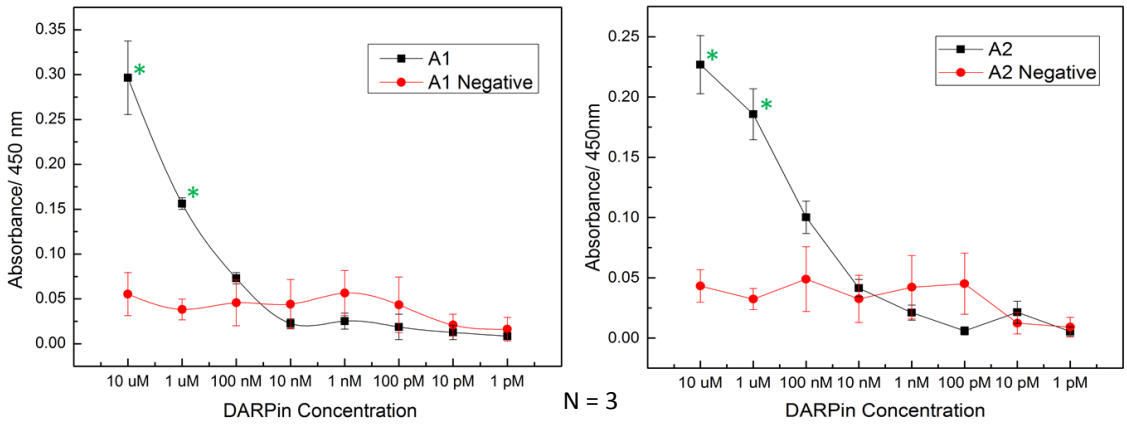


Figure 5.11 DARPin binding ELISA performed against DARPins selected against AcrB SMALPs: 10 µg of detergent solubilised AcrB was aspecifically adsorbed onto the surface of Nunc Maxisorp® plates and treated with a serial dilution of DARPin A1 and A2 from 10 µM – 1 pM. The DARPins showed an ability to bind AcrB SMALPs particularly tightly after two rounds of phage display and two rounds of subsequent validation. The plates were probed with streptavidin-HRP and the absorbance at 450 nm was analysed after incubation with TMB and subsequent exposure to 0.5 M H_2SO_4.The negative control for both ELISAs was the DARPin without the target protein present and all of the readings have been adjusted for background signals by subtracting the absorbance reading produced by BSA alone. Each ELISA was also repeated three times with duplicates of each concentration, error was calculated by the standard deviation of all six of these data points and the significance between the test data and negative control was calculated by a 2 tailed t-test. Data points which are significantly different from the negative have been highlighted with a green asterisk.

The traces produced by ELISAs performed on the DARPins selected against AcrB nanodiscs (see **Figure 5.12**) show that all of the DARPins are able to bind to detergent solubilised AcrB to some extent, however, the best binder when both statistical significance and A_{450nm} readings are taken into account is E3. When compared in terms of significance, D1 is the best as it is significantly different from the negative control at every concentration other than 100 nM, F1 is only significantly different from the negative control down to a concentration of 100 nM, H1s significance ceases after 1 µM, and finally E3 is significantly different from its negative control down to 10 nM. The highest absorbance reading however, shows that E3 performs the best out of the DARPins. When 10 µM of DARPins is used on the ELISA the reading produced by E3 is substantially higher (approximately twice as large) than D1, F1 and H1. Like all of the other ELISAs performed in this Chapter, negative controls were performed without membrane protein present. The negative controls of H1 and F1, however, suffers from a higher than usual degree of fluctuations. This effect is aggravated in the graph due to the small values of the absorbance readings in these results.

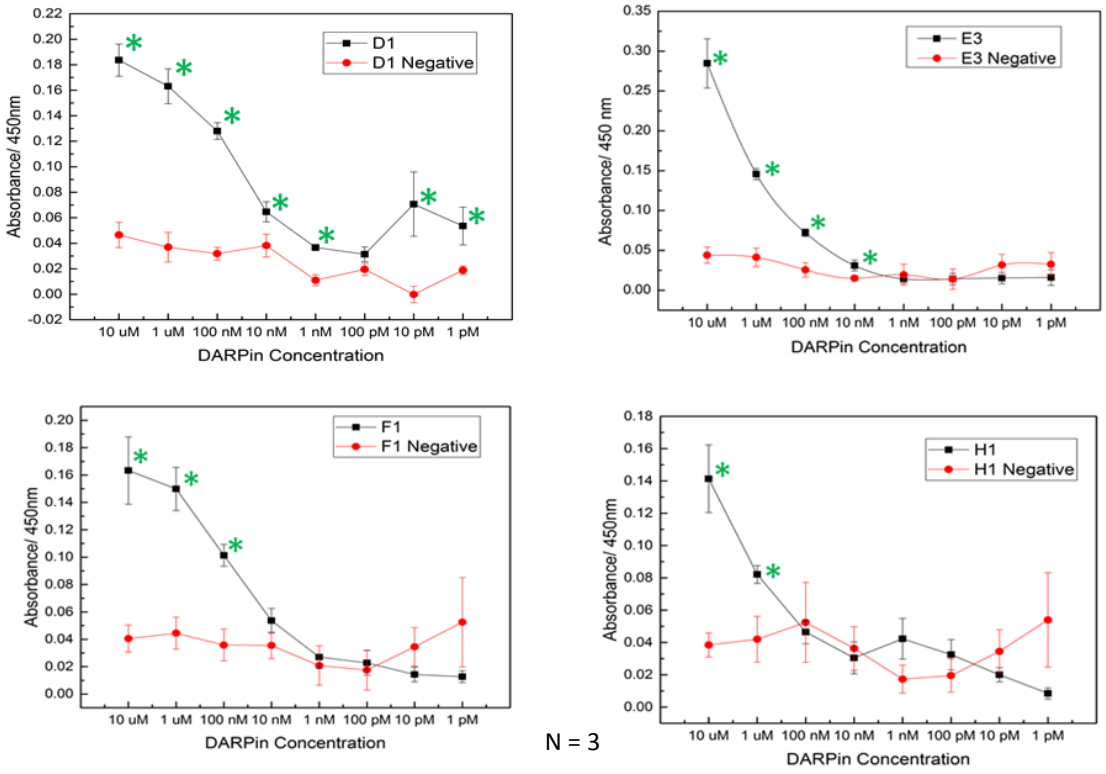


Figure 5.12: <u>DARPin binding ELISA performed against DARPins selected against AcrB nanodiscs:</u> This ELISA was carried out on 10 µg of immobilised AcrB solubilised with DDM using a serial dilution of DARPin from 10 µM to 1 pM. Each ELISA was repeated three times and each concentration was tested in duplicate, the negative control was the DARPin tested in the absence of AcrB and background noise was corrected for by subtracting all of absorbance readings from those produced by BSA alone. Statistical significance was determined by two tailed t-test, all data points which are significantly different from their corresponding negative control are highlighted with a green asterisk.

5.5.3 DARPins selected against NupC SMALPs

The traces produced by the DARPins selected against NupC SMALPs show that 5 of the 6 DARPins successfully bind to the target. DARPin A3 fails to follow the expected data trend but remains significantly different from the negative control throughout the serial dilution, suggesting a concentration independent binding although considering the A_{450nm} readings it is likely that this arbitrary binding is quite poor. All of the DARPins show a significant increase over the various negative controls at most concentrations, more so than the other DARPins which have been tested in this manner. The exception being H10, the negative control of which comes into closer proximity with the test data than any of the other NupC DARPins, thus suggesting that the DARPin does not bind to NupC particularly well. As with the other ELISAs, the data fails to plateau and therefore an affinity for the DARPins cannot be determined. Instead a lower estimate of 100 nM for DARPins C2, G2, H8 and H10 can be given. The traces produced by DARPins A3 and A5 however do not have allow a clear estimate to be given.

When the data trends along with the absorbance readings are considered, C2 and G2 appear to be the most promising DARPins and could be used in co-crystallisation trails (see **Figure 5.13**).

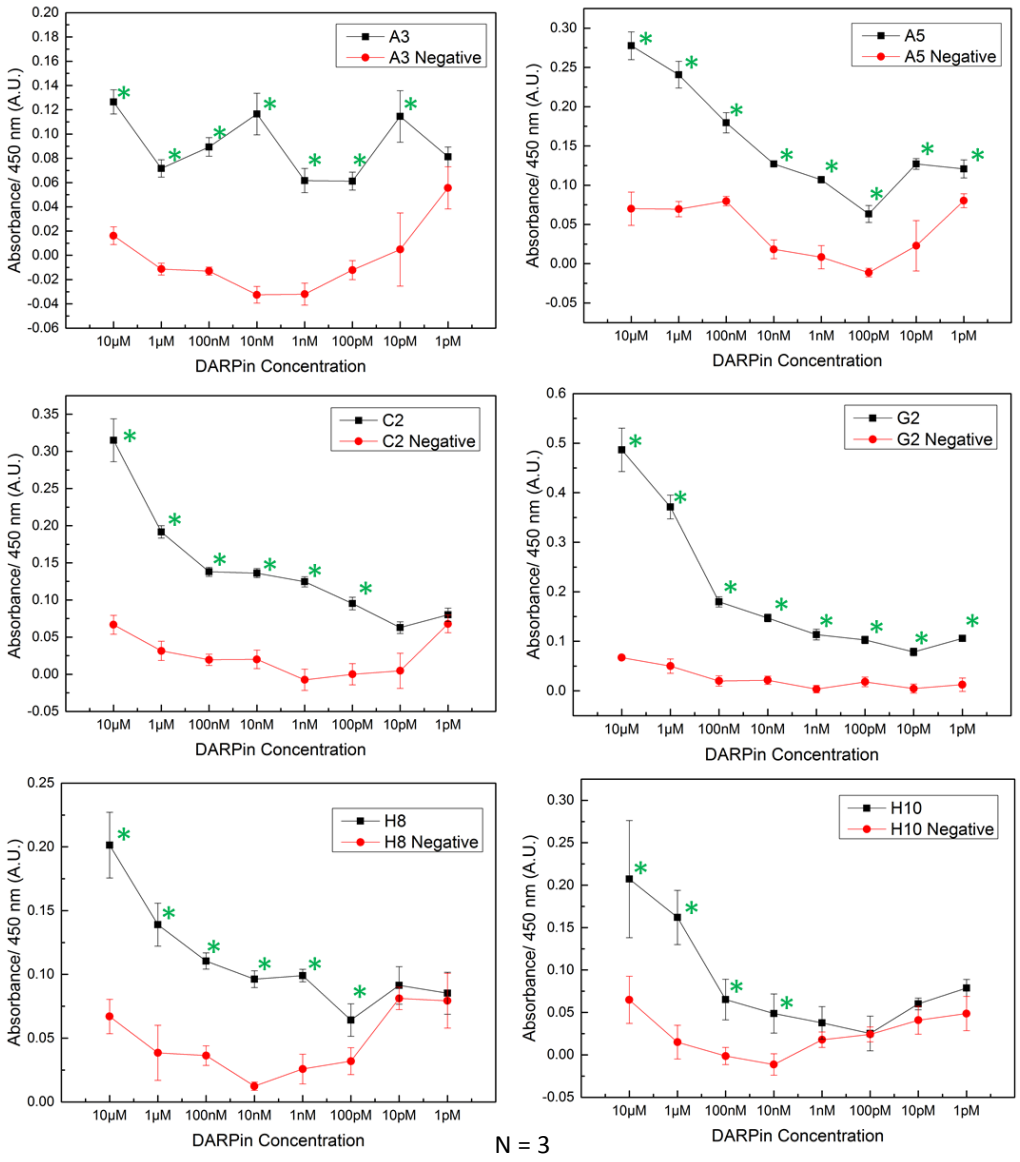


Figure 5.13: <u>DARPin binding ELISA performed against DARPins selected against NupC SMALPs:</u> This ELISA was carried out on 10 µg of immobilised NupC solubilised with DDM using a serial dilution of DARPin from 10 µM to 1 pM. Each ELISA was repeated three times and each concentration was tested in duplicate. The data has been compared with a negative control which consisted of the DARPin tested in the absence of NupC and background noise was corrected for by subtracting all of the absorbance readings from those produced by BSA alone. Statistical significance was determined by two tailed t-test, all data points which are significantly different from their corresponding negative control are highlighted with a green asterisk.

