This is a repository copy of *Membrane protein engineering to the rescue*.

Article:

Membrane protein engineering to the rescue.

Andrea E Rawlings*

Department of Chemistry, The University of Sheffield, Sheffield, UK.

*a.rawlings@sheffield.ac.uk

Abstract

The inherent hydrophobicity of membrane proteins is a major barrier to membrane protein research and understanding. Their low stability and solubility in aqueous environments coupled with poor expression levels make them a challenging area of research. For many years the only way of working with membrane proteins was to optimise the environment to suit the protein, through the use of different detergents, solubilising additives, and other adaptations. However, with innovative protein engineering methodologies the membrane proteins themselves are now being adapted to suit the environment. This mini review looks at the types of adaptations which are applied to membrane proteins from a variety of different fields, including; water solubilising fusion tags, thermostabilising mutation screening, scaffold proteins, stabilising protein chimeras, and isolating water soluble domains.

Introduction

Proteins tend to be classified into one of two broad categories, globular cytoplasmic proteins, and integral membrane proteins (IMPs)[1]. This simple classification is defined by a proteins intrinsic ability to produce a stable, functional structure in either a hydrophilic or hydrophobic environment. Analysis of published genomes suggests that between 15% and 30% of the identified open reading frames encode proteins which would fit into the latter, membrane embedded, category[2,3]. Yet, despite seemingly so many membrane proteins available, these hydrophobic proteins lag significantly behind their water soluble counterparts in terms of characterisation[4]. For instance, less than 1% of the resolved protein structures deposited in the Protein Data Bank (PDB) are membrane proteins[5]. This statistic belies the significant effort and strides made in membrane protein research over the past decades.

Understanding a protein in terms of function, usually requires an understanding of its three dimensional structure, and the current paucity of membrane protein structures produces a knowledge gap which has an impact on a variety of research areas in which membrane proteins play a key role[3]. For instance, membrane proteins are prime targets for therapeutics due to their important roles as key mediators of molecular transport into and out of the cell, as well in other crucial processes such as cell signalling[6]. The knowledge gap in membrane protein research is a major bottleneck to the development of future biomedical and biotechnological advances.

The main barrier to accelerated membrane protein research is their intrinsic hydrophobicity. This single factor produces difficulties at every stage of the protein production pipeline[3]. Membrane proteins have, by their very nature, a predominance of alkyl and aryl residues on their surface which make hydrophobic interactions with the surrounding lipid. This hydrophobic interaction is the main driver of their folding and stability in the lipid environment. Most analytical, characterisation, and structural techniques, such as X-ray crystallography, NMR, and to an increasing extent electron microscopy, rely on relatively large quantities of protein which are soluble in aqueous conditions[3]. As most membrane proteins are produced at very low levels in their natural host, they are difficult to isolate from native membranes. High levels of expressed protein in easily amplified cells such as E. coli, which can then be adequately solubilised, is a necessity.

Expressed membrane proteins must be incorporated into the cell membrane so that they assume the correct fold and retain their functional state. Such a small amount of available membrane relative to the total available volume within the cell places a limitation on the amount of viable membrane protein that any cell can produce. In addition, many membrane proteins produce large insoluble aggregates, inclusion bodies, within the cytoplasm. In order to maximise the amount of the target protein in the membrane, the expression conditions (typically host strain, growth medium, induction temperature and duration) will require extensive time consuming optimisation to achieve sufficient levels of protein for downstream processing[7], and in many cases this may never be realised. This is especially true for membrane transporters. Overexpression of transport proteins can produce an unwanted increase in the movement of molecules and ions across the cell membrane, causing significant toxicity and a reduction in culture growth and protein yield.

Membranes, harbouring the target protein, must be harvested from a cell lysate and solubilised so that they can be purified under aqueous conditions. Conventional solubilisation strategies focus on the addition of detergents[8–10]. These amphiphilic molecules self-assemble in solution, forming micelle structures which can encompass the membrane proteins, burying the hydrophobic surface residues in the core of the micelle, and maintaining the protein's structure close to that present in the natural lipid environment[11]. Finding a detergent which offers good levels of solubilisation whilst retaining the proteins structure requires painstaking optimisation and trial and error testing of a range of detergents

with different chain lengths and head groups[9,12,13]. In addition, the pH of the buffer, salinity levels, and the presence or absence of various other additives will all play a role in determining the final stability of the membrane protein[13,14]. A protein which is only stable and functional in a narrow window of environmental conditions is not only difficult to handle and characterise in the laboratory, but also curbs its potential use in biomedical or biotechnological applications. New additives to replace the use of detergents, such as nanodiscs[15,16], SMALPs[17], bicelles[18] and lipid like peptides (LLP)[19] are now emerging.

With the advent of modern genetic engineering techniques and the ability to alter genetic sequences easily and cheaply, there are ways to overcome many of these challenges for membrane protein production by manipulating the protein itself. This mini review looks at some of the options available and provides examples of where changes to the membrane protein itself have been used with excellent results.

Isolating water soluble domains from a membrane protein.

