New advances in production and functional folding of G-protein-coupled receptors.

Jean-Louis Banères, Jean-Luc Popot, Bernard Mouillac

New advances in production and functional folding of G protein-coupled receptors

Jean-Louis Banères[1,2], Jean-Luc Popot[3,4] and Bernard Mouillac[5-7]

[1]CNRS, UMR-5247, Institut des Biomolécules Max Mousseron, Faculté de Pharmacie, F-34000 Montpellier, France

[2]Universités de Montpellier 1 & 2, 15 avenue Charles Flahault, F-34000 Montpellier, France

[3]CNRS, UMR-7099, Laboratoire de Physico-Chimie Moléculaire des Protéines Membranaires, F-75000 Paris, France

[4]Université Paris-7, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, F-75000 Paris, France

[5]CNRS, UMR-5203, Institut de Génomique Fonctionnelle, Département de Pharmacologie Moléculaire, F-34000 Montpellier, France

[6]INSERM, U661, F-34000 Montpellier, France

[7]Universités de Montpellier 1 & 2, 141 rue de la Cardonille, F-34000 Montpellier, France

Corresponding author: Mouillac, B. (Bernard.Mouillac@igf.cnrs.fr).

Abstract

G protein-coupled receptors (GPCRs), the largest family of integral membrane proteins, participate in the regulation of many physiological functions and are the targets of around 30% of currently marketed drugs. However, knowledge of the structural and molecular bases of GPCR function remains limited, owing to difficulties related to their overexpression, purification and stabilization. The development of new strategies aimed at obtaining large amounts of functional GPCRs is therefore crucial. Here, we review the most recent advances in production and functional folding of GPCRs from *Escherichia coli* inclusion bodies. Major breakthroughs open exciting perspectives for structural and dynamic investigations of GPCRs. In particular, combining targeting to bacterial inclusion bodies with amphipol-assisted folding is emerging as a highly powerful strategy.

Overexpressing membrane proteins for structural and biophysical studies: still a challenge

Structural information on integral membrane proteins (IMPs) remains limited. According to databases of known three-dimensional (3D) structures of IMPs (http://blanco.biomol.uci.edu/Membrane_proteins_xtal.html; http://www.mpdb.tcd.ie; http://www.drorlist.com/nmr/MPNMR.html), the crystal or NMR structures of only ~280 IMPs have been solved to date, as compared to tens of thousands of soluble proteins. Worse, only a small fraction (~35%) of these proteins are eukaryotic, even though IMPs constitute 20-30% of the proteins encoded by human and other eukaryotic genomes. Among IMPs, G protein-coupled receptors (GPCRs), whose transmembrane region is a 7-helix bundle, constitute the largest family [1]. More than 800 GPCRs have been identified, representing 2-3% of the coding sequences in the human genome. They are involved in most essential cellular processes and are the targets of around 30% of current pharmaceutical drugs. It is therefore critical to gain detailed knowledge of their structures and their dynamics in order to understand their functions and/or dysfunctions, as well as to rationally design selective therapeutic compounds. However, except for rhodopsin, whose crystal structure has been solved following its extraction from the retina [2,3], the low natural abundance of GPCRs generally precludes their purification in biochemically relevant amounts. Overexpression is thus a prerequisite to investigating their structure or analyzing their conformational transitions upon interaction with ligands or with signaling proteins like G proteins and arrestins.

Overexpressing GPCRs, however, is still problematic, often resulting in low yield, protein aggregation or misfolding, if not cell toxicity. Consequently, although crystal structures of ligand-bound β_1- and β_2-adrenergic [4,5], adenosine A_{2A} [6], chemokine CXCR4 [7], and dopamine D3 [8] receptors have been recently reported, investigating the structure and dynamics of most GPCRs remains a daunting task. Many overexpression systems have

been tested. GPCRs have been successfully produced by cell-free synthesis [9] and by heterologous expression in mammalian [10] and insect [11] cells, in the photoreceptor cells of *Drosophila* [12], *Xenopus* [13] and mouse [14], and in such microbes as yeasts and bacteria [15,16]. As discussed below, expression in *Escherichia coli* holds great promises, not only due to its simplicity, rapidity, safety, scalability or genetic tractability, but also in terms of quantity and homogeneity of the recombinant protein.