5.5.4 DARPins selected against VcCNT SMALPs

Samples of detergent solubilised VcCNT were produced via the protocol described in **section 2.5.3**, as shown in **Figure 5.14** samples taken throughout the purification procedure were analysed via SDS-PAGE and western blotting. As was the case with previously described purifications, the samples taken throughout the procedure give an insight into the state of the protein. From **Figure 5.14** it is evident that VcCNT expresses well (see intense band in the 'Total' sample of the Coomassie stained gel and its counterpart on the western blot). The vast majority of this protein is successfully solubilised and only a minimal amount of VcCNT did not successfully bind to the resin. The final yield of VcCNT was determined to be 4.35 mg from 100mg of membrane produced in a 30 L fermentation after extensive dialysis and concentration in a viva spin concentrator with a 60 kDa molecular weight cut off.

Like the other ELISAs, the absorbance readings produced at 450 nm have been plotted and compared with a negative control consisting of the DARPin analysed in the absence of immobilised target protein (**Figure 5.15**). All five DARPins produce statistically significant data above 100 nM except for C1 which is only significantly different from its negative control at 10 μM. When the error of H11 is considered it appears to be the most consistent DARPin out of those chosen for further study (along with E3, C2 and G2 in **Figure 5.12** and **Figure 5.13** respectively), as the level of error is minimal across the entire data set. H11 also produces the highest absorbance reading with 10 μM of DARPin out of all 20 DARPins studied, with a value of 0.345. Unfortunately like the previously discussed ELISAs, the range of concentrations used in **Figure 5.15** is not wide enough to produce a plateau in the trace therefore an affinity cannot be determined this way and for all five DARPins only a lower value of 1 μM can be given.

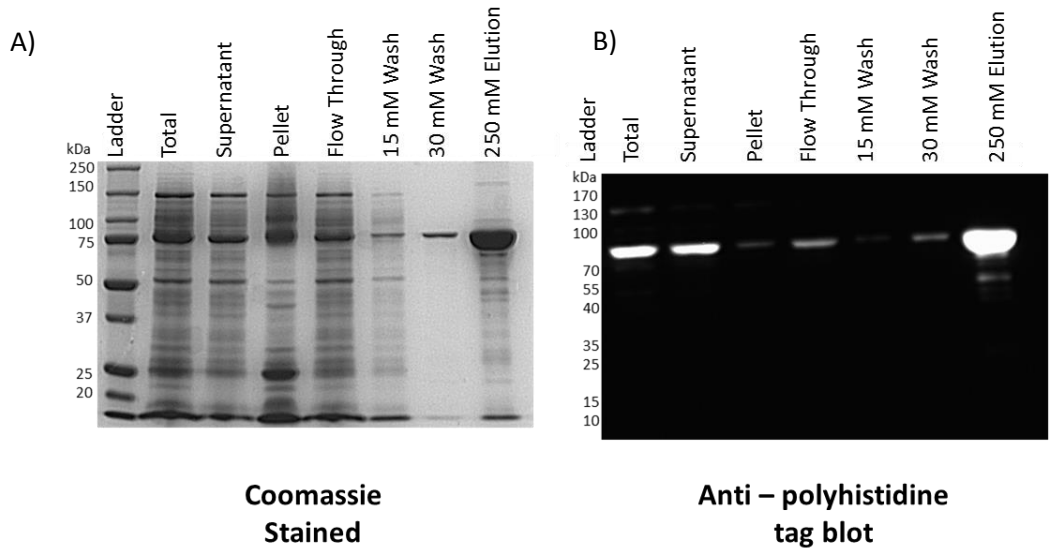
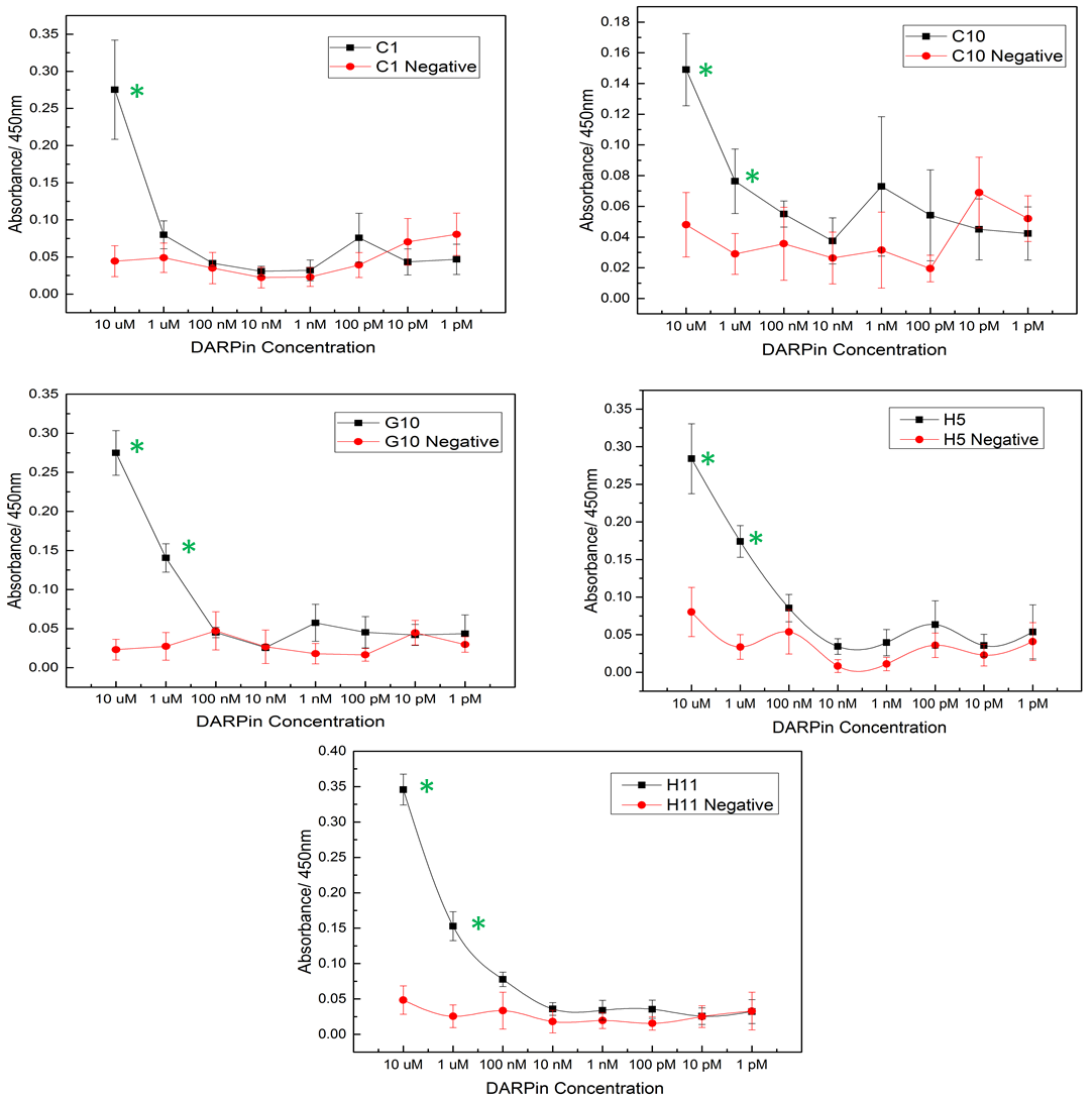


Figure 5.14: <u>IMAC Purification of detergent solubilised VcCNT:</u> Samples taken throughout the purification of VcCNT were analysed via 12% SDS-PAGE stained with Coomassie blue and western blotting probed with an anti-polyhistidine tag antibody conjugated with HRP. Each sample was normalised to a volume of 10 μL, therefore the 'Total' 'Supernatant' and 'Flow Through' represents 0.02% of the original sample. The 'Pellet' represents 0.1% and the two washes represent 0.05%. The eluted protein was subjected to extensive dialysis and concentration in a viva spin concentrator with a 60 kDa molecular weight cut off, it represents 0.33% of the original sample.



N = 3

Figure 5.15: <u>DARPin binding ELISA performed against DARPins selected against VcCNT SMALPs:</u> The ELISAs were performed in the presence of 10 µg of detergent solubilised VcCNT which was aspecifically adsorbed onto the surface of a Nunc Maxisorp® plate and probed with a serial dilution of each DARPin ranging from 10 µM to 1 pM. The negative control consisted of each DARPin analysed in the absence of target protein and all data points were subtracted from the background noise, determined by the analysis of BSA in the absence of DARPin. Each ELISA was repeated three times and each concentration was tested in duplicate in order to calculate the error by standard deviation. Statistical significance was determined by a two- tailed T-test and all data which is statistically significant from its negative control counterpart is highlighted with a green asterisk.

5.7 Discussion and Conclusions

While it was not possible to determine the specific affinity of the DARPins for the various target proteins, the ELISAs confirm that the DARPin are able to bind as they produce significantly higher levels of absorbance at 450 nm in the presence of adsorbed target protein. They do not however confirm that the DARPins are specific for their targets. In order to confirm the specificity of the DARPins, future ELISAs will need to be performed on plates with other membrane proteins adsorbed on their surface. The assays shown in **Figure 5.11, Figure 5.13** and **Figure 5.15** were performed using DARPins which were selected against SMALPs, the ELISAs however were performed on detergent solubilised protein. The results therefore prove that the DARPins are not binding directly to the SMA polymer. This is also true of the assay in **Figure 5.12** which was performed on DARPins selected against AcrB nanodiscs, the fact that significant data was produced in the presence of detergent solubilised AcrB proves that the DARPins are not binding to MSP directly (which is a risk when nanodiscs are used in selection experiments). Unfortunately, the DARPin assay which was carried out on the two DARPins selected against detergent solubilised MPSIL0294-V532C (chemically labelled via biotin maleimide) does not exclude the possibility that the DARPins may bind to the SBP or avitag. The fact that they were performed on unlabelled MPSIL0294-V532C in DDM micelles however does exclude binding to PEG linker.

The seemingly poor affinity of the DARPins isolated in this study are in stark contrast with previous studies which have isolated DARPins in a similar manner as detailed here and have resulted in DARPins with sub-nanomolar affinities (Zahnd, Wyler et al. 2007, Steiner, Forrer et al. 2008). However many of these studies used a mixture of solid and solution based selections and a higher number of phage display cycles thereby reducing the number of background binders.

The poor affinity of the DARPins may also be a result of certain limitations with the assay, for instance, every membrane protein under test had a polyhistidine tag, thereby making it difficult to distinguish if the DARPins are binding directly to the tag or the protein, although this is unlikely as the large surface area of a DARPin most likely would not bind a linear epitope such as a polyhistidine tag. Similarly, the aspecific nature of the DARPin biotinylation procedure may have untold consequences on the affinity of the DARPin to the target. In the case of the DARPins selected against SMALPs and nanodiscs, the low affinity may also be a consequence of the DDM which was present throughout all of the ELISAs. This DDM was added in an attempt to retain the target proteins native confirmation after adsorption, the presence of micelles however may occlude the surface epitopes which the selected DARPins target,

thereby interfering with DARPin binding (see **Chapter 1**). Ideally this assay should be replaced with an alternative which probes the protein in their native membrane environments. In future selection projects this assay works as a first step in the confirmation of the DARPins ability to bind a His-tagged target protein however It would need to be followed by a repeat of the assay with adsorbed protein lacking a polyhistidine tag so that the N-terminal His tag on the DARPins could be detected via an anti-polyhistidine antibody conjugated with HRP, thereby excluding direct binding of the His-tag. It would also be beneficial if this assay was followed up with an assay which does not require the DARPin to be biotinylated such as size exclusion chromatography (SEC), flow cytometry or SEC-MALLs. In either the case biophysical techniques such as surface plasmon resonance (SPR) would be the subsequent validation test in order to elucidate the DARPins affinities.

