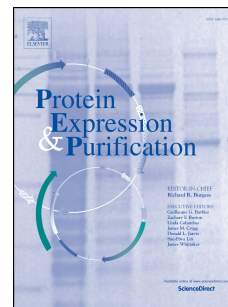


# Journal Pre-proof

Expression and purification of recombinant G protein-coupled receptors: A review

Daniel N. Wiseman, Abigail Otchere, Jaimin H. Patel, Romez Uddin, Naomi L. Pollock, Sarah J. Routledge, Alice J. Rothnie, Cathy Slack, David R. Poyner, Roslyn M. Bill, Alan D. Goddard



PII: S1046-5928(19)30493-0

DOI: <https://doi.org/10.1016/j.pep.2019.105524>

Reference: YPREP 105524

To appear in: *Protein Expression and Purification*

Received Date: 9 September 2019

Revised Date: 28 October 2019

Accepted Date: 29 October 2019

Please cite this article as: D.N. Wiseman, A. Otchere, J.H. Patel, R. Uddin, N.L. Pollock, S.J. Routledge, A.J. Rothnie, C. Slack, D.R. Poyner, R.M. Bill, A.D. Goddard, Expression and purification of recombinant G protein-coupled receptors: A review, *Protein Expression and Purification* (2019), doi: <https://doi.org/10.1016/j.pep.2019.105524>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier Inc.

**Expression and Purification of Recombinant G Protein-Coupled Receptors: A Review.**

Daniel N. Wiseman<sup>a</sup>, Abigail Otchere<sup>a</sup>, Jaimin H. Patel<sup>a</sup>, Romez Uddin<sup>a</sup>, Naomi L. Pollock<sup>b</sup>, Sarah J. Routledge<sup>a</sup>, Alice J. Rothnie<sup>a</sup>, Cathy Slack<sup>a</sup>, David R. Poyner<sup>a</sup>, Roslyn M. Bill<sup>a</sup> & Alan D. Goddard<sup>a\*</sup>

<sup>a</sup> School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK. [d.wiseman1@aston.ac.uk](mailto:d.wiseman1@aston.ac.uk), [otcherea@aston.ac.uk](mailto:otcherea@aston.ac.uk), [patelj47@aston.ac.uk](mailto:patelj47@aston.ac.uk), [uddinr3@aston.ac.uk](mailto:uddinr3@aston.ac.uk), [s.routledge@aston.ac.uk](mailto:s.routledge@aston.ac.uk), [a.rothnie@aston.ac.uk](mailto:a.rothnie@aston.ac.uk), [c.slack@aston.ac.uk](mailto:c.slack@aston.ac.uk), [d.r.poyner@aston.ac.uk](mailto:d.r.poyner@aston.ac.uk), [r.m.bill@aston.ac.uk](mailto:r.m.bill@aston.ac.uk), \*Corresponding author. [a.goddard@aston.ac.uk](mailto:a.goddard@aston.ac.uk)

<sup>b</sup> University of Birmingham, Birmingham, B15 2TT, UK. [n.pollock@bham.ac.uk](mailto:n.pollock@bham.ac.uk)

**Abstract**

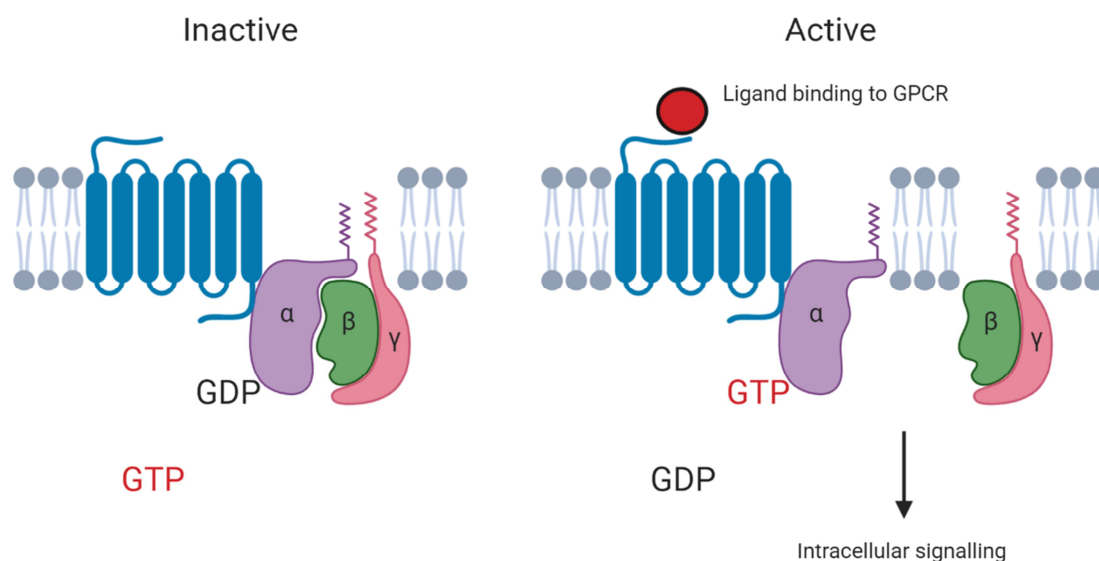
Given their extensive role in cell signalling, GPCRs are significant drug targets; despite this, many of these receptors have limited or no available prophylaxis. Novel drug design and discovery significantly rely on structure determination, of which GPCRs are typically elusive. Progress has been made thus far to produce sufficient quantity and quality of protein for downstream analysis. As such, this review highlights the systems available for recombinant GPCR expression, with consideration of their advantages and disadvantages, as well as examples of receptors successfully expressed in these systems. Additionally, an overview is given on the use of detergents and the styrene maleic acid (SMA) co-polymer for membrane solubilisation, as well as purification techniques.

**Keywords**

Expression; Purification; GPCR; SMALP; Review.

**Introduction**

As the largest family of membrane proteins in the human genome, G protein-coupled receptors (GPCRs) are widely studied due to their involvement in normotypical and pathological cell signalling profiles [1]. Characteristically, as shown in Figure 1, these seven-transmembrane receptors undergo a conformational change upon activation by a ligand, allowing propagation of signalling cascades within the cell [2].



**Figure 1: Ligand induced activation of a G protein-coupled receptor (GPCR).** GPCRs (blue) are transmembrane receptors which activate intracellular signalling pathways, through coupling to G proteins. These heterotrimeric proteins consist of three subunits denoted  $\alpha$ ,  $\beta$  and  $\gamma$ , and are classically activated by a ligand induced conformational change in the GPCR [2]. This movement is proposed to involve a rotational transmembrane helix reorientation, exposing an intracellular binding cleft [3]. GDP is exchanged for GTP on  $G\alpha$ , while the  $\beta\gamma$  complex splits away and is able to signal independently of the  $G\alpha$  subunit. Humans encode 18, 5 and 12 different  $\alpha$ ,  $\beta$  and  $\gamma$  subunits, respectively. These combine into a variety of stimulatory ( $G_s$ ) or inhibitory ( $G_{i/q}$ ) effects on pathways including those dependent on adenylyl cyclase and phospholipase C. Created with Biorender.com.

Understanding the relationship between a GPCR's structure and function will aid further development of ortho- and allosteric molecules against these receptors to affect their pharmacology. While approximately half of all drugs target GPCRs, this is only reflected in a 5% coverage of these receptors, providing significant scope for further structure-based novel drug discovery [4].

While the structure of some GPCRs have been successfully determined, many challenges remain in this field. These include the concepts of homo-dimerization, heteromeric protein-protein interactions and the structural complexity of important motifs [5] – of note, the folding and flexibility of the ligand-binding domain of family B GPCRs [6]. While computational biology has greatly enhanced the versatility of studying GPCR structures [7], classical techniques are also often employed including x-ray crystallography [8], cryo-electron microscopy [9] and nuclear magnetic resonance spectroscopy [10]. A significant drawback to these methods lies in the initial requirement for a high yield and purity of mature, folded target protein.

These challenges have propelled the development and implementation of recombinant membrane protein expression, solubilisation and purification systems over the last two decades - contributing, in no small part, to the increase in resolved structures of membrane proteins in the same time-frame [11]. This review will provide a current summary of the methodology, benefits and drawbacks of the expression systems available to GPCR researchers (Table 1), as well as an overview of applicable solubilisation and purification techniques.

Expression System	Ease of genetic manipulation	Ease of culture	Expense	Scalability	Post-translational modification	Membrane composition	Active GPCR expression
<i>Drosophila melanogaster</i>	Amber	Green	Green	Green	Amber	Amber	Green
<i>E. coli</i>	Green	Green	Green	Green	Red	Amber	Red
Insect cells	Red	Amber	Amber	Amber	Amber	Amber	Green
Mammalian cell-line	Red	Red	Red	Red	Green	Green	Green
Yeast	Green	Green	Green	Green	Amber	Amber	Amber

**Table 1: A comparison of expression systems for GPCRs.** A qualitative assessment of considerations linked to recombinant protein expression systems. Green = positive, amber = moderate, red = negative.

## Expression Systems

### *E. coli*

Native expression of GPCRs is well-known to be restricted to eukaryotic organisms [12] yet *E. coli* has proved an attractive host for expression and purification of a subset of receptors [13]. *E. coli* has become a laboratory workhorse for a number of reasons. Firstly, decades of work have led to well-characterised, rapidly growing cells which are easy to culture and strains have been optimised for protein expression, including membrane protein-specific strains [14]. *E. coli* can be easily grown across a range of scales allowing fermentation to produce large quantities of protein, although it should be noted that there is not always a linear relationship between culture volume and product yield [14]. The genetic tractability of *E. coli* allows a variety of expression plasmids to be used to tune protein expression levels. This can be particularly important with membrane proteins when saturation of the translocon can be a rate-limiting step [13, 15]. Indeed, high level expression can lead to formation of inclusion body and refolding of GPCRs from such environments has met with limited success [16]. Some of these problems can be overcome with judicious strain selection and expression at lower temperatures [17]. However, due to its prokaryotic nature, *E. coli* does not possess a number of features that can be essential for GPCR function. There is a lack of post-translational modification, including glycosylation, which can be essential for ligand binding [18]. Additionally, the lipid membrane environment may not include essential components such as cholesterol [19] and contains a very different lipidome to eukaryotic cells – there is clear evidence for lipid-dependent GPCR activity [20].