One approach to the investigation of membrane proteins is to decide if the membrane spanning region of the protein is actually required at all. Many membrane proteins possess water soluble domains which may be exposed to the cytoplasm, periplasm or even extracellularly. These domains are often sites of catalytic activity. The activity of such domains can provide an easily assayable function for structure-function investigations. Working on a water soluble domain overcomes many of the inherent challenges of membrane proteins, with improved expression levels, greater stability and ease of characterisation. A recent example of a truncated membrane protein investigated in this way is the phosphoethanolamine transferase MCR-1 from Enterobacteriaceae[20]. The crystal structure of this protein is helping to elucidate the mechanisms behind certain antibiotic resistance.

From a protein design perspective, the biggest obstacle is correctly determining the domain boundaries. This is a straightforward procedure where sequence alignments to well characterised protein families exist. However, for some protein groups this can be anything but a trivial exercise. Truncating the protein too early may remove stabilising residues, and too late may mean additional, potentially destabilising residues are fused to the domain of interest. High throughput solubility screening strategies can help. The expression of soluble proteins by random incremental truncation (ESPRIT) is one such method[21]. This library based approach, requiring very little background knowledge of the protein of interest, uses robotic screening to analyse potentially thousands of clones.

In essence the gene encoding the protein of interest is enzymatically degraded, a nucleotide at a time, to produce a spectrum of sequences which become incrementally smaller[21]. Each library member is fused to sequences encoding a C-terminal biotin acceptor peptide [22] and an N-terminal poly-histidine

tag. The biotinylation peptide acts as a solubility reporter, as soluble proteins will be efficiently biotinylated in vivo. Only proteins displaying both a positive signal for the N and C terminal tags will be (a) soluble, (b) in the desired reading frame, and (c) free from proteolytic degradation[21] (Figure 1A). The ESPRIT methodology was used to investigate the phosphatase domain of the sporulation protein SpoIIE from Bacillus subtilis [23]. SpoIIE has an N-terminal membrane spanning domain with ten transmembrane helices (Figure 1B). The remainder of the protein is displayed to the inside of the spore compartment during stage III of the sporulation pathway, where it forms part of a sigma factor cascade, enabling the spore and mother cell compartments to establish different fates[24,25]. By truncating the spoIIE gene incrementally, the domain boundaries were isolated and soluble protein purified, something that was difficult to achieve from analysis of the sequence alone[23] (Figure 1B). Based on this technique the crystal structure of a domain swapped dimer of the phosphatase was determined Figure 1C), confirming it belonged to the PP2C family of protein phosphatases, and allowed an intramolecular signalling surface to be identified[26].

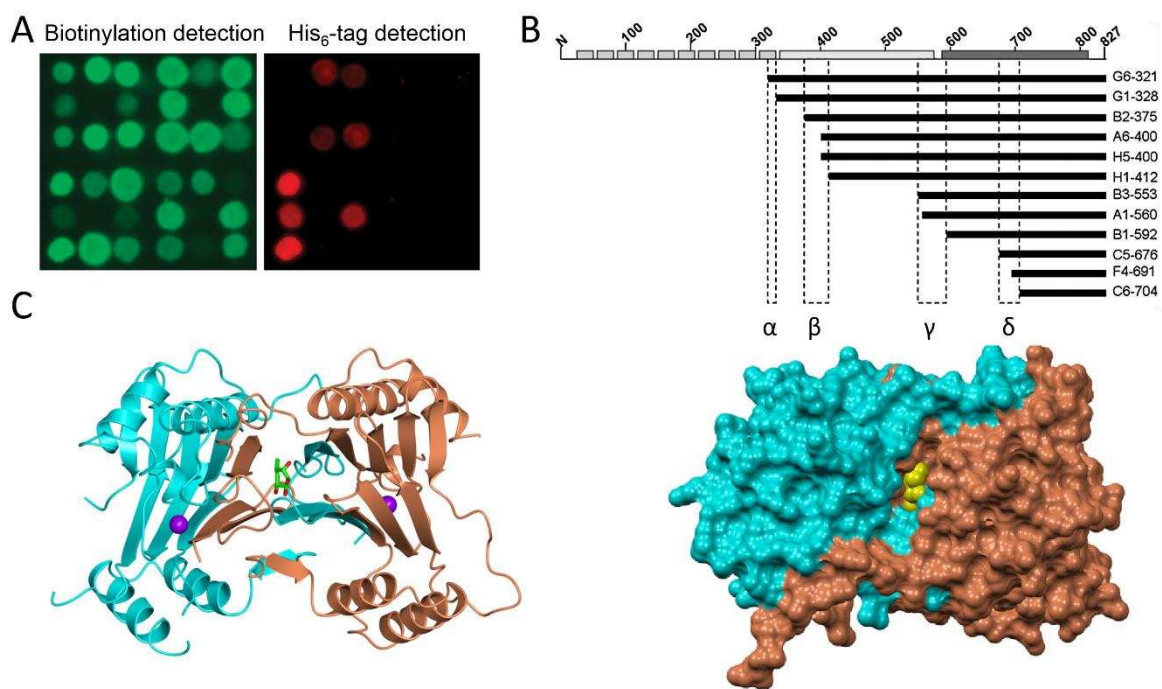**Figure 1.** (A) A portion of the high throughput screening array of SpoIIE truncation mutants. (B) The SpoIIE protein with soluble constructs indicated and residue positions marked. Dashed boxes highlight the postulated domain boundaries α-δ. (C) Cartoon and space filling models of SpoIIE domain swapped dimer, coloured cyan and orange. Figure adapted from Rawlings et al[23] and Levdikov et al with permission[26].