Despite the limitations of the assay, several DARPins have been identified which show promise in their ability to bind to their targets. These DARPins are prime targets for crystallisation chaperones in order to elucidate the structure of VcCNT (DARPin H11), NupC (G2 and C2) and MPSIL0294 (E6). Although it must be considered that while biotinylation of the DARPin was confirmed, the extent of biotinylation was not, the extent of biotinylation could affect the DARPins affinity and performance in the ELISAs. Despite this, G2, C2 and H11 (from **Figure 5.13** and **Figure 5.15** respectively) are considered to be best DARPin binders based upon the consistency that the results displayed in the DARPin binding ELISAs. E6 is also considered one of the most promising DARPins based upon the number of data points which show a significant increase in absorbance compared to its negative control, while retaining the expected absorbance trend (**see Figure 5.10**). In follow up studies the extent of biotinylation should be measured using the HABA method (see **section 2.4.4**).

With the data from the DARPin ELISAs the sequences of the most promising DARPins were aligned with one another (see **Figure 5.16A**). Firstly the sequences of DARPins G2 and C2, selected against NupC SMALPs, show very little similarity overall other than a small cluster of residues at positions 34, 36 and 44 which have an asparagine, an alanine and a histidine residue. Despite the fact that out of the DARPins selected against NupC, only G2 and C2 have all three of these residues in these positions, the importance of them is not clear. The residues do not share any physical characteristics (only two of the residues are hydrophilic, polar and neutral) and despite their presence the DARPin assays in **Figure 5.13** show that G2 produced absorbance readings almost twice as high as C2. Other than this, the only other position with

some similarity between the two is the 69th residue (a valine in G2 and an isoleucine in C2) both of which are aliphatic residues.

DARPins A1 and E3 selected against AcrB SMALPs and nanodiscs, respectively, can also be considered some of the best DARPins which have been identified based upon their consistent absorbance readings. However these DARPins will not be used in co-crystallisation trails as the crystal structure of AcrB has already been elucidated (Murakami, Nakashima et al. 2002), as well as insights into the nature of AcrB in association with DARPins, which showed that the DARPins bound to the periplasmic domain of AcrB (Brandstatter, Sokolova et al. 2011). It was also observed that several the stoichiometry of the DARPin with AcrB was dependent upon the multimeric variant of AcrB that they tested with, although it was shown that the DARPin was able to bind to AcrB in a stronger manner than the interactions between AcrB monomers (Brandstatter, Sokolova et al. 2011). AcrB was chosen to be a as a model target protein to compare SMALP technology, nanodiscs and detergent micelles as a method to present the target during phage display screening. As discussed previously, DDM micelles of AcrB failed to select any DARPins which passed our selection criteria, while it was possible to select four DARPins against nanodiscs AcrB and thirty DARPins again SMALPs AcrB. This suggests that the SMALPs are a better format for isolating potential binding partners. However, the first round validation tests of DARPins selected against SMALP AcrB was performed by binding assays of full phage to SMALP AcrB, and false positives could have been selected, e.g. DARPins binding to the SMA polymer. Thus, I performed the ELISA assay discussed in this chapter, and this confirmed the binding of DARPins A1 and E3 to detergent-solubilised AcrB (**Figure 5.11 and 5.12,** respectively). An alignment of both of their sequences (see **Figure 5.16B**) shows that the tyrosine at position 90 is the only place in which both DARPins have the same residue. However it has already been shown that in this position a tyrosine is one of the preferred residues in all DARPins, therefore there is most likely not any real significance in this. Other than this, the two DARPins have similar residues in three other positions; the isoleucine and leucine in position 33, the valine and isoleucine in position 102 and the arginine and lysine at position 110.

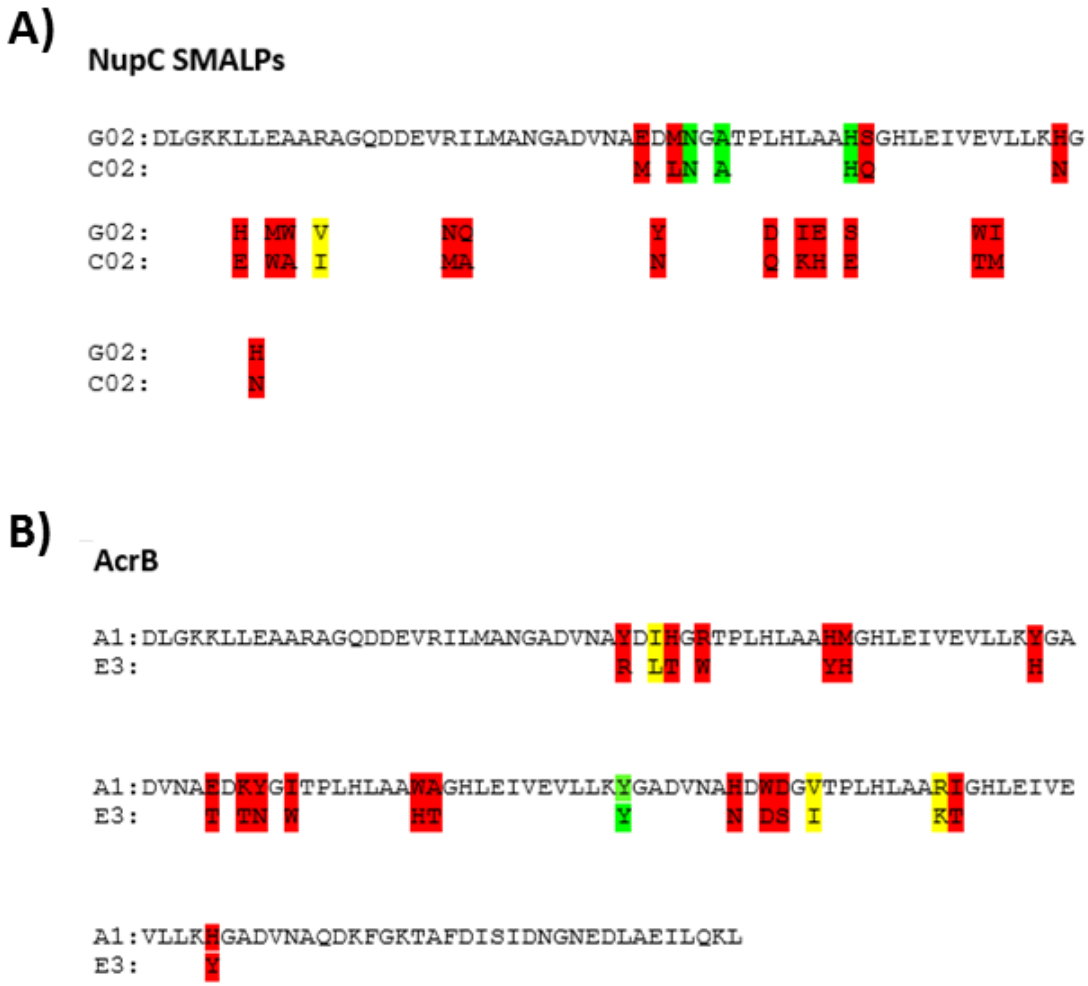


Figure 5.16: <u>Amino acid sequence alignment of the most promising DARPins</u>: The sequences of the most promising DARPins selected against the same target have been aligned in order to highlight any sequence similarity between them. The DARPins aligned are; A) G2 and C2 selected against NupC SMALPs and, B) DARPins A1 and E3 selected against AcrB SMALPs and nanodiscs respectively. Residues highlighted in red are not related to one another, those in yellow share some physical characteristics and those in green are identical.

Thus, it is evident that both nanodiscs and SMALPs AcrB are suitable formats to present a target protein in a phage display screening. However because only one unique DARPin successfully bound AcrB SMALPs and nanodiscs it is not possible to draw meaningful conclusions regarding which format is better.

SMALPs are better in regard to their production however, which is less time intensive than nanodiscs. Also SMA solubilised membrane proteins can generally be stored longer than nanodiscs. Observational evidence has shown that nanodiscs only remain stable for a maximum of 2 weeks when stored at 4°C. It has also been observed in this project that nanodisc construction requires up to two weeks in order to remove all traces of detergent and are unable to withstand storage at -80°C. Considering the vast amount of time required in their construction and the complications in their storage, the fact that the DARPins which are

isolated are of a similar quality as those selected against SMALPs provides a strong argument to suggest that SMALP are a preferable presentation platform for membrane proteins.

Chapter 6

Discussion and future directions

Selecting DARPins against immobilised membrane proteins is a well-established concept (Boersma, Chao et al. 2011, Stefan, Martin-Killias et al. 2011, Mittal, Bohm et al. 2012). Screening experiments often lead to the generation of high affinity binders some of which are capable of effecting protein function and thus show promise as potential drug candidates (Eggel, Baumann et al. 2009, Stefan, Martin-Killias et al. 2011). The screens however, often use detergent solubilised membrane protein which has been immobilised via aspecific adsorption or biotinylation. Depending upon the detergent selected, some potential surface epitopes may be engulfed within the micelle, thereby rendering them inaccessible to any potential binding partner (see **chapter 1**). Therefore, the primary aim of this thesis was to compare alternative methods of membrane protein immobilisation and presentation with regard to their ability to select DARPins capable of binding a panel of model membrane proteins after two rounds of phage display.

To compare alternative methods of protein immobilisation, alternative protein tags were cloned onto the C-terminal of a model membrane protein. The initial model protein was a divalent metal ion transporter from *Enterococcus faecalis* designated MPSIL0294. Three different C-terminal tags were inserted downstream from a polyhistidine tag: a streptavidin binding peptide (SBP), an avitag which had undergone *in vivo* biotinylation and the chemical biotinylation of a C-terminal cysteine residue substituted for the valine residue at position 532 via site directed mutagenesis (see **chapter 3**). A second model membrane protein, AcrB – a component of a multi drug efflux pump from *E.coli*, was chemically biotinylated with NHS-biotin also.

Comparative ELISAs showed that the three MPSIL0294 variants differed in their ability to immobilise on neutravidin coated plates, when probed with an anti polyhistidine antibody conjugated with HRP (see **Figure 3.13**). Out of the three variants, MPSIL0294-V532C showed the highest reading on the plates. Therefore, for the following reasons I hypothesis this higher reading is due to a lower surface coverage, which lead to an increase in the accessibility of the C-terminal polyhistidine tag. Firstly, the MPSIL0294 variants were also tested with FACs analysis after immobilisation on streptavidin coated magnetic beads (see **Figure 3.14**). In contrast to the multiwell plates, the FACs showed that MPSIL0294-V532C was unable to immobilise to the same extent as the other two variants, possibly due to its lower labelling efficiency. I therefore hypothesize that less packing of an immobilised membrane protein on a solid surface may lead to the increased accessibility of epitopes to an antibody.

Each of the MPSIL0294-variants were subjected to two rounds of phage display followed by validation of the resultant selection outputs via phage ELISA (see **Section 3.7**). Out of the three MPSIL0294 variants and detergent solubilised AcrB, MPSIL0294-V532C produced the highest number of 'false-positive' DARPins that bound strongly streptavidin (see **Table 3.6**). This may be due to MPSIL0294-V532Cs lower coverage on the pates leading to an increased exposure of streptavidin to the DARPin library.