Despite these clear limitations [21], there have been a number of reports of GPCRs being successfully expressed in *E. coli* [22-37]. Unmodified GPCRs tend to have low stability and may aggregate in such systems [36]. A key strategy for successful expression and correct folding of GPCRs in *E. coli* is the use of fusion partners [38]. These serve both to direct the correct insertion of the receptor into the membrane whilst also increasing its overall solubility, thereby aiding both expression and purification [29]. Additionally, strategies such as selective mutagenesis to introduce stabilising mutations and the use of insertions or truncations has proven successful in some cases [17, 36]. Indeed, the genetic tractability of *E. coli* can be used to select for variants with increased stability and expression even for relatively intractable receptors [39-41].

There are a number of advantages of the use of *E. coli* for downstream applications. It is relatively easy to conduct isotopic labelling experiments such that the subsequent protein can be used for NMR studies [10]. It should, however, be noted that the relatively low

expression levels of GPCRs in *E. coli* is further impacted by such labelling strategies [10]. However, through optimised expression it has been reported that GPCR expression of up to 50 mg/L can be achieved [38]. The genetic amenability and tools available for *E. coli* open possibilities to select GPCR variants with enhanced expression and stability, generate those “locked” in a particular conformation, and also to, potentially, engineer those with completely novel functions [42]. Despite its prokaryotic nature, *E. coli* has clear potential for at least a subset of GPCRs.

## Yeast

The fission yeast *S. pombe* and baker's yeast *S. cerevisiae* are important tools to express and investigate the signaling and stability of GPCRs [43-49], however, the methylotrophic yeast, *Pichia pastoris* (reclassified as *Komagataella phaffii*), is favored for the overexpression of GPCRs for structural studies [50]. High yields of functional receptors have been expressed [21], including the adenosine 2a receptor [51-53], 5HT5A receptor, beta2-adrenergic receptor [54] and muscarinic acetylcholine receptor M2 subtype (CHRM2) [55]. In addition, high-resolution crystal structures of the histamine H<sub>1</sub> receptor [56] and the adenosine 2a receptor in complex with an antibody Fab fragment [57] have been obtained using the *P. pastoris* expression system as well as other membrane proteins.

This has been feasible due to the ease of manipulation and stable integration of expression vectors into *P. pastoris* coupled with its ability to grow to high cell densities on glycerol and to utilize methanol as the sole carbon source [58]. This system allows high levels of protein expression to be induced under the tightly controlled AOX1 promoter [59, 60]. Other promoters are also available, including the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter as well as emerging novel methanol-inducible, non-methanol inducible and constitutive promoters [61, 62]. Several expression vectors and strains are commercially available to optimize protein expression. Commonly used vectors are the pPIC9K, pPICZ or pPICZalpha, the latter of which contains an  $\alpha$ -MF signal sequence derived from *S. cerevisiae* to enhance protein secretion. Expression vectors generally contain geneticin, kanamycin or zeocin resistance genes, and auxotrophic markers have also been used for GPCR expression [50, 58]. Strains frequently used include the wild type X-33 strain, protease deficient strains (SMD1163, SMD1168 and SMD1165) and auxotrophic strains GS115 and KM71 [11, 50, 58].

*P. pastoris* is able to perform post translational modifications such as disulphide bond formation, and N and O-linked glycosylation. N-linked glycosylation occurs at the Asn-X-Ser/Thr motif on extracellular domains of GPCRs [63], and in some cases is required for cell surface expression [64], ligand binding and cell signalling [65]. While early steps of *P. pastoris* N-linked glycosylation are similar to the process in mammalian cells [66], it may potentially hypermannosylate the protein which can lead to misfolding, although this is less extensive than in *S. cerevisiae* [67]. It may also glycosylate residues of protein where this would not naturally occur [67, 68]. Some GPCRs expressed in *P. pastoris* have therefore been engineered with these sites removed to facilitate crystallization [56, 57]. A study by Yurugi-Kobayashi *et al* 2009 [69] which analysed glycosylation-deficient GPCRs demonstrated that while some receptors were expressed with lower functional levels, others were expressed at levels suitable for structural studies when this approach was combined with culture optimisation. *P. pastoris* strains and vectors with humanized N-glycosylation have been developed [66, 70], but have not as of yet been applied to GPCR expression. Other modifications to GPCRs expressed by *P. pastoris* include codon optimisation, N and C-terminal truncations [52, 57] and T4-lysozyme fusion to intracellular loop 3 (ICL3) [56].

In contrast to mammalian cells, yeast membranes contain ergosterol rather than cholesterol. Membrane cholesterol is thought to be required for the correct function of some GPCRs, and may cause direct conformational changes or indirectly alter membrane properties. Crystal structures suggest that some GPCRs contain specific cholesterol binding sites [71-73]. A humanized *P. pastoris* strain has been engineered which synthesizes cholesterol [74], and could be of benefit for GPCR expression. Cholesteryl hemi-succinate, a cholesterol derivative, can also be added to maintain stability [53, 56, 75].

In summary, yeast possess several advantages over other expression systems, including the ability to perform eukaryotic post-translational modifications while being capable of rapid growth to high cell densities on a large scale in relatively cheap media [76]; these aspects make yeast an appealing host, and their use for GPCR expression has significantly improved knowledge of GPCR structure and function.

### **Insect Cell-Line (Sf9, Sf21, Hi5)**

In GPCR structural studies, insect cells are the most commonly used expression system to achieve milligram quantities of protein [75, 77]. Expression is achieved via infection with a recombinant form of *Autographica californica*; a multiple nuclear polyhedrosis virus. This baculovirus infects the cells and drives the production of the protein of interest, usually via the polyhedrin promoter [78]. The majority of GPCR studies use the *Spodoptera frugiperda* (Sf9) or Sf21 cell-line, in preference to *Trichoplusia ni* (High Five) cells [77]. However protein expression levels can vary between different cell-lines so screening of cell lines is necessary when using insect cells for structural studies. The advantages of insect cell expression include growth in serum-free shaker cultures, which decreases costs and enables relatively easy scale-up, as well as high yield and the ability to perform most post-translational modifications [78].

A range of different systems have been developed to generate the recombinant baculovirus [78]. Once generated the virus requires titration in order to achieve an appropriate multiplicity of infection (MOI); an excess of virus will kill the cells before they can be harvested [78]. This step is often the most difficult as how to directly quantify virus is unclear and often inaccurate. Viral plaque assays are commonly used but take a minimum of two weeks and, often, their accuracy is questionable [78]. Various alternative approaches have been devised, including flow cytometry and qPCR, but cost, time and variability can still be problematic [79]. Many researchers find that using the virus directly to express protein in Sf9 cells and quantifying the expression level is quicker and more accurate. However, once the virus has been generated and titrated it can be stored for a number of years at 4°C and much longer at -80°C. More virus can also be generated by infecting the cells with a high MOI and collecting the cell culture media five days post-infection [80]. Initially, it can take up to a month to produce enough baculovirus to drive large scale expression however, once the virus has been generated, protein can be expressed within a week.

One potential disadvantage of insect cell expression arises from differences in lipid composition compared to mammalian cells. Insect cell membranes are low in cholesterol, have very high phosphatidylinositol content and no phosphatidylserine [75], and as noted below, protein function is highly dependent on lipid environment [81]. There have also been reports that a proportion of the protein produced can be misfolded [82], or that the lytic pathway of viral infection can cause protein degradation [80]. However, overall it remains one of the key approaches for GPCR overexpression.

### **Mammalian Cell-Line (HEK293, COS)**

Membrane proteins are often expressed in insect, bacterial or yeast expression systems due to their high protein yield and expression, which is advantageous for structural studies [83]. High resolution human membrane protein structures have been solved from recombinant proteins derived from these sources, but the protein conformation and modifications may differ from a human protein expressed in human cells. To address these issues, mammalian cell-lines capable of expressing a desired protein have been trialled [84]. The selection of a specific cell-line is determined by whether the expression system represents the near-native environment in which the desired protein is endogenously expressed [83]. GPCRs are heavily post-translationally modified, therefore expressing human GPCRs in mammalian cells is often ideal to characterise their function and pharmacology [75]. For structural biology however, post-translational modifications can be detrimental during crystal formation. Glycosylation sites affect formation of ordered crystals due to the flexibility and heterogeneity of glycan residues [75]. This issue can be solved by mutating the N-glycosylation sites, provided that the conformation of the receptor is still stable [75].

The native environment in which GPCRs are expressed has an impact on the conformation and pharmacology of these receptors. The phospholipid composition of the native lipid bilayer has an allosteric effect on GPCRs. In the case of the human beta 2 adrenergic receptor in different liposomes, synthetic phosphatidylglycerol stabilised the active conformation of the receptor [81]. Due to variations in lipid composition between cell-lines of different expression systems, choosing a human cell-line that has a similar lipid composition to that of the endogenous GPCR is important for receptor pharmacology. As cholesterol can allosterically modulate GPCRs [85], its replacement with ergosterol in yeast expression systems can be detrimental. In the case of the human  $\mu$ -opioid GPCR, ergosterol constrains the receptor in an inactive state, whereas cholesterol stabilises the active state [86]. Post-translational modification is important for GPCR function, where mammalian cells have the correct enzymes for phosphorylation and palmitoylation of human GPCRs [75].