Producing water soluble IMPs with solubilising fusion tags.

A recent development in membrane protein engineering is to retain the native membrane protein sequence, but introduce a fusion tag which overcomes the insolubility of the membrane protein in aqueous solution[27]. Mizrachi et al have pioneered such an approach [28,29], termed SIMPLEx (solubilisation of IMPs with high levels of expression). They employ an amphipathic protein fusion partner, added to the membrane protein termini[29]. In this case the C-terminal lipid binding domain of apolipoprotein was selected due its intrinsic amphipathic nature, and its inherent structural flexibility[30]. To prevent undesired incorporation of the membrane protein in the inner cell membrane of E. coli a decoy protein, such as maltose binding protein (MBP) was added to the N-termini of the IMP. Having such a large soluble protein at the N-termini blocks transfer of the IMP to the membrane, ensuring efficient solubilisation via the apolipoprotein fusion tag[29]. Using this methodology, a range of different IMPs have been produced without the need for additional solubilising additives (detergents for example). The list of IMPs currently explored with this technique extends to cytochrome B5, the ethidium multidrug resistance protein E (EmrE), Claudin-1 and Claudin-3, DsbB and OmpX[29]. Significantly, IMPs from both bacterial and mammalian systems have been tested with equal success[29]. However, careful characterisation is required to ensure that the membrane protein retains the correct structure and function, and that it has not been unintentionally compromised by the addition of the solubilising tag. The ability to produce human IMPs in this way opens up the approach for drug discovery applications, where the challenges of producing adequate quantities of purified membrane protein is a major hurdle[31].

The implications of this methodology have wider applications. For instance, the SIMPLEx approach has been used to produce a water soluble form of the bacterial protein DsbB[29,32]. This enzyme, in the presence of ubiquinone, is capable of catalysing the oxidation of the thiodisulphide oxidoreductase partner enzyme DsbA within the bacterial periplasm[33][34] (Figure 2A). DsbA is exported to the periplasmic space via a defined N-terminal export sequence. DsbB is a membrane bound enzyme which inserts into the inner membrane and thereby limits the oxidising activity of the protein to the periplasm. By producing a water soluble form of DsbB via the addition of the apolipoprotein and decoy fusion tags, this protein can be overexpressed to high levels within the cytoplasm[32]. In that study the partner protein, DsbA, had its periplasmic export sequence removed, meaning it too was retained within the cytoplasm[32].   This allowed the DsbB-DsbA disulphide bond catalysis pathway, which was previously limited to the periplasm, to be fully constituted within the cytoplasm. The modified form of DsbB was still able to interact with DsbA as shown by crosslinking and small angle X-ray scattering (SAXS) studies (Figure 2B)[32]. The recompartmentalisation of this pathway in the cytoplasm was also enzymatically active, displaying disulphide bond forming capability against several different substrate proteins[32]. Excitingly, this result demonstrates that by engineering IMPs in this way, enzymatic activities previously only associated with lipid membranes, can be transplanted

into different cellular compartments and non-native environments without the need for additional solubilising additives.
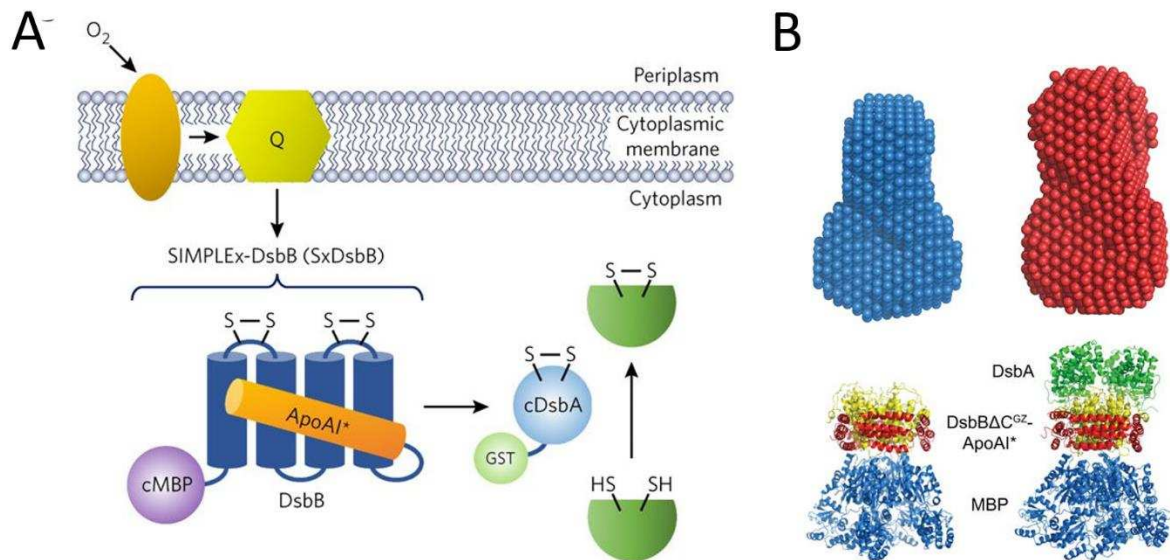