Thus I propose that a decrease in surface packing may lead to the increased generation of DARPins which bind streptavidin. Interestingly, MPSIL0294-V532C was also the only variant capable of selecting DARPins that bound the target, supporting the hypothesis that saturation of the surface may be a hindrance in the pursuit of binding partners during a selection, although more data is required to confirm this. Another comparative phage display which compared different levels of surface coverage would address this issue. This could either be achieved in an identical manner as described above, or by immobilising lower concentrations of protein onto the surface prior to a selection. This is especially interesting due to the fact that such a comparison in phage display has not been attempted before and the importance of factors such as coverage of a surface by a target is often overlooked in the literature.

All of the MPSIL0294 variants were solubilised in DDM; a twelve carbon maltoside which produces large micelles. If the size of large micelles reduce the accessibility of surface epitopes to the DARPins (see **Chapter 1**), this will not be detected in this study. Therefore a comparative phage display using protein solubilised in different detergents would determine if the size of the detergent micelle is a factor in the number of binding partners successfully identified. It is expected that, as the carbon chain of the detergent decreases, more of the target protein will become accessible and therefore produce more binding DARPins.

As stated, this study also aimed to compare alternative methods of membrane protein presentation in regard to their ability to isolate DARPins after phage display. Therefore, a panel of membrane proteins were solubilised in SMALPs and nanodiscs and in a comparative study. The panel of proteins included MPSIL0294-SBP SMALPs and nanodiscs, AcrB SMALPs and nanodiscs, the *E.coli* concentrative nucleoside transporter NupC in SMALPs and NupCs homologue from *Vibrio cholera* VcCNT in SMALPs (see **Chapter 4**). All members of this panel were subjected to chemical biotinylation with NHS-biotin, with the exception of the MPSIL0294-SBP and AcrB nanodisc samples; as the former is able to bind streptavidin directly, while the latter were constructed with a biotinylated MSP. Once again comparative ELISAs were performed on samples of MPSIL0294-SBP nanodiscs and SMALPs to test their ability to

immobilise on neutravidin coated plates. The nanodisc samples followed the trend established by its detergent solubilised counterparts although with a substantially lower magnitude, while the SMALPs failed to produce a significant signal (see **Figure 4.7**). This may theoretically be due to a decrease in the accessibility of the tag to the polyhistidine antibody due to the large size of nanodiscs (9.8 nm) (Bayburt and Sligar 2010) and SMALPS (9-12 nm) and the large molecular weight of the antibody-HRP conjugate.

It might be that a lower coverage or loading, leads to decreased levels of steric hindrance between the individual MPSIL0294 proteins. This hypothesis could be tested via an ELISA that compares increasing surface coverages of MPSIL0294-V532C. This coverage could be monitored using quartz crystal microbalance with dissipation monitoring (QCM-D) with a streptavidin coated QCM chip. ELISAs would then be subsequently performed on the chips themselves, the result of which should display an inverse correlation with the surface coverage if the conclusion is correct. As stated above, the DDM micelle, nanodiscs and SAMLPS are all large constructs. Therefore, the effect sample size has on the success of the ELISA could be tested through use of a smaller probe. For instance, if the samples were immobilised via their His-tags on Ni-NTA covered plates they could be subsequently probed streptavidin-HRP and detected with TMB. Despite the fact that the streptavidin HRP is approximately 100 kDa it is smaller than a full length IgG antibody (150 kDa). This probe is still rather large however; therefore samples of MPSIL0294-avitag or MPSIL0294-V532C nanodiscs could be probed with biotin-HRP which is substantially smaller. Although ELISAs which employed biotin-HRP would need to be performed on blank maxisorp plates in order to ensure the biotin-HRP does not bind directly to the neutravidin. Although a direct comparison between the two experiments would not be accurate as the different probes would have different sensitivities. An alternative follow up to this work would be a comparative ELISA with a smaller target protein. The predicted mass of MPSIL0294 is approximately 60 kDa which is larger than the mass of streptavidin itself. It therefore seems likely that when the protein saturates the streptavidin coated surface it would sterically hinder a large antibody probe from gaining access to its epitope. The ELISA data presented in this study is unable to determine whether the readings are due to the polyhistidine tag on the MPSIL0294-SBP or the his tag of the MSP, therefore repeating this ELISA on empty nanodiscs will show if the MSP polyhistidine tag is accessible. Alternatively, if biotin-HRP is employed as stated earlier, it could also distinguish between the two.

The SMALPs failed to produce a signal using an ELISA with anti-his-HRP, which might be due to the electrostatic interaction between the strong negative charge of the maleic acid moiety and the positive charge of the polyhistidine tag (see **Figure 4.7**). This electrostatic interaction has also been shown to reduce the affinity of the tag to the cobalt resin. The reduced affinity was evident from the elution of SMALPs at low imidazole concentrations which, in compound with the poor expression level of MPSIL0294 and the excess of resin used in its purification, is the most likely reason for its lower purities compared to the detergent-solubilised proteins. The accessibility of the SBP tag on MPSIL0294 was therefore confirmed by binding the protein to streptactin, which proved to be successful, although only a limited fraction of the protein was shown to bind. The protein which failed to bind to the streptactin resin may be due to shielding of the SBP tag by the SMA polymer. It would therefore be an interesting follow up to repeat this streptactin binding assay with a large linker between the C-terminal of the protein and the tag. Distancing both the polyhistidine tag as well as the SBP tag may also alleviate the issues observed during the purification MPSIL0294 SMALPs.

The two rounds of phage display were performed on the panel of membrane proteins followed by two phage ELISA validation tests, the first of which highlighted the number of DARPins capable of binding streptavidin (see **Section 4.4**). The result of this validation test showed that AcrB nanodiscs produced the highest number of streptavidin binders (76% of the selection output) followed by the AcrB SMALPs (64%). The prevalence of streptavidin binders to somewhat unexpected considering the fact that the phage library was initially deselected against blank streptavidin coated surfaces. Nevertheless, for AcrB, the detergent solubilised sample (59%) gave the best results in this respect, with the least numbers of selected DARPins binding to streptavidin. This trend is not seen in samples of MPSIL0294-SBP as the proportion of the selection outputs isolated against DDM solubilised protein which bound streptavidin was 78%, compared with 19% of the nanodiscs and 12.5% of the SMALPs.

Compared to published instances of selection with DARPins, four rounds of selection in ribosome display was deemed sufficient to produce a library of DARPin inhibitors which bound specifically to DDM solubilised AcrB which had been biotinylated. Prior to the selection against the AcrB, the DARPins were subjected to two deselection procedures against neutravidin, BSA and biotinylated maltose binding protein (Sennhauser, Amstutz, Briand, Storchenegger, & Grütter, 2006). Similarly a study performed three rounds of phage display against the Fc domain of human IgG1 and reportedly began to isolate high quality binders after the second round with very little aspecific binding to a streptavidin or neutravidin. In this study a mixture

of solid selection (in which target proteins are immobilised on a solid surface such as a plate) and solution selection (in which target proteins are immobilised onto the surface of beads capable of staying in solution such as streptavidin coated magnetic Dynabeads®) (Steiner, Forrer, & Plückthun, 2008).

Thus when the number of streptavidin or aspecific binders that were isolated in this study is considered, it appears clear that the deselection procedure employed was insufficient. Also considering the low affinity of the binders isolated it appears clear that a phage display protocol that combined solid and solution selections, as well as increasing the number of panning rounds, may have improved the selection outputs. Follow up studies must use a more extensive deselection process, possibly introducing multiple rounds of deselection, with both streptavidin and neutravidin. If subsequent selection aim to generate particularly strong binders then a competitive selection strategy could be employed. In competitive selections, a known associate of the immobilised target protein can be added to a later panning round. In theory the strongest binders will be retained and bind to the target, while weaker binders are outcompeted by the associate. This protocol has been demonstrated to vastly increase the affinity of scFVs against HM-1 killer toxin (HM-1)-neutralizing monoclonal antibodies (nmAb-KT) by adding the toxin HM-1 after four rounds of phage display (Kabir, Krishnaswamy, Miyamoto, Furuichi, & Komiyama, 2009).

The reason for this difference is not clear, therefore repeating the two rounds of phage display with AcrB tagged with SBP would be an essential experiment to perform in order to determine the relationship between solubilisation method and the selection of streptavidin binders. These experiments should also be seen in relation to the coverage or 'loading' efficiency of the target, as discussed above. This relationship would be further clarified if follow up studies increased the number of formats under test to include biotinylated amphipols. In order to increase the applicability of this work it would also be beneficial to include different antibody mimetics such as scFVs, nanobodies and adhirrons (see **section 1.4**) in order to see if solubilisation method has a similar effect on their selection.

The second of the validation tests selected the DARPins that were capable of binding the target and showed that 34% of the selection output against AcrB SMALPs was capable of binding the protein. The relatively high number of binding DARPins was also shown in samples of NupC SMALPs (26%) and particularly VcCNT (74%) (see **Table 4.13**). The study shows that the proportion of DARPin binders selected against SMALPs is substantially higher than both nanodiscs and detergent solubilisation, as only 4.5% of the AcrB nanodisc selection output

bound to the target while no DARPins bound to DDM solubilised AcrB. Of all the DDM solubilised protein tested, the only selection output to produce positive DARPins was that of MPSIL0294-V532C (2.3%). This provides a clear suggestion that detergent free methods of solubilisation are superior and that nanodiscs are inferior to SMALPs. Unfortunately due to the low yield of MPSIL0294-SBP SMALPS and nanodiscs, they could not be included in the second validation test and it was thus not possible to distinguish between genuine DARPin binders and false positives. However, the SMALPs success provides further credence to the notion that I proposed earlier, that a decrease in surface coverage may lead to an increase accessibility to potential epitopes on the surface of a plate. When taken in conjunction with my second hypothesis, that the MPSIL0294-SBP SMALPS may struggle to saturate a streptavidin covered surface due to shielding of the SBP tag by the SMA polymer. A limitation of this hypothesis is the fact that the SMALPs discussed above were immobilised via chemical biotinylation not SBP, therefore, phage display performed on SMALPs constructed with the samples of AcrB tagged with SBP suggested earlier would better test this hypothesis.

The most natural progression from this thesis is a comparative study with a wider range of membrane proteins particularly GPCRs, which have successfully been solubilised in SMALPs and are the leading target of novel pharmaceutical agents (Salon, Lodowski et al. 2011, Jamshad, Charlton et al. 2015). Of particular importance is the inclusion of eukaryotic proteins such as human NRAMP.

20 DARPins were selected for further characterisation and confirmation of their ability to bind to their targets. While the aim of this study was not to produce high affinity DARPins and due to the low number of cycles (two) in the phage display, it is not unexpected that all the DARPins have micro-molar affinity (see **Section 5.5**). In order to isolate confirmation specific DARPins, DDM was present throughout the ELISAs. It is possible that the apparent affinities in the ELISA studies are influenced by biotinylation of the DARPins, which may have reduced their affinity. The number of cycles in future selections will need to increase if high affinity DARPins are required. The ELISAs help to exclude DARPin binding to the MSP of nanodiscs and the SMA copolymer. However, they do not exclude binding to the polyhistidine tag, therefore the next step in this study would be to perform biophysical analysis of these DARPins through SEC and SPR.

The most promising DARPins isolated in this study are H11 selected against VcCNT SMALPs, G2 and C2 (NupC SMALPs), E6 (MPSIL0294-V532C). The use of these DARPins as crystallisation chaperones would also be a good follow up study to this thesis. It may also be beneficial to

perform mutagenic analysis in order to determine their essential residues and subsequently attempt to improve their affinities. As well as including DARPin E6 in the MPSIL0294 Zn uptake assay, to investigate if it's binding has functional consequences on MPSIL0294.