Yields of recombinant protein in non-mammalian expression systems are often higher than mammalian expression systems, therefore optimising the GPCR gene construct for expression is an important first step [75]. GPCR gene constructs are often codon optimized for mammalian cell expression [87]. Kozak sequences (GCCACCATGG) and signal peptide sequences can be fused to the 5' end of the GPCR construct to enhance protein expression and cell surface delivery [87, 88]. The optimized construct can be ligated into a plasmid vector, which can be transfected into mammalian cells transiently, or be used to create stable cell lines [75].

Overall, immortalised mammalian cell-lines are useful to study human GPCRs in their wild-type or mutated form. The immortalised human embryonic kidney 293 (HEK293) can transiently express recombinant proteins and is amongst the most popular human cell-line to use [89, 90]. HEK293 cells exist as adherent cells or suspension cells; the latter are grown at a higher density, which is useful for protein production [89]. Table 2 summarises the currently known 3D structures of GPCRs derived from expression in mammalian cell-lines, as of 30/8/19.

GPCR	Organism	Cell-Line	Resolution, Å	PDB Entry
Angiotensin type II receptor	<i>H. sapiens</i>	Expi293F	2.90	6DO1
CB <sub>1</sub> cannabinoid receptor	<i>H. sapiens</i>	HEK293F	2.80	5TGZ
Cytomegalovirus US28	<i>H. sapiens</i>	HEK	2.89	4XT1
Leukotriene B <sub>4</sub> receptor	<i>C. porcellus</i>	HEK293	3.70	5X33
Rhodopsin	<i>B. taurus</i>	Cos	3.40	2J4Y
		HEK293S-GnTI <sup>-</sup>	3.30	4A4M
		HEK293S	2.36	6FK6
		HEK293	4.38	6QNO
	<i>H. sapiens</i>	HEK293S	3.30	4ZWJ
	<i>H. adansoni</i>	HEK293	2.14	6I9K
Smoothened receptor	<i>H. sapiens</i>	HEK293S-GnTI <sup>-</sup>	3.20	5L7D
		HEK293S	3.84	6OT0
	<i>M. musculus</i>	HEK293	2.80	6O3C

**Table 2: 3D structures of recombinant GPCRs derived from mammalian cell-line expression.** Database query generated with MemProtMD at <https://blanco.biomol.uci.edu/mpstruc/>. Accessed 5/9/19.

### *Drosophila melanogaster*

Each of the conventional expression systems detailed above are not without their limitations. One of the major drawbacks associated with all of these systems is the build-up of immature proteins in the intracellular membranes caused by the cell's failure to properly fold and transport the mature GPCR to the cell surface. This issue can lead to inadequate yields for structural studies thus limiting our understanding of GPCR structure and function [91, 92] and obtaining adequate yields of the mature GPCR often requires optimisation of the expression conditions, increasing cost.

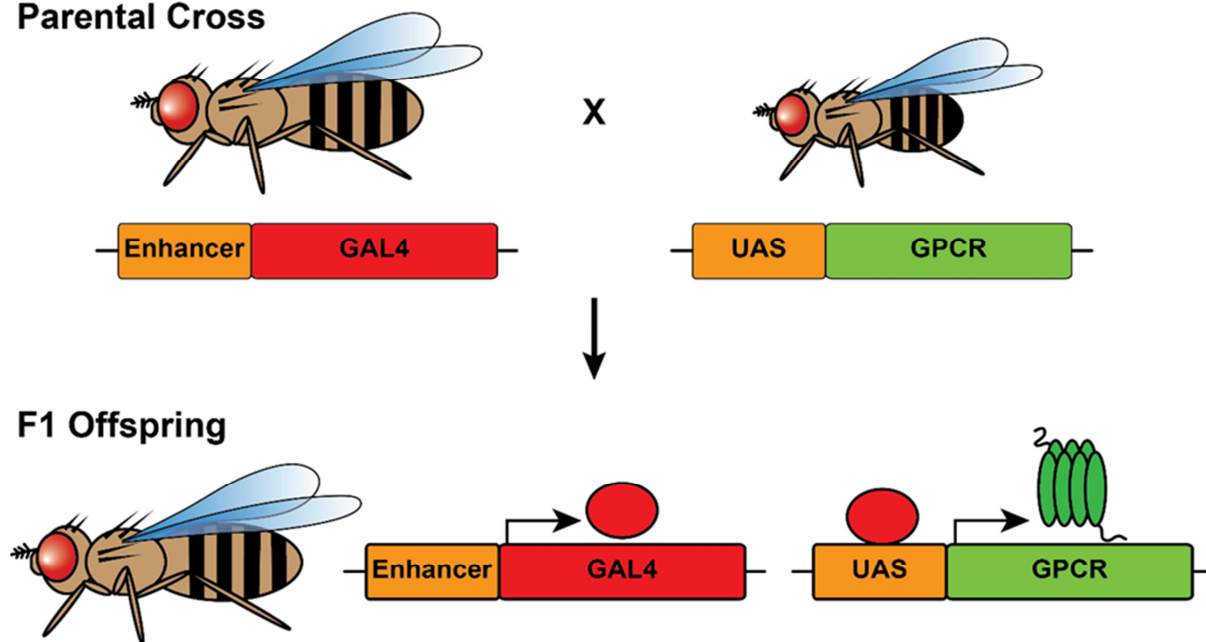
The fruit-fly, *Drosophila melanogaster*, has recently been utilised as an attractive alternative expression system to overcome some of these problems. The system takes advantage of the unique properties and architecture of the fly eye which consists of photoreceptor cells (PRCs) containing membrane stacks called rhabdomeres [93] therefore providing a large surface area for expression and folding of large amounts of membrane-associated proteins [94].

Heterologous expression of proteins within the PRCs is achieved using the well-established GAL4-UAS system [95]. This system allows the tissue-specific expression of transgenes by exploiting the use of the yeast GAL4 protein, a transcription factor that specifically binds to an Upstream Activating Sequence (UAS) to drive expression of its target genes. To express a specific protein in a particular tissue type within the fly, two strains are mated together: the driver-strain which expresses GAL4 from a tissue-specific promoter and the UAS strain which contains the transgene of interest cloned downstream of the GAL4 UAS (Figure 2). In the resulting offspring, the transgene will be expressed in those specific cells that contain GAL4 [95]. By using a driver-strain that specifically expresses GAL4 within the fly eye, heterologous GPCR expression can be restricted to the PRCs. Generating transgenic flies is



relatively easy and comparable in cost to other conventional expression systems. Also, *Drosophila* culture media for rearing experimental animals is relatively inexpensive and the need to work in sterile conditions is eliminated when working with flies [94].

### Parental Cross



**Figure 2. The *Drosophila* GAL4-UAS system for targeted gene expression.** The GAL4-UAS system can be used for targeted expression of GPCRs within the *Drosophila* photoreceptor cells (PRCs). To obtain flies expressing a gene of interest in a tissue specific pattern, two *Drosophila* strains are mated together in the parental cross. The driver strain expresses the yeast GAL4 protein from a tissue-specific enhancer/promoter. The UAS strain contains the gene of interest cloned downstream of the GAL4 Upstream Activating Sequence (UAS). The resulting F1 offspring will express GAL4 protein in a tissue-restricted pattern which will bind to the UAS sequences upstream of the gene of interest to drive its expression in those specific cells. By using a driver strain that expresses GAL4 specifically within the fly eye, heterologous GPCR expression can be restricted to the PRCs.

*Drosophila* PRCs have been successfully used to express a number of GPCRs [96]; the *Drosophila* metabotropic glutamate receptor, DmGluRA being first reported. Overexpression of DmGluRA in PRCs resulted in higher yields of mature receptor than obtained using other conventional methods, including insect cell culture. Moreover, toxicity effects of DmGluRA overexpression were not observed in the host cells, overcoming a major limitation of other expression systems. Expressing mammalian mGluRs in the fly eye produced similar yields as expressing DmGluRA, suggesting that this system can be used to express foreign GPCRs from other species including human, rat and *Chlamydomonas* [94]. Furthermore, expression of the other two classes of GPCRs have now also been successfully reported using this system [96]. Importantly, it should also be noted that scale-up of expression, as required for downstream processes such as crystallisation, can be readily achieved and can sometimes be an issue for other cell culture- based expression systems [76].

Although the use of *Drosophila* for GPCR expression overcomes several of the major drawbacks associated with more conventional expression systems, it is not without its own limitations. Firstly, this method requires access to fly genetics expertise and facilities for *Drosophila* culture that if unavailable will require the need for collaboration with specialised laboratories that can provide these services [94]. Additionally, although this system is capable of post-translationally modifying proteins there are differences in some of the modifications that occur in the fly that may be important for GPCR function. An example of this is *N*-glycosylation, which tends to be less complex in insects and lacking in extended

antennae compared to mammals [97, 98]. Furthermore, regarding purification of membrane-associated proteins expressed in the fly eye, there is currently a lack of reports describing the use of detergent-free purification methods such as SMALPs using this system [94, 96]. This will be an important future development due to the problems of membrane protein stabilisation associated with using detergents [99, 100]. Yet despite these apparent shortfalls, *Drosophila* could prove to be a cheaper and more efficient alternative for functional GPCR expression and purification.