Once efficient expression has been achieved, purifying sufficient amounts of native-like, functional and stable protein still remains a formidable challenge. GPCRs, as all IMPs, have to be handled in aqueous solutions in complex with surfactants, usually detergents. Because detergents tend to be inactivating, identifying a detergent or lipid/detergent mixture that ensures protein homogeneity, functionality and stability is often a limiting step. Nevertheless, several overexpressed GPCRs have been successfully purified in a stable (often engineered) and functional form, allowing their crystallization and structure determination. For instance, the adenosine A_{2A} receptor has been purified in a fully functional form and crystallized in *n*-dodecyl-β-D-maltopyranoside (DDM) mixed with cholesterol hemisuccinate (CHS) [6]. The β_1-adrenergic receptor has been crystallized in octylthioglucoside [4], whereas the β_2-adrenergic receptor was best solubilized, purified and crystallized in DDM [5]. DDM was also compatible with NMR spectroscopy analysis of the β_2-adrenergic receptor [17]. While DDM is often used, it seems that an optimized surfactant environment has to be identified for each GPCR. Amphipols (APols), a new class of surfactants, can efficiently substitute for detergents to stabilize IMPs and offer a very promising alternative medium [18]. This review focuses on GPCR bacterial expression and on their functional folding using APols.

Overexpression of GPCRs in bacteria: targeting the inner membrane or inclusion bodies?

GPCRs have been overexpressed in bacteria following two different approaches. Since GPCRs are plasma membrane proteins, targeting recombinant receptors to the inner membrane of the bacterium was initially considered as the most obvious strategy (**Figure 1**). In most cases, however, this leads to severe cell toxicity and low levels of expression. A more efficient insertion into the bacterial inner membrane can be achieved by fusing the GPCR to a protein helper partner. Thus, coupling *E. coli* β-galactosidase (114 kDa) to the N-terminus of the human β_2-adrenergic receptor led to measurable membrane expression [19]. The combination of *E. coli* maltose-binding protein (MBP, 43 kDa), used as an N-terminal fusion partner, with *E. coli* thioredoxine A (TRX, 10 kDa), added at the C-terminus of the GPCR, has been shown to be particularly well adapted for expression of the rat neurotensin NTS1 or the cannabinoid CB2 receptors [20]. The human adenosine A_{2A} receptor has been highly expressed with only MBP fused at the N-terminus [21]. Fusion of the jellyfish green fluorescent protein (GFP, 27 kDa) to the C-terminus of the human cannabinoid CB1 and bradykinin B_2 receptors has led to efficient membrane expression [22]. Although membrane expression of the neurotensin receptor has been demonstrated to be highly successful and applied to automated large-scale purification [23], the MBP-GPCR-TRX fusion strategy cannot be generally applied without extensive receptor truncations or modifications.

Expression of heterologous proteins in *E. coli* is frequently associated with incorrect folding and accumulation of the recombinant protein in cytoplasmic aggregates named inclusion bodies (IBs). Targeting GPCRs to IBs combines many advantages. IBs are mechanically stable and can be easily isolated from other cell constituents by centrifugation, they are not toxic to the cell, and they are resistant to proteolytic degradation. Production of GPCRs in IBs can be massive (**Table 1**). It implies, however, that the receptors thus expressed have to

be subsequently refolded to their native state, which constitutes a difficult challenge (see below). This strategy has been first successfully developed for the rat olfactory OR5 receptor and several other GPCRs [24], and subsequently improved for the human leukotriene BLT1 and the human serotonin 5-HT_{4A} receptors [25,26]. In most cases, once again, a fusion partner is needed for efficient production (**Figure 1**). Except for the BLT1 receptor, which has been expressed in IBs after fusion to a short T7 tag [25], most GPCRs were coupled to a large fusion partner such as the schistosomal glutathione S-transferase (GST, 25 kDa), and had to be truncated at their N-termini. The serotonin $5HT_{4A}$ receptor was efficiently expressed after fusion to bacterial ketosteroid isomerase (KSI, 12 kDa) [26], but KSI proved to be inefficient for other receptors (J.-L. Banères, unpublished). A recent high-throughput effort at large-scale production of more than 100 GPCRs as bacterial IBs has shown that a majority of them can be expressed in quantities sufficient for solubilization and purification [27,28]. This extensive study evaluated the efficiency of various fusion partners, namely GST, MBP, TRX or the *E. coli* N-utilization substance A (NusA, 50 kDa), to target GPCRs to IBs. Depending on culture conditions, GST and TRX were identified as most efficient, although some GPCRs could be overexpressed without any protein tag.