Of particular interest are DARPins E3, selected against AcrB nanodiscs, and A1, selected against AcrB SMALPS (see **Section 5.5.2**). Both of these DARPins produced absorbance traces which are almost identical to one another with a very similar maximum absorbance reading at 10 µM of DARPin. These two DARPins do not share a high level of sequence similarity in their variable regions. This suggests that DARPins of a similar quality can be selected to both nanodisc and SMALP targets. However, nanodiscs take lot longer to produce and lack stability. As SMALP targets resulted in a higher number of 'positive' DARPins, I conclude that SMALPs are the best presentation method for the selection of DARPins tested in this thesis.

References

Abrantes, M. C., J. Kok and F. Lopes Mde (2013). "EfaR is a major regulator of Enterococcus faecalis manganese transporters and influences processes involved in host colonization and infection." Infect Immun **81**(3): 935-944.

Ahmad, Z. A., S. K. Yeap, A. M. Ali, W. Y. Ho, N. B. M. Alitheen and M. Hamid (2012). "scFv Antibody: Principles and Clinical Application." Clinical and Developmental Immunology **2012**: 15.

Andreu, C. and M. Del Olmo (2013). "Yeast arming by the Aga2p system: effect of growth conditions in galactose on the efficiency of the display and influence of expressing leucine-containing peptides." Appl Microbiol Biotechnol **97**(20): 9055-9069.

Aschner, M. (2006). "The transport of manganese across the blood–brain barrier." NeuroToxicology **27**(3): 311-314.

Baek, H., K. H. Suk, Y. H. Kim and S. Cha (2002). "An improved helper phage system for efficient isolation of specific antibody molecules in phage display." Nucleic Acids Res **30**(5): e18.

Bao, H., F. Duong and C. S. Chan (2012). "A step-by-step method for the reconstitution of an ABC transporter into nanodisc lipid particles." J Vis Exp(66): e3910.

Barrette-Ng, I. H., S. C. Wu, W. M. Tjia, S. L. Wong and K. K. Ng (2013). "The structure of the SBP-Tag-streptavidin complex reveals a novel helical scaffold bridging binding pockets on separate subunits." Acta Crystallogr D Biol Crystallogr **69**(Pt 5): 879-887.

Basit, H., K. Shivaji Sharma, A. Van der Heyden, C. Gondran, C. Breyton, P. Dumy, F. M. Winnik and P. Labbe (2012). "Amphipol mediated surface immobilization of FhuA: a platform for label-free detection of the bacteriophage protein pb5." Chem Commun (Camb).

Bayburt, T. H. and S. G. Sligar (2010). "Membrane protein assembly into Nanodiscs." FEBS Lett **584**(9): 1721-1727.

Bazan, J., I. Calkosinski and A. Gamian (2012). "Phage display--a powerful technique for immunotherapy: 1. Introduction and potential of therapeutic applications." Hum Vaccin Immunother **8**(12): 1817-1828.

Bessette, P. H., J. J. Rice and P. S. Daugherty (2004). "Rapid isolation of high-affinity protein binding peptides using bacterial display." Protein Eng Des Sel **17**(10): 731-739.

Boersma, Y. L., G. Chao, D. Steiner, K. D. Wittrup and A. Pluckthun (2011). "Bispecific designed ankyrin repeat proteins (DARPins) targeting epidermal growth factor receptor inhibit A431 cell proliferation and receptor recycling." J Biol Chem **286**(48): 41273-41285.

Boersma, Y. L. and A. Pluckthun (2011). "DARPins and other repeat protein scaffolds: advances in engineering and applications." Curr Opin Biotechnol **22**(6): 849-857.

Bolanos-Garcia, V. M. and O. R. Davies (2006). "Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography." Biochim Biophys Acta **1760**(9): 1304-1313.

Bork, P., L. Holm and C. Sander (1994). "The Immunoglobulin Fold: Structural Classification, Sequence Patterns and Common Core." J Mol Biol **242**(4): 309-320.

Bowie, J. U. (2001). "Stabilizing membrane proteins." Curr Opin Struct Biol **11**(4): 397-402.

Brandstatter, L., L. Sokolova, T. Eicher, M. A. Seeger, C. Briand, H. J. Cha, M. Cernescu, J. Bohnert, W. V. Kern, B. Brutschy and K. M. Pos (2011). "Analysis of AcrB and AcrB/DARPin ligand complexes by LILBID MS." Biochim Biophys Acta **1808**(9): 2189-2196.

Bremer, E., A. Middendorf, J. Martinussen and P. Valentin-Hansen (1990). "Analysis of the tsx gene, which encodes a nucleoside-specific channel-forming protein (Tsx) in the outer membrane of Escherichia coli." Gene **96**(1): 59-65.

Carmen, S. and L. Jermutus (2002). "Concepts in antibody phage display." Brief Funct Genomic Proteomic **1**(2): 189-203.

Cha, H. J., R. T. Muller and K. M. Pos (2014). "Switch-loop flexibility affects transport of large drugs by the promiscuous AcrB multidrug efflux transporter." Antimicrob Agents Chemother **58**(8): 4767-4772.

Chan, S.-N., N. Abu Bakar, M. Mahmood, C.-L. Ho and N. A. Shaharuddin (2014). "Molecular Cloning and Characterization of Novel Phytocystatin Gene from Turmeric, Curcuma longa." BioMed Research International **2014**: 9.

Chao, Y. and D. Fu (2004). "Kinetic study of the antiport mechanism of an Escherichia coli zinc transporter, ZitB." J Biol Chem **279**(13): 12043-12050.

Cheng, X. and H. Wang (2012). "Multiple targeting motifs direct NRAMP1 into lysosomes." Biochemical and Biophysical Research Communications **419**(3): 578-583.

Dahmane, T., F. Giusti, L. J. Catoire and J. L. Popot (2011). "Sulfonated amphipols: synthesis, properties, and applications." Biopolymers **95**(12): 811-823.

Daugherty, P. S. (2007). "Protein engineering with bacterial display." Curr Opin Struct Biol **17**(4): 474-480.

Denisov, I. G., Y. V. Grinkova, A. A. Lazarides and S. G. Sligar (2004). "Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size." J Am Chem Soc **126**(11): 3477-3487.

Diab, C., C. Tribet, Y. Gohon, J. L. Popot and F. M. Winnik (2007). "Complexation of integral membrane proteins by phosphorylcholine-based amphipols." Biochim Biophys Acta **1768**(11): 2737-2747.

Du, D., Z. Wang, N. R. James, J. E. Voss, E. Klimont, T. Ohene-Agyei, H. Venter, W. Chiu and B. F. Luisi (2014). "Structure of the AcrAB-TolC multidrug efflux pump." Nature **509**(7501): 512-515.

Eggel, A., M. J. Baumann, P. Amstutz, B. M. Stadler and M. Vogel (2009). "DARPins as bispecific receptor antagonists analyzed for immunoglobulin E receptor blockage." J Mol Biol **393**(3): 598-607.

Ehrnstorfer, I. A., E. R. Geertsma, E. Pardon, J. Steyaert and R. Dutzler (2014). "Crystal structure of a SLC11 (NRAMP) transporter reveals the basis for transition-metal ion transport." Nat Struct Mol Biol **21**(11): 990-996.

Eicher, T., H. J. Cha, M. A. Seeger, L. Brandstatter, J. El-Delik, J. A. Bohnert, W. V. Kern, F. Verrey, M. G. Grutter, K. Diederichs and K. M. Pos (2012). "Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop." Proc Natl Acad Sci U S A **109**(15): 5687-5692.

Elvin, J. G., R. G. Couston and C. F. van der Walle (2013). "Therapeutic antibodies: market considerations, disease targets and bioprocessing." Int J Pharm **440**(1): 83-98.

Fagerberg, L., K. Jonasson, G. von Heijne, M. Uhlen and L. Berglund (2010). "Prediction of the human membrane proteome." Proteomics **10**(6): 1141-1149.

Fortier, A., G. Min-Oo, J. Forbes, S. Lam-Yuk-Tseung and P. Gros (2005). "Single gene effects in mouse models of host: pathogen interactions." Journal of Leukocyte Biology **77**(6): 868-877.

Fortin, A., L. Abel, J. L. Casanova and P. Gros (2007). "Host Genetics of Mycobacterial Diseases in Mice and Men: Forward Genetic Studies of BCG-osis and Tuberculosis." Annual Review of Genomics and Human Genetics **8**(1): 163-192.

Franzin, C. M., X. M. Gong, K. Thai, J. Yu and F. M. Marassi (2007). "NMR of membrane proteins in micelles and bilayers: the FXYD family proteins." Methods **41**(4): 398-408.

Frehel, C., F. Canonne-Hergaux, P. Gros and C. De Chastellier (2002). "Effect of Nramp1 on bacterial replication and on maturation of Mycobacterium avium-containing phagosomes in bone marrow-derived mouse macrophages." Cell Microbiol **4**(8): 541-556.

Fridy, P. C., Y. Li, S. Keegan, M. K. Thompson, I. Nudelman, J. F. Scheid, M. Oeffinger, M. C. Nussenzweig, D. Fenyo, B. T. Chait and M. P. Rout (2014). "A robust pipeline for rapid production of versatile nanobody repertoires." Nat Methods **11**(12): 1253-1260.

Friedman, M., E. Nordberg, I. Hoiden-Guthenberg, H. Brismar, G. P. Adams, F. Y. Nilsson, J. Carlsson and S. Stahl (2007). "Phage display selection of Affibody molecules with specific binding to the extracellular domain of the epidermal growth factor receptor." Protein Eng Des Sel **20**(4): 189-199.

Friedrich, M. G., V. U. Kirste, J. Zhu, R. B. Gennis, W. Knoll and R. L. Naumann (2008). "Activity of membrane proteins immobilized on surfaces as a function of packing density." J Phys Chem B **112**(10): 3193-3201.

Gera, N., M. Hussain and B. M. Rao (2013). "Protein selection using yeast surface display." Methods **60**(1): 15-26.

Getz, J. A., T. D. Schoep and P. S. Daugherty (2012). "Peptide discovery using bacterial display and flow cytometry." Methods Enzymol **503**: 75-97.

Giusti, F., J. L. Popot and C. Tribet (2012). "Well-defined critical association concentration and rapid adsorption at the air/water interface of a short amphiphilic polymer, amphipol A8-35: a study by Forster resonance energy transfer and dynamic surface tension measurements." Langmuir **28**(28): 10372-10380.

Glockshuber, R., M. Malia, I. Pfitzinger and A. Pluckthun (1990). "A comparison of strategies to stabilize immunoglobulin Fv-fragments." Biochemistry **29**(6): 1362-1367.

Glover, C. A., V. L. Postis, K. Charalambous, S. B. Tzokov, W. I. Booth, S. E. Deacon, B. A. Wallace, S. A. Baldwin and P. A. Bullough (2011). "AcrB contamination in 2-D crystallization of membrane proteins: lessons from a sodium channel and a putative monovalent cation/proton antiporter." J Struct Biol **176**(3): 419-424.

Gohon, Y., F. Giusti, C. Prata, D. Charvolin, P. Timmins, C. Ebel, C. Tribet and J. L. Popot (2006). "Well-defined nanoparticles formed by hydrophobic assembly of a short and polydisperse random terpolymer, amphipol A8-35." Langmuir **22**(3): 1281-1290.

Goldman, E. R., G. P. Anderson, J. L. Liu, J. B. Delehanty, L. J. Sherwood, L. E. Osborn, L. B. Cummins and A. Hayhurst (2006). "Facile generation of heat-stable antiviral and antitoxin single domain antibodies from a semisynthetic llama library." Anal Chem **78**(24): 8245-8255.