Figure 2. (A) The reconstituted cytoplasmic DsbA/DsbB system. DsbB is solubilised via ApoAI* and the decoy protein MBP. (B) SAXS model of the DbsB-ApoAI*-MBP construct in blue, with the DsbA bound form in red. Crystals structures of the various components have been used to produce models of the fusion protein to fit the SAXS structure (PDB accession numbers for the structures used are: 2K74 for monomeric DsbB, 2LEG for the complex DsbA/DsbB, 2A01 for ApoAI, and 1NL5 for MBP). Figure adapted from Mizrachi et al[32] with permission.

A slightly different approach to generate water soluble membrane proteins is via the introduction of multiple residue substitutions, where typically hydrophobic surface residues are switched for hydrophilic ones. This approach is often guided by computational modelling in order to pick the most appropriate substitutions, and has been successfully used to generate various water soluble forms of membrane proteins including KcsA[35] and the Ca^{2+} pump regulating protein phospholamban[36,37].

Conformational thermostabilisation of GPCRs.

One of the highest priority IMP classes for structural and biophysical characterisation are G protein-coupled receptors (GPCRs). These proteins are found within the membranes of eukaryotic cells[38]. GPCRs are involved in a huge array of different cellular processes. These IMP's feature seven transmembrane spanning helices (TMS) which enable signal transduction across a cell membrane [39]. These proteins feature specific binding sites for ligands, which stimulates the GPCR[40]. The interacting ligands which activate the receptors can be the natural receptor molecule or a synthetic ligand. This ability to elicit a GPCR response with synthetic molecules makes these proteins prime

targets for pharmaceuticals[41], and they represent the site of action for approximately 35% of existing drugs[42].

One of the main difficulties in realising more GPCR crystal structures is the inherent structural flexibility in these proteins[43]. They display a highly dynamic structure, with a large degree of conformation flexibility[44]. For a protein to crystallise it needs to assume a regular, ordered assembly of repeating protein units. Flexibility within the protein structure means each protein has the potential to assume one of a large number of possible conformations; this makes the crystals unlikely to form or have the long range order required for X-ray diffraction studies. An additional problem to overcome is that GPCR's often have quite short loops connecting the different transmembrane spanning helices. The lack of significant hydrophilic regions means a shortage of protein-protein contacts in the crystal lattice, leading to instability of the protein crystal and poor diffraction quality. However, there are protein engineering methodologies which can help in both of these regards[43].

An innovative approach is termed conformational thermostabilisation. This methodology has the ability to produce a stabilised GPCR with a favoured conformation, making it ideal for crystallisation and structural characterisation[43,45]. This technique relies on a program of single residue substitutions, typically switched to alanine, systematically introduced at each position along the protein sequence[46][47]. Each variant is then expressed in cells, detergent solubilised, and heated to just above the melting temperature (Tm) of the native protein. Using a radiolabelled version of a ligand which binds to the GPCR, allows researchers to easily and rapidly determine if the radioligand remains bound to the detergent solubilised protein after heating. This technique relies on the fact that the ligand binding site of GPCRs tends to be formed from residues displayed on various parts of the protein[40]. If the protein is heated and subsequently denatures, the specific architecture of the protein responsible for producing the binding site would be lost[48]. Therefore binding/non-binding of the specific radioligand gives a very convenient indicator of thermostability.

Preparing, expressing and screening these mutations (potentially hundreds of constructs for a single GPCR) is a large undertaking, but using high throughput techniques and automation where necessary can reduce some of the experimental burden[46]. An alternative strategy is to generate random mutations by processes such as error prone PCR and screen for binding on the pool of variants using fluorescent ligands and fluorescently activated cell sorting (FACS) equipment[49,50]. In both cases, by determining the individual substitutions which give rise to the largest increase in Tm, the most stabilising mutations can be collected within a single construct. However, the most elegant addition to this procedure is the use of a specific agonist or antagonist during the thermal screening steps in order to find mutations which stabilise the protein in an active or inactive state. There are challenges to this technique, namely that it requires functional well expressing membrane proteins, as well as ready access to a specific ligand, and the infrastructure to produce and screen a high number of single

residue substitutions (hundreds of constructs in most cases). However, the rewards appear to be make up for these difficulties with numerous examples of structures obtained through this approach[51–55].

Computational advances are also helping to lighten the load of some of these limitations. Using different strategies ranging from free energy calculations to machine learning, allows a suite of candidate mutations to be rapidly screened for stability in silico[56,57]. To reduce the range of potential mutants most systematic mutagenesis strategies only substitute each residue for alanine but it has been estimated that in around 30% of cases alanine may not produce the optimum stabilization[47]. A computational approach allows each residue to be swapped for any desired amino acid, thereby allowing much greater mutational scope to be explored. Computational methodologies appear extremely promising, particularly with the increase in GPCR structures from which to build realistic homology models. Due to the reductions in both cost and time it can bring, many laboratories not equipped for high throughput screening would have the potential to benefit from a thermostabilising methodology.