The use as targeting partner of an α_5 integrin fragment (α_5I, 31 kDa) has allowed many rhodopsin-like GPCRs to be expressed at high levels regardless of their length (from 337 to 472 amino acids), their G protein coupling selectivity, or the nature of their endogenous ligands. This efficient and apparently generic procedure has been successfully applied to expressing the β_3-adrenergic receptor, the vasopressin V2 and V1b and oxytocin OTR receptors, the chemokine CCR5 and CXCR4 and chemokine-like ChemR23 receptors, the ghrelin GHS-1a receptor, the cannabinoid CB1 receptor and the leukotriene BLT1, BLT2, CysLT1 and CysLT2 receptors, without requiring any optimization of either the GPCR coding sequence, the cell culture conditions, or the extraction/purification procedures [29,30].

The α_5I fusion strategy represents an important breakthrough for *in vitro* studies aimed at understanding the molecular bases of GPCR function and structure, and potentially for other membrane protein families.

Comparison of the different expression strategies in *E. coli* (**Table 1**) suggests that targeting GPCRs to IBs, presents an interesting potential in terms of both the amounts produced and general applicability.

Functional folding of GPCRs from IBs using classical surfactants: the state of the art

Following expression, IBs are first solubilized under denaturing conditions. Then, following purification (using a metal-affinity chromatography procedure for instance), the fusion partner used for high-level expression has to be removed, usually through proteolytic cleavage. A notable advantage of α_5I-GPCR fusions regarding this crucial biochemical step is that α_5I keeps the receptor soluble after dialysis in the absence of denaturing agents, greatly facilitating an efficient proteolytic cleavage of the fusion protein [29]. After another purification step in SDS buffers, folding is then initiated by transfer from SDS to other surfactants (**Figure 2**). Folding efficiency depends on the competition between protein aggregation and 3D structure formation as well as on the ability of the receiving surfactant to stabilize the native 3D state of the folded receptor [31].

Efficient folding implies that the solubilized protein is not aggregated to start with. Globular proteins expressed in IBs can be efficiently solubilized by chaotropic agents such as urea or guanidinium chloride. In contrast, the solubilization of IMPs requires harsh detergents such as sodium dodecyl sulfate (SDS) or *N*-lauroylsarcosine (or organic solvents). In SDS, proteins in general and, in particular, IMPs such as bacteriorhodopsin (BR) [32], the μ-opioid receptor [33] or the small multidrug transporter EmrE [34], retain or acquire a significant amount of α-helical secondary structure. Given that some of the helical segments present in

SDS solution are likely to overlap regions that form transmembrane helices in the folded protein, a SDS-solubilized GPCR should probably be considered not as fully unfolded, but rather as partially prefolded, as far as the secondary structure is considered. If we look at the μ-opioid receptor as a GPCR reference, its α-helical content determined in 0.1% SDS solution is around 40% at pH 7-8 [33], a value in agreement with the predicted secondary structure of the full-length protein (50-54%).

GPCR folding is initiated by displacing the denaturing detergent with a milder surfactant. Under these conditions, regions that have a propensity to fold may do so, allowing native-like interactions between folded segments to form. These can be intramolecular, which may lead to correct folding, or intermolecular, leading to aggregation. Finding favorable folding conditions therefore implies identifying a surfactant or surfactant mixture that will favor intramolecular interactions and then efficiently stabilize the native fold of the protein. Various such environments have been reported so far, although the limited number of successful examples makes inferring general rules difficult. The refolding environments include classical detergents and lipid/detergent mixtures, bicelles, lipid vesicles and, finally, original surfactants such as APols (**Figure 2**).