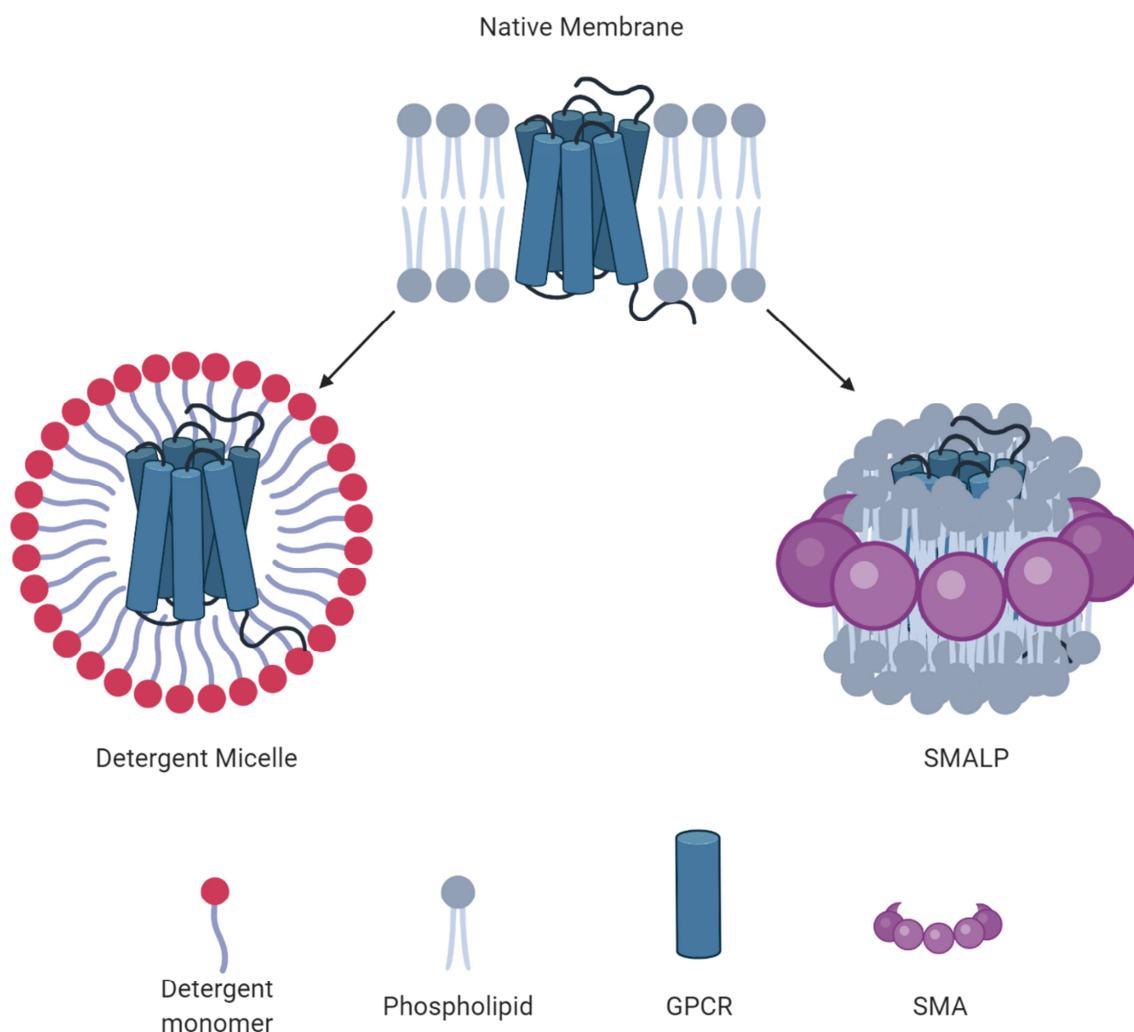
## **Solubilisation and Purification**

### **Detergents**

An important barrier to studying GPCRs is the need to solubilise and purify these membrane proteins away from their native bilayer [101]. Ideally, this process should simultaneously retain target proteins in their folded, functional conformations for further *in vitro* study. Surfactant detergents are able to solubilise and extract membrane proteins due to their amphiphilic nature, improving the aqueous solubility of the protein [102]. A plethora of detergents are commercially available with different physicochemical properties; often, a screen is best performed to identify optimal detergents, likely on a protein-by-protein basis [103].

Briefly, detergents fall into three classes based on their polar head group – ionic, zwitterionic and non-ionic. Ionic detergents such as SDS are regarded as harsh, zwitterions are milder (LDAO) while non-ionic detergents are considered mild. The described harshness is derived from the efficacy of disrupting intra- and inter-molecular interactions. While some factors can be scrutinised, such as the critical micelle concentration (CMC) and hydrophilic-lipophilic balance (HLB), some detergents clearly perform well [104]. Overall, the non-ionic alkyl maltopyranoside detergents DM and DDM have been most successful in contributing to the resolution of membrane protein structures (approximately 45%) [105], and can be regarded as an evidence-based starting point [11, 53, 56, 57, 106].

Detergent monomers will, above the CMC, associate with a biological membrane and undergo a transbilayer mechanism to flip from the outer to inner leaflet [107]. This leads to the formation of lipid-detergent micelles (Figure 3), the efficacy of which depends on the HLB and size/polarity of the detergent molecule. At this point, the bilayer falls apart resulting in the solubilisation of the membrane. A GPCR's protrusion from the bilayer can aid in the incorporation of detergent monomers due to disruptions in the ordered lipid arrangement. However, this can be opposed by the notion of detergent-resistant membranes [108] – especially in regards to GPCR populations in cholesterol-rich lipid rafts [109].



**Figure 3: Solubilised GPCR in a detergent micelle compared to a SMALP.** These diagrams show the interactions of the phospholipid bilayer (grey) or detergent monomers (red) with a GPCR (blue). Importantly, the difference between a detergent micelle and SMALP is shown, with regards to the retention of the GPCR's annular lipids when surrounded by the SMA co-polymer (purple). Created with Biorender.com.

Following solubilisation, it is often necessary to reduce or remove excess detergent to enable purification and further structural or functional analysis. Several methods are sufficient to achieve this including dialysis, size exclusion and affinity chromatography [107].

Despite their utility thus far, detergents are not without their limitations. Understanding of GPCR structure/function has led to the acknowledgement of the native lipid environment. Not only do the lipids surrounding a receptor provide lateral pressure, directly bound lipids can also be essential to influence active/inactive conformations [53]. As such, detergent micelles do not exert the same lateral pressure, and in some cases, the directly bound lipids required for stability/function are removed. It is therefore preferential to adopt techniques which retain these important components to represent a more biologically realistic reflection of GPCRs and their surroundings; and to prevent destabilisation or inactivation during solubilisation [110].

While several new detergents are being designed/developed, other advances reviewed elsewhere include amphipols [111], nano discs [112] and co-polymers including DIBMA [113]

and SMA [114]. Largely, detergents must be chosen empirically, which can be an expensive and protein-demanding approach.

### SMALPs

GPCRs are one of the classes of protein that have most frequently defied the attempts by biochemists to purify and characterise them. Poor thermal stability is often blamed for this, and considerable effort and resources have been expended to generate thermostabilized versions of GPCRs, in particular for structural studies [115]. An alternative outlook is that GPCRs are destabilised by detergents. Replacing these with better membrane mimetics could prevent destabilisation of the proteins. Amongst the alternatives that have been proposed to meet this need are styrene maleic acid lipid particles (SMALPs). Styrene maleic acid (SMA) is an amphipathic co-polymer that, when added to lipids, spontaneously assembles into nanoparticles of ~10 nm diameter [116, 117]. These nanoparticles consist of SMA polymer surrounding a patch of lipid bilayer (Figure 3). When SMA is added to biological membranes a similar self-assembly process forms polymer-bound lipid particles containing membrane proteins [118-122].

One of the first reported successes using the SMALP method was the purification and functional characterisation of the adenosine-2A receptor ( $A_{2A}R$ ) which was overexpressed in both human epithelial kidney (HEK) cells and *Pichia pastoris* [53]. This demonstrated that the protein could be rapidly and effectively purified in a detergent-free manner.

The ligand-binding properties of the  $A_{2A}R$ -SMALPs were used to assay its stability under a variety of conditions. Notably,  $A_{2A}R$ -SMALPs withstood more than 5 freeze-thaw cycles without reduction in their ligand binding ability. Likewise,  $A_{2A}R$ -SMALPs and  $A_{2A}R$  in the membrane both retained 75 % of their specific ligand binding capacity after up to 15 days of incubation at 4 °C. In detergent, this binding declined to 0 % by day 3. Similarly at 37 °C the stability of  $A_{2A}R$  in SMALPs far outstripped that of the detergent-solubilised sample. This remarkable stability under a range of conditions makes  $A_{2A}R$ -SMALPs a much more flexible and useful reagent than the detergent-solubilised equivalent.

$A_{2A}R$  was also used in a study of an alternative styrene co-polymer: styrene maleimide (SMI). This has some similar properties and architecture to SMA, but is acid-compatible and can be used in buffers with pH < 7.8 [123]. By contrast SMA is soluble only above pH 5.8, and is more usually used in buffers of pH 8.  $A_{2A}R$ -SMILPs had ligand-binding properties equivalent to the protein in the cell membrane, indicating that SMI also has potential as a reagent for the detergent-free purification of GPCRs. In the same study, the vasopressin receptor ( $V_{1a}R$ ) also retained its specific ligand-binding properties in SMILPs.

The ability to bind ligands is not the only indicator of GPCR function. Perhaps of more importance is the ability of a purified GPCR to recruit/signal to G-proteins and initiate intracellular signalling cascades [124]. One study using SMA has demonstrated that the melatonin receptor (MT1R) and the ghrelin receptor (GHS-R1a) in SMALPs are capable of G-protein activation, arrestin recruitment and ligand binding [125].