Stabilisation through protein chimeras.

GPCRs, due to their prevalence as therapeutic drug targets, are at the forefront of membrane protein innovations. Over the past decade various teams of researchers have made chimeric proteins, whereby a loop connecting two transmembrane spanning helices is exchanged for a stable, soluble, protein[58–60]. By incorporating the new protein into the IMP a significant increase in hydrophilicity is achieved. If chosen carefully, the new fusion partner can promote the formation of protein crystals. Key requirements for judicious selection of fusion partners are; that the protein is extremely stable, is already structurally characterised, and that the N and C termini are located closely together to allow them to be inserted into a loop of an IMP without distorting the arrangement of the transmembrane helices. A recent example of this is the Orexin receptor, OX_2R[58]. This GPCR is a member of the rhodopsin family and forms part of the sleep-wake cycle in mammals where it binds to neuropeptides of the central nervous system[61]. Yin et al prepared a chimera of OX_2R by substituting the interconnecting loop 3 (IL3) with PGS, a glycogen synthase from Pyrococcus abysii[58]. This 196 residue protein is extremely stable, as it is derived from a thermophilic organism, benefitted from a high resolution crystal structure, and featured termini less than 10 Å apart. The chimeric protein was crystallised and yielded a crystal structure of OX_2R at 2.5 Å resolution[58]. The binding interaction with the anti-insomnia drug Suvorexant was revealed, and novel potential binding sites for future agonists were identified via computational docking to the OR_2X model[58]. A limitation of this approach is that the protein is missing one loop which may have unintended consequences for protein conformation or ligand interaction. This is often minimised by selecting loops which are less likely to form part of the ligand binding site.
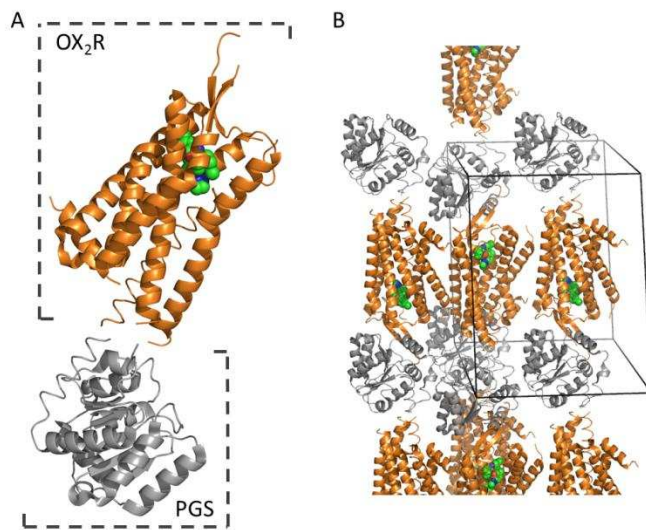
Figure 3. (A) Cartoon structure of the OX$_2$R–PGS fusion protein. The PGS domain is indicated in grey and is located at interconnecting loop 3. The Suvorexant ligand is depicted as spheres and coloured green for carbon, red for oxygen, and blue for nitrogen. (B) Packing of OX$_2$R–PGS in the crystal lattice. Note, many of the crystal contacts occur through the PGS domain[58]. Figure prepared using PyMOL[62].

Membrane protein chimeras with green fluorescent protein (GFP) have been used to find membrane protein constructs with improved stability, expression, or membrane insertion[63–67]. The intrinsic fluorescence of GFP makes it a natural choice of reporter protein, which can be easily assayed using fluorescence detection. In these cases GFP (or a form of split GFP to minimise any stabilising effect from the GFP itself) is typically fused to one of the termini of the membrane protein, allowing the protein to be tracked during various stages of its production and purification. This approach can, for instance, be coupled to targeted or random mutagenesis of the protein to find variants with improved expression levels and detergent stability[63,67].

If the part of the membrane protein which is of interest is one of the interconnecting loop regions and the transmembrane spanning portion is simply an anchor, then there are opportunities to transplant the desired loop and graft it onto a scaffold protein. This approach has been used to investigate the loop regions of bacterial membrane proteins from Magnetotactic bacteria (MTB). These bacteria are able to produce crystals of magnetic iron minerals (magnetite or goethite) within specialised membrane bound organelles in their cells termed magnetosomes[68–70]. The magnetosomes harbour a range of different proteins, embedded within the membrane, which enable the production of the crystals[71,72]. Some of these proteins are found tightly bound to the crystal, interacting via the loops between transmembrane spanning helices. In one example the loop region from MamC was transplanted onto maltose binding protein (MBP)[73]. This allowed the crystal structure of the loop to be obtained that revealed a helical conformation, from which a proposed magnetite interaction site was derived[73].

Protein produced using this approach was tested in in vitro magnetite formation reactions to ascertain whether this construct retained their native functionality. Inspection of the synthesised magnetite nanocrystals showed these proteins were able to influence the properties of synthetic magnetite nanoparticles[73].