Efficient folding in detergents has been reported for a limited number of GPCRs such as the leukotriene receptors BLT1 and BLT2 (**Table 2**). BLT1 was folded as a functional protein to ~30% in LDAO [25] whereas the BLT2 receptor was folded as a functional state to ~4% in DPC/HDM mixtures [29]. In both cases, adding lipids (e.g. asolectin) was required for improving the percentage of functional recovery. Other GPCRs have been reported to efficiently fold in detergent micelles. As stated above, the OR5 receptor was first folded in digitonin before insertion in lipid vesicles [24]. Fluorescence-monitored ligand binding assays demonstrated that about 80% of the folded OR5 receptor bound its lilial ligand [24]. More recently, refolding of the SDS-solubilized parathyroid hormone receptor 1 (PTH-1R) and of

CB1 receptors has been performed by exchanging the SDS for a mixture of the non-ionic detergents DDM and Cymal 6 [28]. In this study, ligand binding assays demonstrated that ~30% of the folded CB1 was functional. The glucagon-like peptide-1 receptor (GLP-1R) has been reported to fold upon transfer from SDS to Brij78 as a functional protein to ~40% [35]. In all these cases, i.e. for the PTH1R, CB1 and GLP-1R, the folding process was carried out in the presence of methyl-β-cyclodextrin, used to strip off SDS. Removal of dodecylsulfate can also be achieved by precipitation using K^+ ions (see below) [36].

The efficiency of alternate membrane-like environments to fold and stabilize GPCRs recovered from IBs has also been explored. Certain mixtures of long-chain and short-chain phospholipids assemble as bilayer discs, called bicelles, which mimic the membrane environment (**Figure 2**). A limited number of GPCRs, specifically the serotonin 5-HT_{4A} [26] and the neuropeptide Y_2 receptors [37], have been folded to a native-like conformation in DMPC/CHAPS bicelles with folding yields of ~25 % and ~65%, respectively.

A few cases of successful folding in lipid vesicles of GPCRs recovered from IBs have also been reported. The first such example was described in the pioneering work of Kiefer's group on the OR5 olfactory receptor, in which the overexpressed receptor was solubilized in the strong, negatively charged detergent *N*-lauroylsarcosine and then folded by transfer to the non-denaturing detergent digitonin [24]. The digitonin-folded receptor was able to bind its ligand (see above), thus providing yet another example of successful folding in detergent. The OR5 receptor was subsequently reconstituted in lipid vesicles by supplementing it with DDM/POPC/POPG mixtures before removing the detergent by adsorption onto hydrophobic beads. Under such conditions, the protein was stabilized in a fully ligand-competent state (~1% of the solubilized and purified material), as assessed by photoaffinity labeling.

As in the case of bicelles, successful folding of GPCRs by direct transfer to lipids is limited to a few examples, such as the human neuropeptide Y_1 receptor [38].

Amphipol-assisted folding of GPCRs: a generic approach?

As summarized in the previous section, folding in detergents or detergent/lipid mixtures has thus far yielded only a handful of functional GPCRs. Moreover, even for those GPCRs that have been folded under such conditions, folding yields are usually low, conditions are highly idiosyncratic, and identifying them is very time-consuming. Developing a more general approach to folding GPCRs recovered from IBs would be of great interest. This has led to testing APols as a possible generic folding medium.

APols were initially designed, synthesized and validated as mild alternatives to classical detergents [39]. They are defined as “amphipathic polymers that are able to keep individual IMPs soluble under the form of small complexes” [18]. What is of interest here is that : (i) most IMPs are more (and generally much more) stable in APols than they are in detergent solutions [18,40], and (ii) APols have proven an efficient medium in which to fold IMPs to their native state [41,42]. APols are relatively short polymers (their mass is typically in the 8-20-kDa range) that carry a high density both of hydrophobic chains and of highly hydrophilic groups. The prototypal APol, named A8-35 (**Figure 3A**) [40,41], remains by far the most thoroughly studied and most widely used APol [18,43,44]. The high solubility of A8-35 in water is due to the presence of carboxylates. As a consequence, A8-35 becomes insoluble in acidic solutions [45,46], a limitation that has prompted the development of alternative chemical structures such as sulfonated APols (SAPols; **Figure 3B**) [47] or glucose-based, non-ionic APols (NAPols; **Figure 3C**) [48,49], both of which are insensitive to pH. In aqueous solutions, APols form small, micelle-like particles, each of which comprises only a few APol molecules (~4 of them in the case of A8-35) [46].

IMP/APol complexes are typically obtained by one of the two following routes. Usually, a native IMP in detergent solution is supplemented with APols. This results in the

formation of ternary complexes [50,51]. The detergent is then removed, yielding small, compact IMP/APol complexes [52], in which the APol forms a thin layer covering the hydrophobic transmembrane surface of the protein [53,54]). Alternatively, a denatured IMP in SDS and/or urea is transferred to APols, during which process it recovers or adopts its native 3D structure (**Figure 2,** and see below). As a rule, APol-trapped IMPs are much more stable than their detergent-solubilized counterparts [18,40]. The underlying mechanisms are several. The most important factor is that APols are less efficient than detergents at disrupting the protein/protein and protein/lipid interactions that determine and stabilize the 3D structure of IMPs [40,47]. This led to the suggestion that, in addition to being less aggressive towards properly folded, native IMPs, APols might provide a favorable medium for the formation or reformation of native-like interactions starting from a denatured protein.