Protein name	GPCR class	Downstream analysis	Source membrane	Polymer used for solubilisation	Reference
A <sub>2a</sub> R	A	Radioligand binding assays, thermal stability	Human endothelial kidney (HEK 293T) cells; <i>Pichia pastoris</i>	2:1 SMA (SMA2000P)	Jamshad <i>et al.</i> [53]
CTR	B	Radioligand binding	Cos cells		Charlton thesis [127]
GHS-1aR	A	Ligand binding via FRET; arrestin recruitment; GTP binding assays	Asolectin proteoliposomes	2:1 SMA (SMA2000P)	Logez <i>et al.</i> [125]
HWbR	A	Crystallisation	<i>Escherichia coli</i>	3:1 SMA (SMA3000P)	Broecker <i>et al.</i> [126]
MT1R	A	Ligand binding; G protein activation; arrestin recruitment	<i>P. pastoris</i>	2:1 SMA (SMA2000P)	Logez <i>et al.</i> [125]
V1aR	A	Radioligand binding	HEK 293T	2:1 SMI	Hall <i>et al.</i> [123]; Charlton thesis [127]

**Table 3: Summary of published methods for the solubilisation of GPCRs by styrene maleic acid and related polymers.**

To date, there is only one high resolution structure of a GPCR in SMALPs, *Haloquadratum walsbyi* bacteriorhodopsin (HWbR). This structure was solved at 2.0 Å resolution using the *in meso* crystallisation (lipidic cubic phase) methodology [126]. Hence, it is likely that the protein-SMALPs integrated into the bilayers of the cubic phase prior to crystallogenesis. Lipids are visible in the structure, but these are identifiable as monooleins, the lipids used to assemble the cubic phase. In a parallel experiment, bR purified using detergent had a remarkably similar structure to the structure derived from SMA-solubilised bR. Therefore in this case there is an argument that using SMA did not provide additional structural information compared to using detergent. By contrast, a recent structure was solved by cryo-electron microscopy at 3.4 Å resolution of a bacterial respiratory supercomplex purified using SMA [127]. This structure did show specific native lipids bound to the protein, which may be of relevance in understanding the subtleties of its structure and function. This hints that cryo-EM may be a viable approach for solving structures of membrane proteins retaining their native lipids.

Following solubilisation, several purification methods may be employed which have been reviewed elsewhere [128]. Summarised in Table 4, these include gel filtration, ion exchange and affinity chromatography, of which the latter is most popular. While affinity to antibodies or ligands such as lectin can be utilised for membrane proteins, GPCRs expressed in the systems discussed generally utilize a poly-histidine tag [53, 56], FLAG tag [54, 129] or HA tag [130] for detection and affinity chromatography purification, and it is possible to obtain the quantities of functional material required for structural studies [56, 57].

Purification Technique	Pros	Cons
Affinity chromatography [131, 132]	Can be used if protein molecular weight, charge or hydrophobicity is unknown. High affinity binding can result in high sample purity.	May require a tag or terminal fusion. Washing may remove weakly bound molecules. SMALPs are sensitive to divalent cations.
Gel filtration [133, 134]	Efficient separation of large and small molecules. Minimal elution volume. No sample loss.	Only separated on size. May require further techniques. Limited resolution due to short chromatogram timescale.
Ion exchange [131]	Only one charge-based interaction. Predictable elution pattern.	Inconsistency between columns. Limited to ionizable groups.

**Table 4: Pros and cons of purification techniques available for GPCRs.**

### Conclusion

GPCRs remain a challenging component of the membrane protein structural biology field. While the sources of difficulty are gradually being lessened as understanding and technology advance, the dearth of structural information is limiting novel drug design and discovery [4]. Computational biology has greatly enhanced the ability to predict and manipulate GPCR structure, and how this affects their functions. However, *in silico* experiments remain only a component of the holistic study of membrane proteins; expression and purification are largely required for downstream biochemical and biophysical analysis [8, 9, 135, 136].

As such, and discussed in this review, the expression systems available to GPCR researchers each come with their own benefits and drawbacks. While ease of culture and genetic amenability are undoubtedly attractive qualities, they clearly do not entirely make up for biologically important characteristics such as post-translational modifications. There will always seemingly remain a compromise with the expression system of choice, if only the expense. Regardless of these drawbacks, each traditional system will be preferred for application to certain techniques. For example, post-translational modification may be less desired with regards to crystallisation, but more so for trafficking and functionality.

GPCR	Organism	Expression level, pmol/mg total MP
CCR5 Chemokine receptor	<i>H. sapiens</i>	555
DmGluRA Metabotropic glutamate receptor	<i>D. melanogaster</i>	226
mGluR5 Metabotropic glutamate receptor	<i>R. norvegicus</i>	192
Rh1 Rhodopsin	<i>D. melanogaster</i>	502
V2R Vasopressin receptor	<i>H. sapiens</i>	>1000

**Table 5: Expression levels of recombinant GPCRs in the photoreceptor cells of *Drosophila melanogaster*.** Data obtained from Panneels *et al.* 2011 [94].

More recent advancements, including expression in the eyes of *Drosophila*, offer a promising solution for a scalable production of functional recombinant membrane proteins (Table 5). Currently, as of September 2019, only five PDB entries were derived from expression in *Drosophila* – none of which were GPCRs. Future work to broaden the diversity and characterisation of varied GPCRs would invaluablely reinforce the use of this emerging technique. Overall, the field is currently in a much stronger position than a few decades ago, and will undoubtedly continue to build upon the methods reviewed here.

### Acknowledgements

This review was written and edited by Daniel Wiseman. I would like to thank my colleagues Abigail Otchere, Jaimin Patel, Romez Uddin, Naomi Pollock, Sarah Routledge and Alan Goddard for submission of text. I would also like to thank my colleagues Alice Rothnie, Cathy Slack, David Poyner, Roslyn Bill and Alan Goddard for support in editing and proofreading.

### References

1. Kobilka, B. (2007). G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta*. 1768(4): pp794-807.
2. Bockaert, J. & Pin, J. P. (1999). Molecular tinkering of G protein-coupled receptors: an evolutionary success. *Embo Journal*. 18(7): pp1723-1729.
3. Trzaskowski, B., Latek, D., Yuan, S., Ghoshdastider, U., Debinski, A., & Filipek, S. (2012). Action of molecular switches in GPCRs--theoretical and experimental studies. *Current Medicinal Chemistry*. 19(8): pp1090–1109. doi:10.2174/092986712799320556.
4. Santos, R., Ursu, O., Gaulton, A., Bento, P. A., Donadi, R. S., Bologa, C. G., Harlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T. I. & Overington, J. P. (2017). A Comprehensive Map of Molecular Drug Targets. *Nature Reviews Drug Discovery*. 16: pp19-34.
5. Ladds, G., Goddard, A. D., Hill, C., Thornton, S. & Davey, J. (2007). Differential effects of RGS proteins on  $G\alpha_q$  and  $G\alpha_{11}$  activity. *Cellular Signalling*. 19(1): pp103-113.
6. Liang, Y. L., Khoshouei, M., Deganutti, G., Glukhova, A., Koole, C., Peat, T. S., Radjainia, M., Plitzko, J. M., Baumeister, W., Miller, L. J., Hay, D. L., Christopoulos, A., Reynolds, C. A., Wooten, D. & Sexton, P. M. (2018). Cryo-EM structure of the active,  $G_s$ -protein complexed, human CGRP receptor. *Nature*. 561: pp492-497.
7. Coudrat, T., Simms, J., Christopoulos, A., Wootten, D. & Sexton, P. M. (2017). Improving virtual screening of G protein-coupled receptors via ligand-directed modelling. *PLoS Computational Biology*. 13(11): e1005819.
8. Birch, J., Axford, D., Foadi, J., Meyer, A., Eckhardt, A., Thielmann, Y. & Moraes, I. (2018). The fine art of integral membrane protein crystallisation. *Methods*. 147: pp150-162.
9. Thongin, N., Kargas, V., Clews, J. & Ford, R. C. (2018). Cryo-electron microscopy of membrane proteins. *Methods*. 147: pp176-186.
10. Tapaneeyakorn, S., Goddard, A. D., Oates, J., Willis, C. L. & Watts, A. (2011). Solution-and solid-state NMR studies of GPCRs and their ligands. *Biochimica et Biophysica Acta-Biomembranes*. 1808(6): pp1462-1475.
11. Dilworth, M. V., Piel, M. S., Bettany, K. E., Ma, P., Luo, J., Sharples, D., Poyner, D. R., Gorss, S. R., Moncoq, K., Henderson, P. J. F. & Miroux, B. (2018). Microbial expression systems for membrane proteins. *Methods*. 147: pp3-39.

12. de Mendoza, A., Sebe-Pedros, A. & Ruiz-Trillo, I. (2014). The evolution of the GPCR signaling system in eukaryotes: modularity, conservation, and the transition to metazoan multicellularity. *Genome Biology and Evolution*. 6(3): pp606-619.
13. Zhang, Z., Kuipers, G., Niemiec, L., Baumgarten, T., Slotboom, D. J., de Gier, J. W. & Hjelm, H. (2015). High-level production of membrane proteins in *E. coli* BL21 (DE3) by omitting the inducer IPTG. *Microbial Cell Factories*. 14(1): p142.
14. Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology*. 260(3): pp289-298.
15. Wagner, S., Baars, L., Ytterberg, A. J., Klussmeier, A., Wagner, C. S., Nord, O., Nygren, P. A., van Wijk K. J. & de Gier, J. W. (2007). Consequences of membrane protein overexpression in *Escherichia coli*. *Molecular and Cell Proteomics*. 6(9): pp1527-1550.
16. Baneres, J. L., Martin, A., Hullot, P., Girard, J. P., Rossi J. C. & Parello, J. (2003). Structure-based analysis of GPCR function: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. *Journal of Molecular Biology*. 329(4): pp801-814.
17. Link, A. J., Skretas, G., Strauch, E. M., Chari, N. S. & Georgiou, G. (2008). Efficient production of membrane-integrated and detergent-soluble G protein-coupled receptors in *Escherichia coli*. *Protein Science*. 17(10): pp1857-1863.
18. Wheatley, M. & Hawtin, S. R. (1999). Glycosylation of G-protein-coupled receptors for hormones central to normal reproductive functioning: its occurrence and role. *Human Reproduction Update*. 5(4): pp356-364.
19. Goddard, A. D. & Watts, A. (2012). Regulation of G protein-coupled receptors by palmitoylation and cholesterol. *BMC Biology*. 10: p27.
20. Goddard, A. D., Dijkman, P. M., Adamson, R. J. & Watts, A. (2013). Lipid-dependent GPCR dimerization. *Methods in Cell Biology*. 117: pp341-357.
21. Lundstrom, K., Wagner, R., Reinhart, C., Desmyter, A., Cherouati, N., Magnin, T., Zeder-Lutz, G., Courtot, M., Prual, C., Andre, N., Hassaine, G., Michel, H., Cambillau, C. & Pattus, F. (2006). Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. *Journal of Structural and Functional Genomics*. 7(2): pp77-91.
22. Bertin, B., Freissmuth, M., Breyer, R. M., Schutz, W., Strosberg, A. D. & Marullo, S. (1992). Functional expression of the human serotonin 5-HT<sub>1A</sub> receptor in *Escherichia coli*. Ligand binding properties and interaction with recombinant G protein alpha-subunits. *Journal of Biological Chemistry*. 267(12): pp8200-8206.
23. Freissmuth, M., Selzer, E., Marullo, S., Schutz, W. & Strosberg, A. D. (1991). Expression of two human beta-adrenergic receptors in *Escherichia coli*: functional interaction with two forms of the stimulatory G protein. *PNAS USA*. 88(19): pp8548-8552.
24. Furukawa, H. & Haga, T. (2000). Expression of functional M2 muscarinic acetylcholine receptor in *Escherichia coli*. *Journal of Biochemistry*. 127(1): pp151-161.
25. Grisshammer, R., Duckworth, R. & Henderson, R. (1993). Expression of a rat neurotensin receptor in *Escherichia coli*. *Biochemical Journal*. 295(Pt 2): pp571-576.
26. Hill, R. A. & Sillence, M. N. (1997). Improved membrane isolation in the purification of beta 2-adrenoceptors from transgenic *Escherichia coli*. *Protein Expression and Purification*. 10(1): pp162-167.
27. Kiefer, H., Krieger, J., Olszewski, J. D., von Heijne, G., Prestwich, G. D. & Breer, H. (1996). Expression of an olfactory receptor in *Escherichia coli*: purification, reconstitution, and ligand binding. *Biochemistry*. 35(50): pp16077-16084.