Goluch, E. D., A. W. Shaw, S. G. Sligar and C. Liu (2008). "Microfluidic patterning of nanodisc lipid bilayers and multiplexed analysis of protein interaction." Lab Chip **8**(10): 1723-1728.

Goswami, T., A. Bhattacharjee, P. Babal, S. Searle, E. Moore, M. Li and J. M. Blackwell (2001). "Natural-resistance-associated macrophage protein 1 is an H+/bivalent cation antiporter." Biochem J **354**(Pt 3): 511-519.

Grinkova, Y. V., I. G. Denisov and S. G. Sligar (2010). "Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers." Protein Eng Des Sel **23**(11): 843-848.

Gronwall, C. and S. Stahl (2009). "Engineered affinity proteins--generation and applications." J Biotechnol **140**(3-4): 254-269.

He, M. and F. Khan (2005). "Ribosome display: next-generation display technologies for production of antibodies in vitro." Expert Rev Proteomics **2**(3): 421-430.

Hmila, I., D. Saerens, R. Ben Abderrazek, C. Vincke, N. Abidi, Z. Benlasfar, J. Govaert, M. El Ayeb, B. Bouhaouala-Zahar and S. Muyldermans (2010). "A bispecific nanobody to provide full protection against lethal scorpion envenoming." FASEB J **24**(9): 3479-3489.

Holliger, P., T. Prospero and G. Winter (1993). ""Diabodies": small bivalent and bispecific antibody fragments." Proc Natl Acad Sci U S A **90**(14): 6444-6448.

Huber, T., D. Steiner, D. Rothlisberger and A. Pluckthun (2007). "In vitro selection and characterization of DARPins and Fab fragments for the co-crystallization of membrane proteins: The Na(+)-citrate symporter CitS as an example." J Struct Biol **159**(2): 206-221.

Jabado, N., P. Cuellar-Mata, S. Grinstein and P. Gros (2003). "Iron chelators modulate the fusogenic properties of Salmonella-containing phagosomes." Proc Natl Acad Sci U S A **100**(10): 6127-6132.

Jamshad, M., J. Charlton, Y. P. Lin, S. J. Routledge, Z. Bawa, T. J. Knowles, M. Overduin, N. Dekker, T. R. Dafforn, R. M. Bill, D. R. Poyner and M. Wheatley (2015). "G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent." Biosci Rep **35**(2).

Jamshad, M., Y. P. Lin, T. J. Knowles, R. A. Parslow, C. Harris, M. Wheatley, D. R. Poyner, R. M. Bill, O. R. Thomas, M. Overduin and T. R. Dafforn (2011). "Surfactant-free purification of

membrane proteins with intact native membrane environment." Biochem Soc Trans **39**(3): 813-818.

Johnson, Z. L., C. G. Cheong and S. Y. Lee (2012). "Crystal structure of a concentrative nucleoside transporter from Vibrio cholerae at 2.4 A." Nature **483**(7390): 489-493.

Jung, S. H., I. Pastan and B. Lee (1994). "Design of interchain disulfide bonds in the framework region of the Fv fragment of the monoclonal antibody B3." Proteins **19**(1): 35-47.

Justesen Bo, H. and T. Günther-Pomorski (2014). Chromatographic and electrophoretic methods for nanodisc purification and analysis. Reviews in Analytical Chemistry. **33:** 165.

Kang, H. J., C. Lee and D. Drew (2013). "Breaking the barriers in membrane protein crystallography." Int J Biochem Cell Biol **45**(3): 636-644.

Kaspar, M., E. Trachsel and D. Neri (2007). "The antibody-mediated targeted delivery of interleukin-15 and GM-CSF to the tumor neovasculature inhibits tumor growth and metastasis." Cancer Res **67**(10): 4940-4948.

Keefe, A. D. (2001). "Protein selection using mRNA display." Curr Protoc Mol Biol **Chapter 24**: Unit 24 25.

Khafizov, K., R. Staritzbichler, M. Stamm and L. R. Forrest (2010). "A Study of the Evolution of Inverted-Topology Repeats from LeuT-Fold Transporters Using AlignMe." Biochemistry **49**(50): 10702-10713.

Kieke, M. C., B. K. Cho, E. T. Boder, D. M. Kranz and K. D. Wittrup (1997). "Isolation of anti-T cell receptor scFv mutants by yeast surface display." Protein Eng **10**(11): 1303-1310.

Klammt, C., I. Maslennikov, M. Bayrhuber, C. Eichmann, N. Vajpai, E. J. Chiu, K. Y. Blain, L. Esquivies, J. H. Kwon, B. Balana, U. Pieper, A. Sali, P. A. Slesinger, W. Kwiatkowski, R. Riek and S. Choe (2012). "Facile backbone structure determination of human membrane proteins by NMR spectroscopy." Nat Methods **9**(8): 834-839.

Knowles, T. J., R. Finka, C. Smith, Y.-P. Lin, T. Dafforn and M. Overduin (2009). "Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer." J Am Chem Soc **131**(22): 7484-7485.

Kragh-Hansen, U., M. le Maire and J. V. Moller (1998). "The mechanism of detergent solubilization of liposomes and protein-containing membranes." Biophys J **75**(6): 2932-2946.

Krogh, A., B. Larsson, G. von Heijne and E. L. Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." J Mol Biol **305**(3): 567-580.

Kummer, L., C. W. Hsu, O. Dagliyan, C. MacNevin, M. Kaufholz, B. Zimmermann, N. V. Dokholyan, K. M. Hahn and A. Pluckthun (2013). "Knowledge-based design of a biosensor to quantify localized ERK activation in living cells." Chem Biol **20**(6): 847-856.

Kunji, E. R., M. Harding, P. J. Butler and P. Akamine (2008). "Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography." Methods **46**(2): 62-72.

Lafuse, W. P., G. R. Alvarez and B. S. Zwilling (2000). "Regulation of Nramp1 mRNA stability by oxidants and protein kinase C in RAW264.7 macrophages expressing Nramp1(Gly169)." Biochem J **351 Pt 3**: 687-696.

Lam-Yuk-Tseung, S., G. Govoni, J. Forbes and P. Gros (2003). "Iron transport by Nramp2/DMT1: pH regulation of transport by 2 histidines in transmembrane domain 6." Blood **101**(9): 3699-3707.

Le Bon, C., J. L. Popot and F. Giusti (2014). "Labeling and functionalizing amphipols for biological applications." J Membr Biol **247**(9-10): 797-814.

le Maire, M., P. Champeil and J. V. Moller (2000). "Interaction of membrane proteins and lipids with solubilizing detergents." Biochim Biophys Acta **1508**(1-2): 86-111.

Lee, S. Y., J. H. Choi and Z. Xu (2003). "Microbial cell-surface display." Trends Biotechnol **21**(1): 45-52.

Leitz, A. J., T. H. Bayburt, A. N. Barnakov, B. A. Springer and S. G. Sligar (2006). "Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology." Biotechniques **40**(5): 601-602, 604, 606, passim.

Li, J., A. Mahajan and M. D. Tsai (2006). "Ankyrin repeat: a unique motif mediating protein-protein interactions." Biochemistry **45**(51): 15168-15178.

Lichtenberg, D., H. Ahyayauch and F. M. Goni (2013). "The mechanism of detergent solubilization of lipid bilayers." Biophys J **105**(2): 289-299.

List, T. and D. Neri (2013). "Immunocytokines: a review of molecules in clinical development for cancer therapy." Clin Pharmacol **5**: 29-45.

Loewen, S. K., S. Y. M. Yao, M. D. Slugoski, N. N. Mohabir, R. J. Turner, J. R. Mackey, J. H. Weiner, M. P. Gallagher, P. J. F. Henderson, S. A. Baldwin, C. E. Cass and J. D. Young (2004). "Transport of physiological nucleosides and anti-viral and anti-neoplastic nucleoside drugs by recombinant Escherichia coli nucleoside-H+ cotransporter (NupC) produced in Xenopus laevis oocytes." Molecular Membrane Biology **21**(1): 1-10.

Ma, C. (2013). Prokaryote proteins as experimentally amenable models for eukaryote transporters Doctor of Philosophy Thesis, University of Leeds.

Ma, J. F. and J. Furukawa (2003). "Recent progress in the research of external Al detoxification in higher plants: a minireview." J Inorg Biochem **97**(1): 46-51.

Mackenzie, B. and M. D. Garrick (2005). "Iron Imports. II. Iron uptake at the apical membrane in the intestine." American Journal of Physiology - Gastrointestinal and Liver Physiology **289**(6): G981-G986.

Mackenzie, B. and M. A. Hediger (2004). "SLC11 family of H+-coupled metal-ion transporters NRAMP1 and DMT1." Pflugers Arch **447**(5): 571-579.

Makino, T., G. Skretas and G. Georgiou (2011). "Strain engineering for improved expression of recombinant proteins in bacteria." Microb Cell Fact **10**: 32.

Marks, J. D., H. R. Hoogenboom, T. P. Bonnert, J. McCafferty, A. D. Griffiths and G. Winter (1991). "By-passing immunization. Human antibodies from V-gene libraries displayed on phage." J Mol Biol **222**(3): 581-597.

McCafferty, J., A. D. Griffiths, G. Winter and D. J. Chiswell (1990). "Phage antibodies: filamentous phage displaying antibody variable domains." Nature **348**(6301): 552-554.

Mikolosko, J., K. Bobyk, H. I. Zgurskaya and P. Ghosh (2006). "Conformational flexibility in the multidrug efflux system protein AcrA." Structure **14**(3): 577-587.

Milovnik, P., D. Ferrari, C. A. Sarkar and A. Plückthun (2009). "Selection and characterization of DARPins specific for the neurotensin receptor 1." Protein Engineering Design and Selection **22**(6): 357-366.

Mittal, A., S. Bohm, M. G. Grutter, E. Bordignon and M. A. Seeger (2012). "Asymmetry in the homodimeric ABC transporter MsbA recognized by a DARPin." J Biol Chem **287**(24): 20395-20406.

Mohanty, A. K. and M. C. Wiener (2004). "Membrane protein expression and production: effects of polyhistidine tag length and position." Protein Expr Purif **33**(2): 311-325.

Monegal, A., D. Ami, C. Martinelli, H. Huang, M. Aliprandi, P. Capasso, C. Francavilla, G. Ossolengo and A. de Marco (2009). "Immunological applications of single-domain llama recombinant antibodies isolated from a naive library." Protein Eng Des Sel **22**(4): 273-280.

Monroe, N., G. Sennhauser, M. A. Seeger, C. Briand and M. G. Grutter (2011). "Designed ankyrin repeat protein binders for the crystallization of AcrB: plasticity of the dominant interface." J Struct Biol **174**(2): 269-281.

Moraes, I., G. Evans, J. Sanchez-Weatherby, S. Newstead and P. D. Stewart (2014). "Membrane protein structure determination - the next generation." Biochim Biophys Acta **1838**(1 Pt A): 78-87.

Murakami, S., R. Nakashima, E. Yamashita, T. Matsumoto and A. Yamaguchi (2006). "Crystal structures of a multidrug transporter reveal a functionally rotating mechanism." Nature **443**(7108): 173-179.

Murakami, S., R. Nakashima, E. Yamashita and A. Yamaguchi (2002). "Crystal structure of bacterial multidrug efflux transporter AcrB." Nature **419**(6907): 587-593.

Muyldermans, S. (2013). "Nanobodies: natural single-domain antibodies." Annu Rev Biochem **82**: 775-797.

Nada Jabado, S. l.-Y.-T., John R. Forbes and Philippe Gros (2004). Mouse Natural Resistance Associated Macrophage Protein 1 (Nramp1): A Key Player in Host Innate Immunity against Infections. The NRAMP Family. P. G. Mathieu Cellier, kluwer academic / plenum publishers**:** 216.