Conclusions

Membrane proteins are some of the most challenging biomolecules to produce, isolate and characterise, and yet they often perform some of the most interesting and useful roles within the protein function repertoire. This makes them a rich source of reagents for use in fields as wide ranging as pharmaceuticals to nanoparticle formation. Their inherent hydrophobicity and stability within lipid environments creates obstacles to their wider use and exploitation. However, advances within the field of protein engineering are providing new tools to overcome these difficulties. Proteins can be edited to yield water soluble domains[23], fused to solubilising protein tags to produce water soluble variants[29], or have their stability enhanced to aid in crystallisation or detergent compatibility[46]. The ingenuity of protein engineers has led to many new innovations and it is truly an exciting time to be working in the area of membrane protein research.

[1]   Kubicek J, Block H, Maertens B, Spriestersbach A, Labahn J. (2014) Expression and Purification of Membrane Proteins. Methods Enzymol. **541**, 117–140.

[2]   Almen MS, Nordstrom KJ V, Fredriksson R, Schioth HB. (2009) Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. Bmc Biol. **7**.

[3]   Bill RM, Henderson PJF, Iwata S, Kunji ERS, Michel H, Neutze R, et al. (2011) Overcoming barriers to membrane protein structure determination. Nat. Biotechnol. **29**, 335–340.

[4]   Moraes I, Evans G, Sanchez-Weatherby J, Newstead S, Stewart PDS. (2014) Membrane protein structure determination The next generation. Biochim. Biophys. Acta-Biomembranes. **1838**, 78–87.

[5]   Shimizu K, Cao W, Saad G, Shoji M, Terada T. (2018) Comparative analysis of membrane protein structure databases. Biochim. Biophys. Acta - Biomembr. **1860**, 1077–1091.

[6]   Overington JP, Al-Lazikani B, Hopkins AL. (2006) Opinion - How many drug targets are

there? Nat. Rev. Drug Discov. **5**, 993–996.

[7]   Deacon SE, Roach PCJ, Postis VLG, Wright GSA, Xia XB, Phillips SE V, et al. (2008) Reliable scale-up of membrane protein over-expression by bacterial auto-induction: From microwell plates to pilot scale fermentations. Mol. Membr. Biol. **25**, 588–598.

[8]   Loll PJ. (2014) Membrane proteins, detergents and crystals: what is the state of the art? Acta Crystallogr. Sect. F, Struct. Biol. Commun. **70**, 1576–83.

[9]   Orwick-Rydmark M, Arnold T, Linke D. (2016) The Use of Detergents to Purify Membrane Proteins. Current Protocols in Protein Science. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2016. p. 4.8.1-4.8.35.

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

[11]   le Maire M, Champeil P, Moller J V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents. Biochim. Biophys. Acta. **1508**, 86–111.

[12]   Linke D. (2014) Explanatory chapter: choosing the right detergent. Methods in enzymology. 2014. p. 141–148.

[13]   Postis VLG, Deacon SE, Roach PCJ, Wright GSA, Xia X, Ingram JC, et al. (2008) A high-throughput assay of membrane protein stability. Mol. Membr. Biol. **25**, 617–624.

[14]   Mus-Veteau I, Tate C. (2010) Practical Considerations of Membrane Protein Instability during Purification and Crystallisation. Heterologous Expression of Membrane Proteins. Humana Press; 2010. p. 187–203.

[15]   Inagaki S, Ghirlando R, Grisshammer R. (2013) Biophysical characterization of membrane proteins in nanodiscs. Methods. **59**, 287–300.

[16]   Lee SC, Pollock NL. (2016) Membrane proteins: is the future disc shaped? Biochem. Soc. Trans. **44**, 1011–1018.

[17]   Esmaili M, Overduin M. (2018) Membrane biology visualized in nanometer-sized discs formed by styrene maleic acid polymers. Biochim. Biophys. Acta - Biomembr. **1860**, 257–263.

[18]   Whiles JA, Deems R, Vold RR, Dennis EA. (2002) Bicelles in structure-function studies of membrane-associated proteins. Bioorg. Chem. **30**, 431–442.

[19]   Veith K, Martinez Molledo M, Almeida Hernandez Y, Josts I, Nitsche J, Löw C, et al. (2017) Lipid-like Peptides can Stabilize Integral Membrane Proteins for Biophysical and Structural Studies. ChemBioChem. **18**, 1735–1742.

[20]   Stojanoski V, Sankaran B, Prasad BVV, Poirel L, Nordmann P, Palzkill T. (2016) Structure of the catalytic domain of the colistin resistance enzyme MCR-1. BMC Biol. **14**, 81.

[21]   Yumerefendi H, Tarendeau F, Mas PJ, Hart DJ. (2010) ESPRIT: An automated, library-based method for mapping and soluble expression of protein domains from challenging targets. J. Struct. Biol. **172**, 66–74.

[22]   Beckett D, Kovaleva E, Schatz PJ. (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein Sci. **8**, 921–9.

[23]   Rawlings AE, Levdikov VM, Blagova E, Colledge VL, Mas PJ, Tunaley J, et al. (2010) Expression of soluble, active fragments of the morphogenetic protein SpoIIE from Bacillus subtilis using a library-based construct screen. Protein Eng. Des. Sel. **23,** 817-25.