APol-mediated IMP folding was first demonstrated using as models urea-solubilized OmpA and FomA, two β-barrel outer membrane proteins (OMPs) from the eubacteria *E. coli* and *Fusobacterium nucleatum*, respectively, and a paradigmatic α-helical IMP, BR, a light-driven proton pump from the plasma membrane of the archaebacterium *Halobacterium salinarium* [41]. In the latter case, the SDS-solubilized BR was folded by precipitating dodecylsulfate as its potassium salt [36] in the presence of A8-35. The rationale behind the choice of this apparently particular procedure is to proceed as rapidly as possible to the exchange of SDS for APols, leaving protein little chance to explore misfolding or aggregation opportunities offered by partially denaturing environments. Precipitation achieves this goal much more efficiently than dialysis, adsorption of the detergent onto BioBeads or cyclodextrins, or exchange of surfactants after immobilization of the protein onto an affinity chromatography column.

It is truly remarkable that APols favored the folding of two families of IMPs with completely different structures as OMPs and BR, suggesting that the approach could be quite

general. This led to testing it on GPCRs recovered under denaturing conditions (namely in the presence of SDS) from *E. coli* IBs [42]. Conditions initially established to refold BR were applied essentially without any changes to folding six GPCRs, namely the leukotriene B_4 receptors BLT1 and BLT2, the serotonin receptor 5-HT_{4A}, the cannabinoid CB1 receptor [42] and, more recently, the ghrelin GHSR-1a and the vasopressin V2 receptors (J.-L. Banères and B. Mouillac, unpublished). Folding yields between 30 and 50% were systematically achieved, depending on the receptor considered (these determinations were based on ligand binding studies). They rose up to 60-70% in the presence of lipids (**Table 2**). It has been observed that the presence of lipids increases the stability of APol-trapped GPCRs [42]. One possibility is that they do so by binding to sites that form when the transmembrane surface achieves its native state. Thereby, they would contribute to driving folding towards the latter. As observed for most APol-trapped MPs, GPCRs folded in A8-35 are significantly more stable than those kept in lipid/detergent mixtures [42], which is of great interest for subsequent investigations [55]. The BLT1 and GHSR-1a receptors have also been folded in NAPols, with yields similar to those achieved in A8-35 (J.-L. Banères, unpublished). Although less thoroughly studied than A8-35, NAPols can be of interest when the purified proteins under study must be handled or studied at acidic pH, when their ligands tend to interact with polyanions such as A8-35, as is the case with ghrelin and vasopressin, or when studying the kinetics of interaction of G proteins with activated GPCRs, which is slowed down in the presence of A8-35 (J.-L. Banères, unpublished).

From a fundamental point of view, we note that seven α-helical IMPs (BR and six GPCRs) have now been successfully refolded into a synthetic polymer, APol A8-35, whose chemical structure and supramolecular organization bear no similarity, beyond the amphiphilic character, to lipid bilayers. This indicates that, at least for these proteins, neither an environment mimicking the highly complex and anisotropic lipid bilayer nor even the

presence of lipids is required for transmembrane helices to form and correctly orient and pack with each other, and for the polypeptide to adopt its functional 3D structure. This is consistent with the general notion that all of the chemical information needed for proteins, including IMPs, to correctly fold is stored in their sequences.