Nakashima, R., K. Sakurai, S. Yamasaki, K. Nishino and A. Yamaguchi (2011). "Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket." Nature **480**(7378): 565-569.

Oliver, R. C., J. Lipfert, D. A. Fox, R. H. Lo, S. Doniach and L. Columbus (2013). "Dependence of micelle size and shape on detergent alkyl chain length and head group." PLoS One **8**(5): e62488.

Otzen, D. (2011). "Protein–surfactant interactions: A tale of many states." Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1814**(5): 562-591.

Patching, S. G., S. A. Baldwin, A. D. Baldwin, J. D. Young, M. P. Gallagher, P. J. Henderson and R. B. Herbert (2005). "The nucleoside transport proteins, NupC and NupG, from Escherichia coli: specific structural motifs necessary for the binding of ligands." Org Biomol Chem **3**(3): 462-470.

Pavlidou, M., K. Hanel, L. Mockel and D. Willbold (2013). "Nanodiscs allow phage display selection for ligands to non-linear epitopes on membrane proteins." PLoS One **8**(9): e72272.

Pepper, L. R., Y. K. Cho, E. T. Boder and E. V. Shusta (2008). "A decade of yeast surface display technology: where are we now?" Comb Chem High Throughput Screen **11**(2): 127-134.

Pinner, E., S. Gruenheid, M. Raymond and P. Gros (1997). "Functional Complementation of the Yeast Divalent Cation Transporter Family SMF by NRAMP2, a Member of the Mammalian Natural Resistance-associated Macrophage Protein Family." Journal of Biological Chemistry **272**(46): 28933-28938.

Pluckthun, A. (2012). "Ribosome display: a perspective." Methods Mol Biol **805**: 3-28.

Pluckthun, A. (2015). "Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy." Annu Rev Pharmacol Toxicol **55**: 489-511.

Popot, J. L., E. A. Berry, D. Charvolin, C. Creuzenet, C. Ebel, D. M. Engelman, M. Flotenmeyer, F. Giusti, Y. Gohon, Q. Hong, J. H. Lakey, K. Leonard, H. A. Shuman, P. Timmins, D. E. Warschawski, F. Zito, M. Zoonens, B. Pucci and C. Tribet (2003). "Amphipols: polymeric surfactants for membrane biology research." Cell Mol Life Sci **60**(8): 1559-1574.

Pos, K. M. (2009). "Drug transport mechanism of the AcrB efflux pump." Biochim Biophys Acta **1794**(5): 782-793.

Postis, V., S. Rawson, J. K. Mitchell, S. C. Lee, R. A. Parslow, T. R. Dafforn, S. A. Baldwin and S. P. Muench (2015). "The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy." Biochim Biophys Acta **1848**(2): 496-501.

Postis, V., S. Rawson, J. K. Mitchell, S. C. Lee, R. A. Parslow, T. R. Dafforn, S. A. Baldwin and S. P. Muench (2015). "The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy." Biochimica et Biophysica Acta (BBA) - Biomembranes **1848**(2): 496-501.

Rahman, M., F. Ismat, M. J. McPherson and S. A. Baldwin (2007). "Topology-informed strategies for the overexpression and purification of membrane proteins." Mol Membr Biol **24**(5-6): 407-418.

Rajesh, S., T. Knowles and M. Overduin (2011). "Production of membrane proteins without cells or detergents." N Biotechnol **28**(3): 250-254.

Rawlings, A. E., J. P. Bramble, A. A. S. Tang, L. A. Somner, A. E. Monnington, D. J. Cooke, M. J. McPherson, D. C. Tomlinson and S. S. Staniland (2015). "Phage display selected magnetite

interacting Adhirons for shape controlled nanoparticle synthesis." Chemical Science **6**(10): 5586-5594.

Richer, E., P. Courville, I. Bergevin and M. F. Cellier (2003). "Horizontal gene transfer of "prototype" Nramp in bacteria." J Mol Evol **57**(4): 363-376.

Ries, J., C. Kaplan, E. Platonova, H. Eghlidi and H. Ewers (2012). "A simple, versatile method for GFP-based super-resolution microscopy via nanobodies." Nat Methods **9**(6): 582-584.

Ruigrok, V. J., M. Levisson, M. H. Eppink, H. Smidt and J. van der Oost (2011). "Alternative affinity tools: more attractive than antibodies?" Biochem J **436**(1): 1-13.

Russel, M., H. B. Lowman and T. Clackson (2004). Introduction to phage biology and phage display. Phage display: A Practical Approach. T. Clackson and H. B. Lowman. USA, Oxford University Press.

Salon, J. A., D. T. Lodowski and K. Palczewski (2011). "The significance of G protein-coupled receptor crystallography for drug discovery." Pharmacol Rev **63**(4): 901-937.

Schaffitzel, C., J. Hanes, L. Jermutus and A. Pluckthun (1999). "Ribosome display: an in vitro method for selection and evolution of antibodies from libraries." J Immunol Methods **231**(1-2): 119-135.

Schierle, C. F., M. Berkmen, D. Huber, C. Kumamoto, D. Boyd and J. Beckwith (2003). "The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway." J Bacteriol **185**(19): 5706-5713.

Schuler, M. A., I. G. Denisov and S. G. Sligar (2013). "Nanodiscs as a new tool to examine lipid-protein interactions." Methods Mol Biol **974**: 415-433.

Secco, D., C. Wang, B. A. Arpat, Z. Wang, Y. Poirier, S. D. Tyerman, P. Wu, H. Shou and J. Whelan (2012). "The emerging importance of the SPX domain-containing proteins in phosphate homeostasis." New Phytol **193**(4): 842-851.

Seddon, A. M., P. Curnow and P. J. Booth (2004). "Membrane proteins, lipids and detergents: not just a soap opera." Biochim Biophys Acta **1666**(1-2): 105-117.

Seelig, B. (2011). "mRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries." Nat Protoc **6**(4): 540-552.

Sendamarai, A. K., R. S. Ohgami, M. D. Fleming and C. M. Lawrence (2008). "Structure of the membrane proximal oxidoreductase domain of human Steap3, the dominant ferrireductase of the erythroid transferrin cycle." Proc Natl Acad Sci U S A **105**(21): 7410-7415.

Sennhauser, G., P. Amstutz, C. Briand, O. Storchenegger and M. G. Grutter (2007). "Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors." PLoS Biol **5**(1): e7.

Sennhauser, G. and M. G. Grutter (2008). "Chaperone-assisted crystallography with DARPins." Structure **16**(10): 1443-1453.

Sharma, K. S., G. Durand, F. Gabel, P. Bazzacco, C. Le Bon, E. Billon-Denis, L. J. Catoire, J. L. Popot, C. Ebel and B. Pucci (2012). "Non-ionic amphiphilic homopolymers: synthesis, solution properties, and biochemical validation." Langmuir **28**(10): 4625-4639.

Silacci, M., S. Brack, G. Schirru, J. Marlind, A. Ettorre, A. Merlo, F. Viti and D. Neri (2005). "Design, construction, and characterization of a large synthetic human antibody phage display library." Proteomics **5**(9): 2340-2350.

Simon, M., R. Frey, U. Zangemeister-Wittke and A. Pluckthun (2013). "Orthogonal assembly of a designed ankyrin repeat protein-cytotoxin conjugate with a clickable serum albumin module for half-life extension." Bioconjug Chem **24**(11): 1955-1966.

Smith, G. P. (1985). "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." Science **228**(4705): 1315-1317.

Soltes, G., M. Hust, K. K. Ng, A. Bansal, J. Field, D. I. Stewart, S. Dubel, S. Cha and E. J. Wiersma (2007). "On the influence of vector design on antibody phage display." J Biotechnol **127**(4): 626-637.

Stefan, N., P. Martin-Killias, S. Wyss-Stoeckle, A. Honegger, U. Zangemeister-Wittke and A. Pluckthun (2011). "DARPins recognizing the tumor-associated antigen EpCAM selected by phage and ribosome display and engineered for multivalency." J Mol Biol **413**(4): 826-843.

Stein, C., C. Kellner, M. Kugler, N. Reiff, K. Mentz, M. Schwenkert, B. Stockmeyer, A. Mackensen and G. H. Fey (2010). "Novel conjugates of single-chain Fv antibody fragments specific for stem cell antigen CD123 mediate potent death of acute myeloid leukaemia cells." Br J Haematol **148**(6): 879-889.

Steiner, D., P. Forrer and A. Pluckthun (2008). "Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display." J Mol Biol **382**(5): 1211-1227.

Steiner, D., P. Forrer and A. Plückthun (2008). "Efficient Selection of DARPins with Sub-nanomolar Affinities using SRP Phage Display." J Mol Biol **382**(5): 1211-1227.

Stuart, M. C. and E. J. Boekema (2007). "Two distinct mechanisms of vesicle-to-micelle and micelle-to-vesicle transition are mediated by the packing parameter of phospholipid-detergent systems." Biochim Biophys Acta **1768**(11): 2681-2689.

Tamaskovic, R., M. Simon, N. Stefan, M. Schwill and A. Pluckthun (2012). "Designed ankyrin repeat proteins (DARPins) from research to therapy." Methods Enzymol **503**: 101-134.

Theurillat, J. P., B. Dreier, G. Nagy-Davidescu, B. Seifert, S. Behnke, U. Zurrer-Hardi, F. Ingold, A. Pluckthun and H. Moch (2010). "Designed ankyrin repeat proteins: a novel tool for testing epidermal growth factor receptor 2 expression in breast cancer." Mod Pathol **23**(9): 1289-1297.

Thie, H., T. Schirrmann, M. Paschke, S. Dubel and M. Hust (2008). "SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli." N Biotechnol **25**(1): 49-54.

Thomas, W. D. and G. P. Smith (2010). "The case for trypsin release of affinity-selected phages." Biotechniques **49**(3): 651-654.

Tiede, C., A. A. Tang, S. E. Deacon, U. Mandal, J. E. Nettleship, R. L. Owen, S. E. George, D. J. Harrison, R. J. Owens, D. C. Tomlinson and M. J. McPherson (2014). "Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications." Protein Eng Des Sel **27**(5): 145-155.

Tribet, C., R. Audebert and J. L. Popot (1996). "Amphipols: polymers that keep membrane proteins soluble in aqueous solutions." Proc Natl Acad Sci U S A **93**(26): 15047-15050.

Trinder, D., C. Fox, G. Vautier and J. K. Olynyk (2002). "Molecular pathogenesis of iron overload." Gut **51**(2): 290-295.

Vaish, A., V. Silin, M. L. Walker, K. L. Steffens, S. Krueger, A. A. Yeliseev, K. Gawrisch and D. J. Vanderah (2013). "A generalized strategy for immobilizing uniformly oriented membrane proteins at solid interfaces." Chem Commun (Camb) **49**(26): 2685-2687.

Valent, Q. A., P. A. Scotti, S. High, J. W. de Gier, G. von Heijne, G. Lentzen, W. Wintermeyer, B. Oudega and J. Luirink (1998). "The Escherichia coli SRP and SecB targeting pathways converge at the translocon." EMBO J **17**(9): 2504-2512.

Vaziri, H., S. A. Baldwin, J. M. Baldwin, D. G. Adams, J. D. Young and V. L. Postis (2013). "Use of molecular modelling to probe the mechanism of the nucleoside transporter NupG." Mol Membr Biol **30**(2): 114-128.

Veesler, D., S. Blangy, C. Cambillau and G. Sciara (2008). "There is a baby in the bath water: AcrB contamination is a major problem in membrane-protein crystallization." Acta Crystallogr Sect F Struct Biol Cryst Commun **64**(Pt 10): 880-885.