28. Yeliseev, A. A., Wong, K. K., Soubias, O. & Gawrisch, K. (2005). Expression of human peripheral cannabinoid receptor for structural studies. *Protein Science*. 14(10): pp2638-2653.
29. Attrill, H., Harding, P. J., Smith, E., Ross, S. & Watts, A. (2009). Improved yield of a ligand-binding GPCR expressed in *E. coli* for structural studies. *Protein Expression and Purification*. 64(1): pp32-38.
30. Baneres, J. L., Mesnier, D., Martin, A., Joubert, L., Dumuis, A. & Bockaert, J. (2005). Molecular characterization of a purified 5-HT<sub>4</sub> receptor: a structural basis for drug efficacy. *Journal of Biological Chemistry*. 280(21): pp20253-20260.
31. Busuttil, B. E., Turney, K. L. & Frauman, A. G. (2001). The expression of soluble, full-length, recombinant human TSH receptor in a prokaryotic system. *Protein Expression and Purification*. 23(3): pp369-373.
32. Inagaki, S., Ghirlando, R., White, J. F., Gvozdenovic-Jeremic, J., Northup, J. K. & Grisshammer, R. (2012). Modulation of the interaction between neurotensin receptor NTS1 and Gq protein by lipid. *Journal of Molecular Biology*. 417(1-2): pp95-111.
33. Krepiy, D., Wong, K., Gawrisch, K. & Yeliseev, A. (2006). Bacterial expression of functional, biotinylated peripheral cannabinoid receptor CB2. *Protein Expression and Purification*. 49(1): pp60-70.
34. Tian, C., Breyer, R. M., Kim, H. J., Karra, M. D., Friedman, D. B., Karpay, A. & Sanders, C. R. (2005). Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor. *Journal of the American Chemical Society*. 127(22): pp8010-8011.
35. Weiss, H. M. & Grisshammer, R. (2002). Purification and characterization of the human adenosine A (2a) receptor functionally expressed in *Escherichia coli*. *European Journal of Biochemistry*. 269(1): pp82-92.
36. Mallipeddi, S., Zvonok, N. & Makriyannis, A. (2018). Expression, Purification and Characterization of the Human Cannabinoid 1 Receptor. *Scientific Reports*. 8(1): p2935.
37. Michalke, K., Huyghe, C., Lichiere, J., Graviere, M. E., Siponen, M., Sciara, G., Lepaul, I., Wagner, R., Magg, C., Rudolph, R., Cambillau, C. & Desmyter, A. (2010). Mammalian G protein-coupled receptor expression in *Escherichia coli*: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1. *Analytical Biochemistry*. 401(1): pp74-80.
38. Petrovskaya, L. E., Shulga, A. A., Bocharova, O. V., Ermolyuk, Y. S., Kryukova, E. A., Chupin, V. V., Blommers, M. J., Arseniev, A. S. & Kirpichnikov, M. P. (2010). Expression of G-protein coupled receptors in *Escherichia coli* for structural studies. *Biochemistry (Moscow)*. 75(7): pp881-891.
39. Dodevski, I. & Pluckthun, A. (2011). Evolution of three human GPCRs for higher expression and stability. *Journal of Molecular Biology*. 408(4): pp599-615.
40. Klenk, C., Ehrenmann, J., Schutz, M. & Pluckthun, A. (2016). A generic selection system for improved expression and thermostability of G protein-coupled receptors by directed evolution. *Scientific Reports*. 6: p21294.
41. Sarkar, C. A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E. & Pluckthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *PNAS USA*. 105(39): pp14808-14813.
42. Skretas, G. & Georgiou, G. (2008). Engineering G protein-coupled receptor expression in bacteria. *PNAS USA*. 105(39): pp14747-14748.
43. Ficca, A. G., Testa, L. & Tocchini Valentini, G. P. (1995). The human beta 2-adrenergic receptor expressed in *Schizosaccharomyces pombe* retains its pharmacological properties. *FEBS Letters*. 377(2): pp140-144.

44. Sander, P., Grunewald, S., Reilinder, H. & Michel, H. (1994). Expression of the human D2s dopamine receptor in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*: a comparative study. *FEBS Letters*. 344: pp41-46.
45. Iguchi, Y., Ishii, J., Nakayama, H., Ishikura, A., Izawa, K., Tanaka, T., Ogino, C. & Kondo, A. (2010). Control of signalling properties of human somatostatin receptor subtype-5 by additional signal sequences on its amino-terminus in yeast. *Journal of Biochemistry*. 147(6): pp875-84. doi: 10.1093/jb/mvq023.
46. Hashi, H., Nakamura, Y., Ishii, J. & Kondo, A. (2018). Modifying expression modes of human neurotensin receptor type 1 alters sensing capabilities for agonists in yeast signaling biosensor. *Biotechnology Journal*. 13(4): e1700522.
47. Togawa, S., Ishii, J., Ishikura, A., Tanaka, T., Ogino, C. & Kondo, A. (2010). Importance of asparagine residues at positions 13 and 26 on the amino-terminal domain of human somatostatin receptor subtype-5 in signalling. *Journal of Biochemistry*. 147(6): pp867-73. doi: 10.1093/jb/mvq022.
48. Weston, C., Poyner, D. R., Patel, V., Dowell, S. & Ladds, G. (2014). Investigating G protein signalling bias at the glucagon-like peptide-1 receptor in yeast. *British Journal of Pharmacology*. 171(15): pp3651-65. doi: 10.1111/bph.12716.
49. Ladds, G., Davis, K., Hillhouse, E. W. & Davey, J. (2003). Modified yeast cells to investigate the coupling of G protein-coupled receptors to specific G proteins. *Molecular Microbiology*. 47(3): pp781-792.
50. Byrne, B. (2015). *Pichia pastoris* as an expression host for membrane protein structural biology. *Current Opinion in Structural Biology*. 32: pp9-17
51. Fraser, N. J. (2006). Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expression and Purification*. 49(1): pp129-37.
52. Singh, S., Gras, A., Fiez-Vandal, C., Ruprecht, J., Rana, R., Martinez, M., Strange, P. G., Wagner, R. & Byrne, B. (2008). Large-scale functional expression of WT and truncated human adenosine A2A receptor in *Pichia pastoris* bioreactor cultures. *Microbial Cell Factories*. 7: p28.
53. Jamshad, M., Charlton, J., Lin, Y. P., Routledge, S. J., Bawa, Z., Knowles, T. J., Overduin, M., Dekker, N., Dafforn, T. R., Bill, R. M., Poyner, D. R. & Wheatley, M. (2015). G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent. *Bioscience Reports*. 16;35(2): e00188. doi: 10.1042/BSR20140171.
54. Weiss, H. M., Haase, W., Michel, H., & Reiländer, H. (1998). Comparative biochemical and pharmacological characterization of the mouse 5HT5A 5-hydroxytryptamine receptor and the human beta2-adrenergic receptor produced in the methylotrophic yeast *Pichia pastoris*. *Biochemical Journal*. 330(Pt3): pp1137–1147. doi: 10.1042/bj3301137.
55. Asada, H., Uemura, T., Yurugi-Kobayashi, T., Shiroishi, M., Shimamura, T., Tsujimoto, H., Ito, K., Sugawara, T., Nakane, T., Nomura, N., Murata, T., Haga, T., Iwata, S. & Kobayashi, T. (2011). Evaluation of the *Pichia pastoris* expression system for the production of GPCRs for structural analysis. *Microbial Cell Factories*. 22;10: p24. doi: 10.1186/1475-2859-10-24.
56. Shimamura, T., Shiroishi, M., Weyand, S., Tsujimoto, H., Winter, G., Katritch, V., Abagyan, R., Cherezov, V., Liu, W., Han, G. W., Kobayashi, T., Stevens, R. C., & Iwata, S. (2011). Structure of the human histamine H1 receptor complex with doxepin. *Nature*. 475(7354): pp65-70. doi: 10.1038/nature10236.
57. Hino, T., Arakawa, T., Iwanari, H., Yurugi-Kobayashi, T., Ikeda-Suno, C., Nakada-Nakura, Y., Kusano-Arai, O., Weyand, S., Shimamura, T., Nomura, N., Cameron, A. D., Kobayashi, T., Hamakubo, T., Iwata, S. & Murata, T. (2012). G-protein-coupled