From a more practical perspective, conditions initially established to refold BR have been applied without much change to refolding of six distinct GPCRs, with functional yields between 30 and 70%. Should this approach turn out to be sufficiently general and easy to implement, as these data suggest, it would represent an important breakthrough for *in vitro* studies aimed at understanding the molecular bases of the function of rhodopsin-like GPCRs and, possibly, of many other IMPs. It is to be noted in this context that all of those IMPs that have been refolded to date using APols, although they display different length (for instance 262 amino acids for BR *versus* 472 amino acids for CB1), have relatively simple structures. In particular, none of them displays extended, complex extramembrane domains like GPCRs from classes B (secretin receptor-like) and C (glutamate receptor-like) or from adhesion and frizzled families. It is currently an open question whether APols would favor or interfere with the folding of such large N-terminal extracellular structures. It would be of interest to express and fold some reference GPCRs from classes B or C following the α_5I-amphipol expression-folding strategy described above to assess whether the presence of their large soluble N-termini may influence expression and functional folding. As indicated before, the class B GPCR GLP-1R has been accumulated in bacterial IBs and functionally folded by transfer from SDS to Brij78 in the presence of methyl-β-cyclodextrin [35]. However the N-terminal domain of GLP-1R is rather short (predicted to be 122 amino acids). In addition, trials to overexpress metabotropic glutamate receptors from the GPCR class C in *E. coli* IBs were not successful [27], but the α_5I fusion strategy has not been applied to these targets so far. Scrambling of disulfide bridges has not been a problem until now, but will undoubtedly be in

some cases. It may possibly be alleviated by careful control of the redox potential during and after folding, and/or by genetic engineering.

What to do next?

Once a GPCR has been folded using APols, it can be studied in this environment, or transferred to another one. Most biochemical and biophysical techniques can be applied to APol-trapped MPs (**Table 3**) [18]. There are reasons to believe that studying protein-protein interactions like oligomerization of GPCRs or recruitment of non-membrane protein partners from signaling complexes can be performed in APols. First, the binding of large soluble toxins and of antibodies to APol-trapped IMPs has already been described [40,56], and these data are promising considering GPCRs and their signaling proteins. Second, GPCRs expressed in IBs and subsequently folded in detergent:lipid mixed micelles can be isolated as dimers that can be used for characterizing molecular events that occur upon activation [29, 57,58]. In the same way, trapping with APols does not prevent GPCRs from assembling into dimers [42]. Purified GPCRs can also interact functionally with signaling proteins. This is true for G proteins with receptors folded in mixed micelles [29,57,58] but also applies to receptors folded in APols. Indeed, both G proteins (Gs/Gq) and arrestins can bind to APols-trapped vasopressin V2 and ghrelin GHSR-1a GPCRs, respectively (J.-L. Banères and B. Mouillac, unpublished). Cryo-electron microscopy can be applied to APol-trapped IMP complexes [59] and supercomplexes (T. Althoff, PhD thesis, University of Frankfurt-am-Main, 2011), whose structure can then be solved by single-particle image analysis. This approach could conceivably be applied to studying the arrangement of GPCRs associated into dimers (oligomers) and/or interacting with their associated signaling proteins. Since complexation by APols is compatible with ligand binding studies [18,42,55,56,60], trapping

with an appropriately functionalized APol would provide a straightforward and very general approach to immobilizing GPCRs onto solid supports for ligand screening [56].

There are cases where one will wish to transfer the folded GPCR to another environment. Direct transfer of IMPs from APols to lipid vesicles or black films has been demonstrated [41,61]. The procedure, however, is unlikely to be applicable to GPCRs, which are fragile proteins and will stand great risks of being denatured in the process. A likely safer route would be to exchange the APol for detergent or lipid/detergent mixed micelles, which is readily possible [50,51,62], and then to proceed to a classical reconstitution.

Crystallization of APol-trapped IMPs is still in its infancy [18]. Crystallization of a GPCR that has been folded using APols should probably best be attempted after transferring the receptor either to a detergent solution, to bicelles, to a lipidic cubic phase or sponge phase. NMR, on the other hand, appears as a particularly promising route to studying the structure of ligands bound to APol-trapped GPCRs, ligand-induced conformational transitions and, possibly, at least some aspects of GPCR structure. Several small IMPs in complexes with APols have been studied to date by solution NMR (T. Dahmane, PhD thesis, University of Paris 7, 2007; P. Bazzacco, PhD thesis, Université of Paris 7, 2009) [53,54,63,64]. GPCR/APol complexes, because of their large size, remain difficult to study *in toto*. However, transferred nuclear Overhauser effects (trNOEs) can be exploited to determine the structure of GPCR-bound ligands. In a recent study, deuterated BLT2 receptor was folded and stabilized using a partially deuterated version of the amphipol A8-35. One of its ligands, the leukotriene LTB_4, was then added in its hydrogenated form, and its receptor-bound structure determined from the magnitude of 89 trNOE signals [55]. The range of NMR studies applicable to APol-trapped GPCRs ought to be extended by developments in the chemistry of APols, such as the availability of a perdeuterated version of A8-35 (F. Giusti, unpublished) or of pH-insensitive APols such as SAPols (**Figure 3B**; T. Dahmane, PhD thesis, Université

Paris 7, 2007) [47] and non-ionic NAPols (**Figure 3C**; P. Bazzacco, PhD thesis, Université Paris 7, 2009) [48,49].