[24]   Barák I, Behari J, Olmedo G, Guzmán P, Brown DP, Castro E, et al. (1996) Structure and function of the Bacillus SpoIIE protein and its localization to sites of sporulation septum assembly. Mol. Microbiol. **19**, 1047–60.

[25]   Barák I, Wilkinson AJ. (2005) Where asymmetry in gene expression originates. Mol. Microbiol. **57**, 611–620.

[26]   Levdikov VM, Blagova EV, Rawlings AE, Jameson K, Tunaley J, Hart DJ, et al. (2012) Structure of the phosphatase domain of the cell fate determinant SpoIIE from Bacillus subtilis. J. Mol. Biol. **415,** 343-58.

[27]   Rawlings AE. (2016) Membrane proteins: Always an insoluble problem? Biochem. Soc. Trans. **44**, 790-5.

[28]   Mizrachi D. (2015) Creation of Water-Soluble Integral Membrane Proteins using an Engineered Amphipathic Protein "Shield". Biophys. J. **108**, 38A.

[29]   Mizrachi D, Chen YJ, Liu JY, Peng HM, Ke AL, Pollack L, et al. (2015) Making water-soluble integral membrane proteins in vivo using an amphipathic protein fusion strategy. Nat. Commun. **6**, 6826.

[30]   Gursky O, Atkinson D. (1996) Thermal unfolding of human high-density apolipoprotein A-1: Implications for a lipid-free molten globular state. Proc. Natl. Acad. Sci. U. S. A. **93**, 2991–2995.

[31]   Hardy D, Bill RM, Jawhari A, Rothnie AJ. (2016) Overcoming bottlenecks in the membrane protein structural biology pipeline. Biochem. Soc. Trans. **44**, 838–844.

[32]   Mizrachi D, Robinson M-P, Ren G, Ke N, Berkmen M, DeLisa MP. (2017) A water-soluble DsbB variant that catalyzes disulfide-bond formation in vivo. Nat. Chem. Biol. **13**, 1022–1028.

[33] Inaba K, Ito K. (2008) Structure and mechanisms of the DsbB–DsbA disulfide bond generation machine. Biochim. Biophys. Acta - Mol. Cell Res. **1783**, 520–529.

[34] Hatahet F, Boyd D, Beckwith J. (2014) Disulfide bond formation in prokaryotes: History, diversity and design. Biochim. Biophys. Acta - Proteins Proteomics **1844**, 1402–1414.

[35] Slovic AM, Kono H, Lear JD, Saven JG, DeGrado WF. (2004) Computational design of water-soluble analogues of the potassium channel KcsA. Proc. Natl. Acad. Sci. **101**, 1828–1833.

[36] Slovic AM, Summa CM, Lear JD, DeGrado WF. (2003) Computational design of a water-soluble analog of phospholamban. Protein Sci. **12**, 337–348.

[37] Slovic AM, Stayrook SE, North B, DeGrado WF. (2005) X-ray Structure of a Water-soluble Analog of the Membrane Protein Phospholamban: Sequence Determinants Defining the Topology of Tetrameric and Pentameric Coiled Coils. J. Mol. Biol. **348**, 777–787.

[38] Fredriksson R, Lagerström MC, Lundin L-G, Schiöth HB. (2003) The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. Mol. Pharmacol. **63**, 1256–1272.

[39] Pierce KL, Premont RT, Lefkowitz RJ. (2002) Seven-transmembrane receptors. Nat. Rev. Mol. Cell Biol. **3**, 639–650.

[40] Deupi X, Kobilka B. (2007) Activation of G Protein–Coupled Receptors. Advances in protein chemistry. 2007. p. 137–166.

[41] Jazayeri A, Dias JM, Marshall FH. (2015) From G Protein-coupled Receptor Structure Resolution to Rational Drug Design. J. Biol. Chem. **290**, 19489–95.

[42] Sriram K, Insel PA. (2018) G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? Mol. Pharmacol. **93**, 251–258.

[43] Tate CG, Schertler GF. (2009) Engineering G protein-coupled receptors to facilitate their structure determination. Curr. Opin. Struct. Biol. **19**, 386–395.

[44] Grisshammer R. (2017) New approaches towards the understanding of integral membrane proteins: A structural perspective on G protein-coupled receptors. Protein Sci. **26**, 1493–1504.

[45] Errey JC, Doré AS, Zhukov A, Marshall FH, Cooke RM. (2015) Purification of Stabilized GPCRs for Structural and Biophysical Analyses. Humana Press, New York, NY; 2015. p. 1–15.

[46] Magnani F, Serrano-Vega MJ, Shibata Y, Abdul-Hussein S, Lebon G, Miller-Gallacher J, et al.

(2016) A mutagenesis and screening strategy to generate optimally thermostabilized membrane proteins for structural studies. Nat. Protoc. **11**, 1554–1571.

[47]   Serrano-Vega MJ, Magnani F, Shibata Y, Tate CG. (2008) Conformational thermostabilization of the 1-adrenergic receptor in a detergent-resistant form. Proc. Natl. Acad. Sci. **105**, 877–882.