- receptor inactivation by an allosteric inverse-agonist antibody. *Nature*. 482(7384): pp237-40. doi: 10.1038/nature10750.
58. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Applied Microbiology and Biotechnology*. 98(12): pp5301-17. doi: 10.1007/s00253-014-5732-5.
59. Cregg, J. M., Barringer, K. J., Hessler, A. Y. & Madden, K. R. (1985). *Pichia pastoris* as a host system for transformations. *Molecular and Cellular Biology*. 5: pp3376-3385.
60. Bollok, M., Resina, D., Valero, F. & Ferrer, P. (2009). Recent Patents on the *Pichia Pastoris* Expression System: Expanding the Toolbox for Recombinant Protein Production. *Recent Patents on Biotechnology*. 3: pp192-201
61. Prielhofer, R., Maurer, M., Klein, J., Wenger, J., Kiziak, C., Gasser, B. & Mattanovich, D. (2013). Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microbial Cell Factories*. 24;12: p5. doi: 10.1186/1475-2859-12-5.
62. Xu, N., Zhu, J., Zhu, Q., Xing, Y., Cai, M., Jiang, T., Zhou, M. & Zhang, Y. (2018). Identification and characterization of novel promoters for recombinant protein production in yeast *Pichia pastoris*. *Yeast*. 35(5): pp379-385. doi: 10.1002/yea.3301.
63. Dong, C., Filipeanu, C. M., Duvernay, M. T. & Wu, G. (2007). Regulation of G protein-coupled receptor export trafficking. *Biochimica et Biophysica Acta*. 1768(4): pp853-70.
64. Chen, Q., Miller, L. J. & Dong, M. (2010). Role of N-linked glycosylation in biosynthesis, trafficking, and function of the human glucagon-like peptide 1 receptor. *American Journal of Physiology-Endocrinology and Metabolism*. 299(1): E62-8. doi: 10.1152/ajpendo.00067.2010.
65. Nehring, R. B., Richter, D. & Meyerhof, W. (2000). Glycosylation affects agonist binding and signal transduction of the rat somatostatin receptor subtype 3. *Journal de Physique*. 94(3-4): pp185-192.
66. Choi, B. K., Bobrowicz, P., Davidson, R. C., Hamilton, S. R., Kung, D. H., Li, H., Miele, R. G., Nett, J. H., Wildt, S. & Gerngross, T. U. (2003). Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *PNAS USA*. 29;100(9): pp5022-7.
67. Vieira Gomes, A. M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F. & Parachin, N. S. (2018). Comparison of Yeasts as Hosts for Recombinant Protein Production. *Microorganisms*. 29;6(2). pii: E38. doi: 10.3390/microorganisms6020038.
68. Laron, Z. (2001). Insulin-like growth factor 1 (IGF-1): A growth hormone. *Molecular Pathology*. 54: pp311–316. doi: 10.1136/mp.54.5.311.
69. Yurugi-Kobayashi, T., Asada, H., Shiroishi, M., Shimamura, T., Funamoto, S., Katsuta, N., Ito, K., Sugawara, T., Tokuda, N., Tsujimoto, H., Murata, T., Nomura, N., Haga, K., Haga, T., Iwata, S. & Kobayashi, T. (2009). Comparison of functional non-glycosylated GPCRs expression in *Pichia pastoris*. *Biochemical and Biophysical Research Communications*. 380(2): pp271-276.
70. Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. (2009). Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nature Protocols*. 4: pp58–70.
71. Pucadyil, T. J. & Chattopadhyay, A. (2004). Cholesterol modulates ligand binding and G-protein coupling to serotonin (1A) receptors from bovine hippocampus. *Biochimica et Biophysica Acta*. 1663(1-2): pp188-200.
72. Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V.-P., Chien, E. Y. T., Velasquez, J., Kuhn, P. & Stevens, R. C. (2008). A specific cholesterol binding

- site is established by the 2.8 Å structure of the human  $\beta$ 2-adrenergic receptor. *Structure* 16: pp897–905. doi: 10.1016/j.str.2008.05.001.
73. Paila, Y. D., Tiwari, S. & Chattopadhyay, A. (2009). Are specific nonannular cholesterol binding sites present in G-protein coupled receptors? *Biochimica et Biophysica Acta*. 1788(2): pp295-302. doi: 10.1016/j.bbamem.2008.11.020.
74. Hirz, M., Richter, G., Leitner, E., Wriessnegger, T. & Pichler, H. (2013). A novel cholesterol-producing *Pichia pastoris* strain is an ideal host for functional expression of human Na,K-ATPase  $\alpha$ 3 $\beta$ 1 isoform. *Applied Microbiology and Biotechnology*. 97(21): pp9465-78. doi: 10.1007/s00253-013-5156-7.
75. Milić, D. & Veprintsev, D. B. (2015). Large-scale production and protein engineering of G protein-coupled receptors for structural studies. *Frontiers in Pharmacology*. 31;6: p66. doi: 10.3389/fphar.2015.00066.
76. Macauley-Patrick, S., Fazenda, M. L., McNeil, B. & Harvey, L. M. (2005). Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 22(4): pp249-70.
77. Saarenpaa, T., Jaakola, V. P. & Goldman, A. (2015). Baculovirus-mediated expression of GPCRs in insect cells. *Methods in Enzymology*. 556: pp185-218.
78. McKenzie, E. A. & Abbott, W. M. (2018). Expression of recombinant proteins in insect and mammalian cells. *Methods*. 147: pp40-49.
79. Roldao, A., Oliveira, R., Carrondo, M. J. & Alves, P. M. (2009). Error assessment in recombinant baculovirus titration: evaluation of different methods. *Journal of Virological Methods*. 159(1): pp69-80.
80. Aloia, A. L., Glatz, R. V., McMurchie, E. J. & Leifert, W. R. (2009). GPCR expression using baculovirus-infected Sf9 cells. *Methods in Molecular Biology*. 552: pp115-129.
81. Dawaliby, R., Trubbia, C., Delporte, C., Masureel, M., van Antwerpen, P., Kobilka, B. K. & Govaerts, C. (2016). Allosteric regulation of G protein-coupled receptor activity by phospholipids. *Nature Chemical Biology*. 12(1): pp35-39.
82. Thomas, J. A. & Tate, C. G. (2014). Quality control in eukaryotic membrane protein overproduction. *Journal of Molecular Biology*. 426(24): pp4139-4154.
83. Andrell, J. & Tate, C. G. (2013). Overexpression of membrane proteins in mammalian cells for structural studies. *Molecular Membrane Biology*. 30(1): pp52-63.
84. Jiang, L., Teng, G. M. K., Chan, E. Y. M., Au, S. W. N., Wise, H., Lee, S. S. T. & Cheung, W-T. (2012). Impact of cell type and epitope tagging on heterologous expression of G-Protein Coupled Receptor: A Systematic Study on Angiotensin Type II Receptor. *PLOS ONE*. 7(10): e47016.
85. Guixa-Gonzalez, R., Albasanz, J. L., Rodriguez-Espigares, I., Pastor, M., Sanz, F., Marti-Solano, M., Manna, M., Martinez-Seara, H., Hildebrand, P. W., Martin, M. & Selent, J. (2017). Membrane cholesterol access into a G-protein-coupled receptor. *Nature Communications*. 8: p14505.
86. Lagane, B., Gaibelet, G., Meilhoc, E., Masson, J. M., Cezanne, L. & Lopez, A. (2000). Role of sterols in modulating the human  $\mu$ -opioid receptor function in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*. 275(43): pp33197-33200.
87. Chakraborty, R., Xu, B., Bhullar, R. P. & Chelikani, P. (2015). Expression of G protein-coupled receptors in Mammalian cells. *Methods in Enzymology*. 556: pp267-281.
88. Zucchelli, S., Patrucco, L., Persichetti, F., Gustincich, S. & Cotella, D. (2016). Engineering Translation in Mammalian Cell Factories to Increase Protein Yield: The Unexpected Use of Long Non-Coding SINEUP RNAs. *Computational and Structural Biotechnology Journal*. 14: pp404-410.

89. Thomas, P. & Smart, T. G. (2005). HEK293 cell line: a vehicle for the expression of recombinant proteins. *Journal of Pharmacological and Toxicological Methods*. 51(3): pp187-200.
90. Subedi, G. P., Johnson, R. W., Moniz, H. A., Moremen, K. W. & Barb, A. (2015). High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension. *Journal of Visualised Experiments*. (106): e53568.
91. Griffith, D. A., Delipala, C., Leadsham, J., Jarvis, S. M. & Oesterhelt, D. (2003). A novel yeast expression system for the overproduction of quality-controlled membrane proteins. *FEBS Letters*. 553(1-2): pp45-50.
92. Panneels, V., Eroglu, C., Cronet, P. & Sinning, I. (2003). Pharmacological characterization and immunoaffinity purification of metabotropic glutamate receptor from *Drosophila* overexpressed in Sf9 cells. *Protein Expression and Purification*. 30(2): pp275-282.
93. Kumar, J. P. & Ready, D. F. (1995). Rhodopsin plays an essential structural role in *Drosophila* photoreceptor development. *Development*. 121(12): pp4359-4370.
94. Panneels, V., Kock, I., Krijnse-Locker, J., Rezgouei, M. & Sinning, I. (2011). *Drosophila* photoreceptor cells exploited for the production of eukaryotic membrane proteins: receptors, transporters and channels. *PLoS One*. 6(4): e18478.
95. Phelps, C. B. & Brand, A. H. (1998). Ectopic gene expression in *Drosophila* using GAL4 system. *Methods*. 14(4): pp367-379.
96. Eroglu, C., Cronet, P., Panneels, V., Beaufils, P. & Sinning, I. (2002). Functional reconstitution of purified metabotropic glutamate receptor expressed in the fly eye. *EMBO Reports*. 3(5): pp491-496.
97. Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J. B., Salvi, D., Seigneurin-Berny, D., Richaud, P., Joyard, J., Pignol, D., Sabaty, M., Desnos, T., Pebay-Peyroula, E., Darrouzet, E., Vernet, T. & Rolland, N. (2011). Heterologous expression of membrane proteins: choosing the appropriate host. *PLoS One*. 6(12): e29191.
98. Schiller, B., Hykollari, A., Yan, S., Paschinger, K. & Wilson, I. B. (2012). Complicated N-linked glycans in simple organisms. *Journal of Biological Chemistry*. 393(8): pp661-673.
99. Lee, S. C., Knowles, T. J., Postis, V. L., Jamshad, M., Parslow, R. A., Lin, Y. P., Goldman, A., Sridhar, P., Overduin, M., Muench, S. P. & Dafforn, T. R. (2016). A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature Protocols*. 11(7): pp1149-1162.
100. Prive, G. G. (2007). Detergents for the stabilization and crystallization of membrane proteins. *Methods*. 41(4): pp388-397.
101. Seddon, A. M., Curnow, P. & Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta-Biomembranes*. 1666: pp105-117.
102. le Maire, M., Champeil, P. & Moller, J. V. (2000). Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et Biophysica Acta*. 1508: pp86-111.
103. Arachea, B. T., Sun, Z., Potente, N., Malik, R., Isailovic, D. & Viola, R. E. (2012). Detergent selection for enhanced extraction of membrane proteins. *Protein Expression and Purification*. 86: pp12-20.
104. Anatrace. *Detergents and Their Uses in Membrane Protein Science*. Pdf available at: <https://www.anatrace.com/Technical-Documentation/Catalogs/Anatrace-Detergents-Booklet-FINAL> (Accessed: 25/6/19).