Acknowledgements

We would like to thank Drs. T. Dahmane and L. Catoire for their critical reading of the manuscript.

References

1. Lagerstrom, M.C. and Schioth, H.B (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov*. 7, 339-357
2. Hofmann, K.P. *et al*. (2009) A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci*. 34, 540-552
3. Rosenbaum, D.M. *et al*. (2009) The structure and function of G protein-coupled receptors. *Nature* 459, 356-363
4. Warne, T. *et al*. (2008) Structure of a β_1-adrenergic G protein-coupled receptor. *Nature* 454, 486-491
5. Rasmussen, S.G. *et al*. (2007) Crystal structure of the human β_2-adrenergic G protein-coupled receptor. *Nature* 450, 383-387
6. Jaakola, V.P. *et al*. (2008) The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* 322, 1211-1217
7. Wu, B. *et al*. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* DOI: 10.1126/science.1194396 (http://www.sciencemag.org)

8. Chien, E.Y.T. *et al*. (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* DOI: 10.1126/science.1197410 (http://www.science.org)

9. Klammt, C. *et al*. (2007) Cell-free production of G protein-coupled receptors for functional and structural studies. *J. Struct. Biol*. 158, 482-493

10. Reeves, P.J. *et al*. (2002) Structure and function in rhodopsin : a tetracycline inducible system in stable mammalian cell lines for high-level expression of opsin mutants. *Proc. Natl. Acad. Sci. U.S.A*. 99, 13413-13418

11. Mouillac, B. *et al*. (1992) Agonist-modulated palmitoylation of β_2-adrenergic receptor in Sf9 cells. *J. Biol. Chem*. 267, 21733-21737

12. Panneels, V. and Sinning, I. (2010) Membrane protein expression in the eyes of transgenic flies. *Methods Mol. Biol*. 601, 135-147

13. Zhang, L. *et al*. (2005) Expression of functional G protein-coupled receptors in photoreceptors of transgenic Xenopus laevis. *Biochemistry* 44, 14509-14518

14. Li, N. *et al*. (2007) Heterologous expression of the adenosine A1 receptor in transgenic mouse retina. *Biochemistry* 46, 8350-8359

15. Sarramegna, V. *et al*. (2005) Solubilization, purification and mass spectrometry analysis of the human mu-opioid receptor expressed in Pichia pastoris. *Protein Expr. Purif*. 43, 85-93

16. Lundstrom, K. (2005) Structural genomics of GPCRs. *Trends Biotechnol*. 23, 103-108

17. Bokoch, M.P. *et al*. (2010) Ligand-induced regulation of the extracellular surface of a G protein-coupled receptor. *Nature* 463, 108-112

18. Popot, J-L. (2010) Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu. Rev. Biochem*. 79, 13.1-13.39

19. Marullo, S. *et al.* (1988) Human β2-adrenergic receptors expressed in *Escherichia coli* membranes retain their pharmacological properties. *Proc. Natl. Acad. Sci. U.S.A.* 85, 7551-7555

20. Tucker, J. and Grisshammer, R. (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem. J*. 317, 891-899

21. Weiss, H.M. and Grisshammer, R. (2002) Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur. J. Biochem*. 269, 82-92

22. Skretas, G. and Georgiou, G. (2009) genetic analysis of G protein-coupled receptor expression in *Escherichia coli*. *Biotechnol. Bioeng*. 102, 357-367

23. White, J.F. *et al*. (2004) Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett*. 564, 289-293

24. Kiefer, H. *et al*. (1996) Expression of an olfactory receptor in Escherichia coli: purification, reconstitution and ligand binding. *Biochemistry* 35, 16077-16084

25. Banères, J-L. *et al*. (2003) Structure-based analysis of GPCR function: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. *J. Mol. Biol*. 329, 801-814

26. Banères, J-L. *et al*. (2005) Molecular characterization of a purified 5-HT4 receptor. *J. Biol. Chem*. 280, 20253-20260