[48]   Vaidehi N, Grisshammer R, Tate CG. (2016) How Can Mutations Thermostabilize G-Protein-Coupled Receptors? Trends Pharmacol. Sci. **37**, 37–46.

[49]   Sarkar CA, Dodevski I, Kenig M, Dudli S, Mohr A, Hermans E, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. Proc. Natl. Acad. Sci. **105**, 14808–14813.

[50]   Dodevski I, Plückthun A. (2011) Evolution of Three Human GPCRs for Higher Expression and Stability. J. Mol. Biol. **408**, 599–615.

[51]   Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, et al. (2008) Structure of a β1-adrenergic G-protein-coupled receptor. Nature. **454**, 486–491.

[52]   Doré AS, Okrasa K, Patel JC, Serrano-Vega M, Bennett K, Cooke RM, et al. (2014) Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. Nature. **511**, 557–562.

[53]   Doré AS, Robertson N, Errey JC, Ng I, Hollenstein K, Tehan B, et al. (2011) Structure of the Adenosine A2A Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. Structure. **19**, 1283–1293.

[54]   Hollenstein K, Kean J, Bortolato A, Cheng RKY, Doré AS, Jazayeri A, et al. (2013) Structure of class B GPCR corticotropin-releasing factor receptor 1. Nature. **499**, 438–443.

[55]   Srivastava A, Yano J, Hirozane Y, Kefala G, Gruswitz F, Snell G, et al. (2014) High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. Nature. **513**, 124–127.

[56]   Popov P, Peng Y, Shen L, Stevens RC, Cherezov V, Liu Z-J, et al. (2018) Computational design of thermostabilizing point mutations for G protein-coupled receptors. Elife. **7**.

[57]   Bhattacharya S, Lee S, Grisshammer R, Tate CG, Vaidehi N. (2014) Rapid Computational Prediction of Thermostabilizing Mutations for G Protein-Coupled Receptors. J. Chem. Theory Comput. **10**, 5149–5160.

[58]   Yin J, Mobarec JC, Kolb P, Rosenbaum DM. (2015) Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. Nature. **519**, 247–250.

[59]   Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, et al. (2007) GPCR Engineering Yields High-Resolution Structural Insights into 2-Adrenergic Receptor Function. Science. **318**, 1266–1273.

[60]   Chun E, Thompson AA, Liu W, Roth CB, Griffith MT, Katritch V, et al. (2012) Fusion Partner Toolchest for the Stabilization and Crystallization of G Protein-Coupled Receptors. Structure. **20**, 967–976.

[61]   Li J, Hu Z, de Lecea L. (2014) The hypocretins/orexins: integrators of multiple physiological functions. Br. J. Pharmacol. **171**, 332–50.

[62]   Schrodinger L. (2010) The PyMOL molecular graphics system, version 1.8.

[63]   Errasti-Murugarren E, Rodríguez-Banqueri A, Vázquez-Ibar JL. (2017) Split GFP Complementation as Reporter of Membrane Protein Expression and Stability in E. coli: A Tool to Engineer Stability in a LAT Transporter. Humana Press, New York, NY; 2017. p. 181–195.

[64]   Drew D, Lerch M, Kunji E, Slotboom D-J, de Gier J-W. (2006) Optimization of membrane protein overexpression and purification using GFP fusions. Nat. Methods. **3**, 303–313.

[65]   Hattori M, Hibbs RE, Gouaux E. (2012) A Fluorescence-Detection Size-Exclusion Chromatography-Based Thermostability Assay for Membrane Protein Precrystallization Screening. Structure. **20**, 1293–1299.

[66]   Kawate T, Gouaux E. (2006) Fluorescence-Detection Size-Exclusion Chromatography for Precrystallization Screening of Integral Membrane Proteins. Structure. **14**, 673–681.

[67]   Rodríguez-Banqueri A, Kowalczyk L, Palacín M, Vázquez-Ibar JL. (2012) Assessment of membrane protein expression and stability using a split green fluorescent protein reporter. Anal. Biochem. **423**, 7–14.

[68]   Blakemore RP. (1975) Magnetotactic bacteria. Science. **190**, 377–379.

[69]   Bazylinski DA, Frankel RB. (2004) Magnetosome formation in prokaryotes. Nat Rev Microbiol. **2**, 217–230.

[70]   Faivre D, Schuler D. (2008) Magnetotactic bacteria and magnetosomes. Chem Rev. **108**, 4875–4898.

[71]   Grünberg K, Müller E-C, Otto A, Reszka R, Linder D, Kube M, et al. (2004) Biochemical and proteomic analysis of the magnetosome membrane in Magnetospirillum gryphiswaldense. Appl Env. Microbiol. **70**, 1040–1050.

[72]   Komeili A, Komelli A. (2007) Molecular mechanisms of magnetosome formation. Annu Rev

Biochem. **76**, 351–366.

[73]   Nudelman H, Valverde-Tercedor C, Kolusheva S, Gonzalez TP, Widdrat M, Grimberg N, et
        al. (2016) Structure-function studies of the magnetite-biomineralizing magnetosome-
        associated protein MamC. J. Struct. Biol. **194**, 244–252.