White, J. K., A. Stewart, J. F. Popoff, S. Wilson and J. M. Blackwell (2004). "Incomplete glycosylation and defective intracellular targeting of mutant solute carrier family 11 member 1 (Slc11a1)." Biochem J **382**(Pt 3): 811-819.

Wilson, D. S., A. D. Keefe and J. W. Szostak (2001). "The use of mRNA display to select high-affinity protein-binding peptides." Proc Natl Acad Sci U S A **98**(7): 3750-3755.

Zahnd, C., P. Amstutz and A. Pluckthun (2007). "Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target." Nat Methods **4**(3): 269-279.

Zahnd, C., E. Wyler, J. M. Schwenk, D. Steiner, M. C. Lawrence, N. M. McKern, F. Pecorari, C. W. Ward, T. O. Joos and A. Pluckthun (2007). "A designed ankyrin repeat protein evolved to picomolar affinity to Her2." J Mol Biol **369**(4): 1015-1028.

Zhang, A. S., F. Canonne-Hergaux, S. Gruenheid, P. Gros and P. Ponka (2008). "Use of Nramp2-transfected Chinese hamster ovary cells and reticulocytes from mk/mk mice to study iron transport mechanisms." Exp Hematol **36**(10): 1227-1235.

Zoonens, M. and J. L. Popot (2014). "Amphipols for each season." J Membr Biol **247**(9-10): 759-796.

Zoonens, M., F. Zito, K. L. Martinez and J.-L. Popot (2014). Amphipols: A General Introduction and Some Protocols. Membrane Proteins Production for Structural Analysis. I. Mus-Veteau, Springer New York**:** 173-203.

Appendix 1

General lab chemicals were procured from the following suppliers:

All lipids were bought from Avanti Polar Lipids Inc. http://www.avantilipids.com/

BDH Laboratory Supplies Merck Ltd, http://www.bdhme.com/product-lab.htm

PageRular™ plus pre-stained ladder, precision plus protein™ standards were purchased on **Bio-Rad Microscience Ltd**., http://www.bio-rad.com/

Disposable PD-10 desalting columns and the superdex 200 10/300 GL columns were purchased from **GE Healthcare UK**, http://www.gelifesciences.com/

Carbenicillin, Isopropyl β-D-1-thiogalactopyranoside (IPTG), Dithiothreitol (DTT was purchased from **Melford Laboratories Ltd**., http://www.melford.co.uk/

The anti polyhistidine antibody/ HRP conjugate was purchased from **R&D systems**

All other chemicals were purchased from **Sigma-Aldrich Chemical Company**, http://www.sigmaaldrich.com/

Supersignal® west pico chemiluminescence substrate, *EZ-Link® Maleimide-PEG$_{11}$-Biotin,* Hispur™ cobalt resin were purchased from **Thermo Fisher Scientific**, http://www.piercenet.com/

Reagents for molecular biology, protein purification were obtained from the following manufacturers:

BDH Laboratory Supplies Merck Ltd, http://www.bdhme.com/product-lab.htm

Bio-Rad Microscience Ltd., http://www.bio-rad.com/

Fisher Scientific UK Ltd, http://www.fisher.co.uk/

DDM was purchased from Glycon Biochemicals, http://www.glycon.de/

http://www.merckmillipore.co.uk/chemicals

Fluozin-1 was purchased from Invitrogen Life Technologies, http://www.invitrogen.com/site/us/en/home.html

Melford Laboratories Ltd., http://www.melford.co.uk/

Merck Millipore (also for Calbiochem®, Novabiochem® and Novagen®),

Restriction endonucleases, buffers and enzymes New England Biolabs, http://www.neb.com/nebecomm/default.asp

Wizard® SV gel and PCR clean up system kit, Wizard® Plus SV miniprep kit were purchased from Promega, http://www.promega.com/

Qiagen, http://www.qiagen.com/default.aspx

EDTA free protease inhibitor cocktail was purchased from Roche Diagnostics, http://www.roche.co.uk/portal/uk/diagnostics

Severn Biotech, Ltd, http://www.severnbiotech.com/

Sigma-Aldrich Chemical Company, http://www.sigmaaldrich.com/

Thermo Fisher Scientific, http://www.piercenet.com/

Appendix 2

The DNA sequence of the gene encoding His_6-tagged MPSIL0294, aligned with the corresponding amino acid sequence:

```
ATGCCTAGGAAAAATTCAGAAGAACATGAACCAAAGCAAAGACATCATTTGATTGAATAT
 M  P  R  K  N  S  E  E  H  E  P  K  Q  R  H  H  L  I  E  Y
GCAAATGGTCCTTCGCTTGAAGAAATTAATGGCACCATCGACGTGCCTAAAAATATGAGT
 A  N  G  P  S  L  E  E  I  N  G  T  I  D  V  P  K  N  M  S
TTTTGGAAAACGTTATTTGCTTACTCAGGTCCAGGAGCATTGGTAGCAGTGGGGTATATG
 F  W  K  T  L  F  A  Y  S  G  P  G  A  L  V  A  V  G  Y  M
GATCCAGGAAACTGGTCTACTTCAATTACTGGGGGACAAAATTTTCAATATTTATTGATG
 D  P  G  N  W  S  T  S  I  T  G  G  Q  N  F  Q  Y  L  L  M
TCGATTATTTTAATTTCCAGTTTGATTGCAATGTTGCTCCAATATATGGCTGCTAAATTA
 S  I  I  L  I  S  S  L  I  A  M  L  L  Q  Y  M  A  A  K  L
GGCATTGTTTCACAGATGGATTTAGCACAAGCGATTCGTGCTAGAACTAGTAAGACGTTA
 G  I  V  S  Q  M  D  L  A  Q  A  I  R  A  R  T  S  K  T  L
GGTATTGTATTATGGATTTTAACAGAGTTAGCAATTATGGCTACAGATATCGCTGAAGTT
 G  I  V  L  W  I  L  T  E  L  A  I  M  A  T  D  I  A  E  V
ATCGGGGGCGCAATTGCCTTATATTTATTATTTCATATTCCTTTAGGTCTGGCTGTCTTC
 I  G  G  A  I  A  L  Y  L  L  F  H  I  P  L  G  L  A  V  F
ATTACGGTATTTGATGTTTTACTTTTGTTGTTATTGACAAAAATTGGTTTTAGAAAAATT
 I  T  V  F  D  V  L  L  L  L  L  L  T  K  I  G  F  R  K  I
GAAGCTATTGTTGTTGCTTTAATTGTTGTTATTTTTGTGATTTTTGCTTATCAAGTGGCA
 E  A  I  V  V  A  L  I  V  V  I  F  V  I  F  A  Y  Q  V  A
TTGTCAAATCCAGTATGGGGAGATGTAATTAAAGGGCTGGTTCCTAGCGCAGAAGCTTTT
 L  S  N  P  V  W  G  D  V  I  K  G  L  V  P  S  A  E  A  F
```

TCTACATCACATGCGGTGAATGGGCAAACGCCACTGACAGGTGCATTGGGGATTATTGGT

S T S H A V N G Q T P L T G A L G I I G

GCAACAGTGATGCCTCATAATTTGTATTTACATTCTTCTGTTGTGCAAAGTCGTAAAATT

A T V M P H N L Y L H S S V V Q S R K I

GATCGTAAAGATAAAACCGATATTCAACGCGCTTTACGCTTTTCAACTTGGGATTCTAAT

D R K D K T D I Q R A L R F S T W D S N

ATTCAATTAACGATGGCATTTTTTGTTAACTCTTTATTATTGATTATGGGTGTGGCCGTT

I Q L T M A F F V N S L L L I M G V A V

TTTAAATCCGGCAGTGTCAAAGATCCTTCATTCTTTGGTTTGTTTGATGCTTTATCAAAT

F K S G S V K D P S F F G L F D A L S N

CCAGCGGTTATGAGTAATTCTATTTTAGCGCATATTGCAGGTTCTGGAATTTTATCAATT

P A V M S N S I L A H I A G S G I L S I

TTATTTGCGGTAGCCTTATTGGCATCAGGACAAAATTCAACAATTACAGGAACGTTAACT

L F A V A L L A S G Q N S T I T G T L T

GGTCAAATCATTATGGAAGGGTTTATTCATATGCGCGTACCAATTTGGTTACGTCGGATG

G Q I I M E G F I H M R V P I W L R R M

GTCACACGTTTGTTATCTGTTATTCCTGTCTTAATCTGTGTTTTAATGACCAGCGGAAAA

V T R L L S V I P V L I C V L M T S G K

AGTACAGTGGAGGAGCATATTGCGATTAATAATTTAATGAACAATTCACAAGTTTTTCTA

S T V E E H I A I N N L M N N S Q V F L

GCATTTGCTTTGCCATTTTCGATGTTGCCTTTACTGATGTTTACAGATAGCCGTGTTGAA

A F A L P F S M L P L L M F T D S R V E

ATGGGTGAACATTTTAAAAACTCGTGGTTAATTAAATTGTTAGGCTGGGTTTCTGTCATT

M G E H F K N S W L I K L L G W V S V I

GGCTTAATTTACCTCAATATGAAAGGTTTACCTGATCAAATTGAAGGATTCTTCGGTGAT

G L I Y L N M K G L P D Q I E G F F G D

AATCCGACAGCGAGTCAAATTACGTTAGCCGATAATATTGCCTATGTCATCATAGCACTC

N P T A S Q I T L A D N I A Y V I I A L

GTCATCCTCTTGTTAGTTTGGACGGTTGTGGAATTATATAAAGGCGATAAACGATATGCA

V I L L L V W T V V E L Y K G D K R Y A

CAGCAGCTTGCAGCTATGGAGCAACAAGTAGAGGAGGTTAAACCTGCAGGACTGGAAGTA

Q Q L A A M E Q Q V E E V K P A G L E V

CTATTTCAAGGACCACATATGCGTGGCAGCCACCATCACCATCACCATCACCAT

L F Q G P H M R G S H H H H H H H H

Note: Highlighted in blue is the codon which was altered by site-directed mutagenesis in order to construct pVA3. Highlighted in yellow is the sequence of the Arg-Gly-Ser-His_6-tag.

The DNA sequence of the Avitag-encoding region present in plasmid pVA1, aligned with the corresponding amino acid sequence:

CTGGAAGTTCTGTTCCAGGGTCCGCATATGGGCAGCGGTCTGAACGACATCTTCGAAGCTCAGAAAATCGAATGGCACGAA

L E V L F Q G P H M G S G L N D I F E A Q K I E W H E

GGCAGCCGTGGCAGCCACCATCACCATCACCATCACCAT

G S R G S H H H H H H H H

Note: Highlighted in grey is an HRV cleavage site. In red is the sequence of the avitag with the lysine to which biotin becomes attached highlighted in blue. The Arg-Gly-Ser-His_6-tag is highlighted in yellow.

The DNA sequence of the SBP tag-encoding region present in plasmid pVA2, aligned with the corresponding amino acid sequence:

CTGGAAGTTCTGTTCCAGGGTCCGCATATGGGCAGCATGGACGAAAAAACCACCGGTTGGCGTGGTGGTCACGTTGTTGAA

L E V L F Q G P H M G S M D E K T T G W R G G H V V E

GGTCTGGCTGGTGAACTGGAACAGCTGCGTGCTCGTCTGGAACACCACCCGCAGGGTCAGCGTGAACCGGGCAGCCGTGGC

G L A G E L E Q L R A R L E H H P Q G Q R E P G S R G

AGCCACCATCACCATCACCATCACCAT

S H H H H H H H H

Note: Grey highlighting indicates the position of an HRV cleavage site. The SBP tag responsible for binding streptavidin is highlighted in green. The Arg-Gly-Ser-His_6-tag is highlighted in yellow.