105. Lyons, J. A., Shahsavari, A., Paulsen, P. A., Pederson, B. P. & Nissen, P. (2016). Expression strategies for structural studies of eukaryotic membrane proteins. *Current Opinion in Structural Biology*. 38: pp137-144.
106. Parker, J. L. & Newstead, S. (2012). Current trends in  $\alpha$ -helical membrane protein crystallization: An update. *Protein Science*. 21: pp1358-1365.
107. Anandan, A. & Vrielink, A. (2016). 'Detergents in Membrane Protein Purification and Crystallisation', in Moraes, I. (ed.) *The Next Generation in Membrane Protein Structure Determination. Advances in Experimental Medicine and Biology*. 922: pp13-28.
108. Schuck, S., Honsho, M., Ekroos, K., Shevchenko, A. & Simons, K. (2003). Resistance of cell membranes to different detergents. *PNAS USA*. 100(10): pp5795-5800.
109. Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. (2017). The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nature Reviews Molecular Cell Biology*. 18: pp361-374.
110. Popot, J. L. (2010). Amphipols, nanodiscs and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annual Review of Biochemistry*. 79: pp737-775.
111. Popot, J.L., Berry, E. A., Charvolin, D., Creuzenet, C., Ebel, C., Engelman, D. M., Flöetenmeyer, M., Giusti, F., Gohon, Y., Hervé, P., Hong, Q., Lakey, J. H., Leonard, K., Shuman, H. A., Timmins, P., Warschawski, D. E., Zito, F., Zoonens, M., Pucci, B. & Tribet, C. (2003). Amphipols: Polymeric surfactants for membrane biology research. *Cellular and Molecular Life Sciences*. 60(8): pp1559-1574.
112. McLean, M. A., Gregory, M. C. & Sligar, S. G. (2018). Nanodiscs: A Controlled Bilayer Surface for the Study of Membrane Proteins. *Annual Review of Biophysics*. 47: pp107-124.
113. Oluwole, A. O., Danielczak, B., Meister, A., Babalola, J. O., Vargas, C. & Keller, S. (2017). Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer. *Angewandte Chemie International Edition in English*. 56(7): pp1919-1924.
114. Stroud, Z., Hall, S. C. L. & Dafforn, T. R. (2018). Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. *Methods*. 147: pp106-117.
115. Muk, S., Ghosh, S., Achuthan, S., Chen, X., Yao, X., Sandhu, M., Griffor, M. C., Fennell, K. F., Che, Y., Shanmugasundaram, V., Qiu, X., Tate, C. G. & Vaidehi, N. (2019). Machine learning for prioritization of thermostabilizing mutations for G-Protein Coupled Receptors. *BioRxiv*. 715375.
116. Knowles, T. J., Finka, R., Smith, C., Lin, Y-P., Dafforn, T. & Overduin, M. (2009) Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer. *Journal of the American Chemical Society*. 131(22): pp7484-7485.
117. Lee, S. C., Knowles, T. J., Postis, V. L. G., Jamshad, M., Parslow, R. A., Lin, Y-P., Goldman, A., Sridhar, P., Overduin, M., Muench, S. P. & Dafforn, T. R. (2016). A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature Protocols*. 11: pp1149-1162.
118. Scheidelaar, S., Koorengel, M. C., Pardo, J. D., Meeldijk, J. D., Breukink, E. & Killian, J. A. (2015). Molecular model for the solubilisation of membranes into nanodisks by styrene maleic acid copolymers. *Biophysical Journal*. 108(2): pp279-290.
119. Pollock, N. L., Lee, S. C., Patel, J. H., Gulamhussein, A. A. & Rothnie, A. J. (2017). Structure and function of membrane proteins encapsulated in a polymer-

- bound lipid bilayer. *Biochimica et Biophysica Acta-Biomembranes*. 1860(4): pp809-817.
120. Simon, K. S., Pollock, N. L. & Lee, S. C. (2018). Membrane protein nanoparticles: the shape of things to come. *Biochemical Society Transactions*. 46(6): pp1495-1504.
121. Lee, S. C. & Pollock, N. L. (2016). Membrane proteins: is the future disc shaped? *Biochemical Society Transactions*. 44(4): pp1011-1018.
122. Wheatley, M., Charlton, J., Jamshad, M., Routledge, S. J., Bailey, S., La-Borde, P. J., Azam, M. T., Logan, R. T., Bill, R. M., Dafforn, T. R. & Poyner, D. R. (2015). GPCR–styrene maleic acid lipid particles (GPCR–SMALPs): their nature and potential. *Biochemical Society Transactions*. 44(2): pp619-623.
123. Hall, S. C. L., Tognoloni, C., Charlton, J., Bragginton, E. C., Rothnie, A. J., Sridhar, P., Wheatley, M., Knowles, T. J., Arnold, T., Edler, K. J. & Dafforn, T. R. (2018). An acid-compatible co-polymer for the solubilisation of membranes and proteins into lipid bilayer-containing nanoparticles. *Nanoscale*. 10: pp10609-10619.
124. Damian, M., Pons, V., Renault, P., M'Kadmi, C., Delort, B., Hartmann, L., Kaya, A. I., Louet, M., Gagne, D., Slah, K. B. H., Denoyelle, S., Ferry, G., Boutin, J. A., Wagner, R., Fehrentz, J-A., Martinez, J., Marie, J., Floquet, N., Gales, C., Mary, S., Hamm, H. E. & Baneres, J-L. (2018). GHSR-D2R heteromerization modulates dopamine signalling through and effect on G protein conformation. *PNAS*. 115(17): pp4501-4506.
125. Logez, C., Damian, M., Legros, C., Dupre, C., Guery, M., Mary, S., Wagner, R., M'Kadmi, C., Nosjean, O., Fould, B., Marie, J., Fehrentz, J-A., Martinez, J., Ferry, G., Boutin, J. A. & Baneres, J-L. (2015). Detergent-free isolation of functional G-Protein Coupled Receptors into nanomatrix lipid particles. *Biochemistry*. 55: pp38-48.
126. Broecker, J., Eger, B. T. & Ernst, O. P. (2017). Crystallogenesis of membrane proteins mediated by polymer-bounded lipid nanodiscs. *Structure*. 25(2): 99384-99392.
127. Charlton, J. (2016). *Solubilisation and characterisation of G protein-coupled receptors using styrene maleic acid polymer*. PhD thesis. University of Birmingham, Birmingham. Viewed 13/8/19, <<https://etheses.bham.ac.uk/id/eprint/6525/1/Charlton16PhD.pdf>>.
128. Lin, S. H. & Guidotti, G. (2009). Purification of membrane proteins. *Methods in Enzymology*. 463: pp619-629.
129. André, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G., Magnin, T., Pattus, F., Michel, H., Wagner, R., Reinhart, C. (2006). Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Science*. 15(5): pp1115-1126. doi: 10.1110/ps.062098206.
130. McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G. & Foord, S. M. (1998). RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. *Nature*. 393(6683): p333.
131. Ninfa, A., Ballou, D. & Benore, M. (2009). *Fundamental Laboratory Approaches for Biochemistry and Biotechnology*. 2<sup>nd</sup> ed. Wiley.
132. Kumar, P. (2018). *Fundamentals and Techniques of Biophysics and Molecular Biology*. 2<sup>nd</sup> ed. New Delhi: Pathfinder Publication.
133. Garret, R. H. & Grisham, C. M. (2013). *Biochemistry*. 5<sup>th</sup> ed. Belmont, CA: Brooks/Cole, Cengage Learning.
134. Skoog, D. A., Holler, F. J. & Crouch, S. R. (2006). 'Liquid Chromatography' in *Principles of instrumental analysis*. 6<sup>th</sup> ed. Belmont, CA: Thomson Brooks/Cole.

135. Castell, O. K., Dijkman, P. M., Wiseman, D. N. & Goddard, A. D. (2018). Single molecule fluorescence for membrane proteins. *Methods*. 147: pp221-228.
136. Uddin, R., Simms, J. & Poyner, D. R. (2018). Functional characterisation of G protein-coupled receptors. *Methods*. 147: pp213-220.

Journal Pre-proof



## Highlights

- 1713% increase in unique membrane protein structures since 2000.
- Variety of expression methods available to GPCR researchers, now including the eyes of *Drosophila melanogaster*.
- Development of novel solubilisation strategies including the styrene maleic acid copolymer.

Journal Pre-proof