27. Michalke, K. *et al*. (2009) Mammalian G protein-coupled receptor expression in *Escherichia coli*: I. High throughput large-scale production as inclusion bodies. *Anal. Biochem*. 386, 147-155

28. Michalke, K. *et al*. (2009) 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. *Anal. Biochem*. 401, 74-80

29. Arcemisbéhère, L. *et al*. (2010) Leukotriene BLT2 receptor monomers activate the G_{i2} GTP-binding protein more efficiently than dimers. J. Biol. Chem. 285, 6337-6347

30. Mouillac, B. and Banères, J-L. (2010) Mammalian membrane receptors expression as inclusion bodies in *Escherichia coli*. Methods Mol. Biol. 601, 39-48

31. Kiefer, H. (2003) In vitro folding of alpha-helical membrane proteins. *Biochim. Biophys. Acta*. 1610, 57-62

32. Huang, K.-S. et al. (1981) Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. *J. Biol. Chem.* 256, 3802-3809

33. Muller, I. *et al.* (2008) The full-length mu-opioid receptor: a conformational study by circular dichroism in trifluoroethanol and membrane-mimetic environments. *J. Membr. Biol.* 223, 49-57

34. Miller, D. *et al*. (2009) In vitro unfolding and refolding of the small multidrug transporter EmrE. *J. Mol. Biol*. 393, 815-832

35. Schröder-Tittmann, K. *et al.* (2010) Recombinant expression, *in vitro* refolding, and biophysical characterization of the human glucagon-like peptide-1 receptor. *Biochemistry* 49, 7956-7965

36. Popot, J-L. *et al*. (1987) Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process. *J. Mol. Biol*. 198, 655-676

37. Schmidt, P. *et al*. (2010) A reconstitution protocol for the in vitro folded human G protein-coupled Y2 receptor into lipid environment. *Biophys. Chem*. 150, 29-36

38. Schimmer, S. *et al.* (2010) Functional characterization of the in vitro folded human y(1) receptor in lipid environment. *Protein Pept. Lett*. 17, 605-609

39. Tribet, C. *et al*. (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc. Natl. Acad. Sci. USA* 93, 15047-15050

458 40. Popot, J-L. *et al.* (2003) Amphipols: polymeric surfactants for membrane biology
459 research. *Cell. Mol. Life Sci*. 60, 1559-1574
460 41. Pocanschi, C.L. *et al.* (2006) Amphipathic polymers: tools to fold integral membrane
461 proteins to their active form. *Biochemistry* 45, 13954-13961
462 42. Dahmane, T. *et al.* (2009) Amphipol-assisted *in vitro* folding of G protein-coupled
463 receptors. *Biochemistry* 48, 6516-6521
464 43. Gohon, Y. and Popot, J-L. (2003) Membrane protein-surfactant complexes. *Curr.*
465 *Opin. Colloid Interface Sci.* 8, 15-22
466 44. Sanders, C.R. *et al*. (2004) French swimwear for membrane proteins. *ChemBioChem*
467 5, 423-426
468 45. Gohon, Y. *et al.* (2004) Partial specific volume and solvent interactions of amphipol
469 A8-35. *Anal. Biochem.* 334, 318-334
470 46. Gohon, Y. *et al.* (2006) Well-defined nanoparticles formed by hydrophobic assembly
471 of a short and polydisperse random terpolymer, amphipol A8-35. *Langmuir* 22, 1281-
472 1290
473 47. Picard, M. *et al*. (2006) Protective and inhibitory effects of various types of amphipols
474 on the Ca^{2+}-ATPase from sarcoplasmic reticulum: a comparative study. *Biochemistry*
475 45, 1861-1869
476 48. Sharma, K.S. *et al*. (2008) Glucose-based amphiphilic telomers designed to keep
477 membrane proteins soluble in aqueous solutions: synthesis and physicochemical
478 characterization. *Langmuir* 24, 13581-13590
479 49. Bazzacco, P. *et al.* (2009) Trapping and stabilization of integral membrane proteins by
480 hydrophobically grafted glucose-based telomers. *Biomacromolecules* 10, 3317-3326

481 50. Zoonens, M. *et al*. (2007) Dynamics of membrane protein/amphipol association
482 studied by Förster resonance energy transfer. Implications for *in vitro* studies of
483 amphipol-stabilized membrane proteins. *Biochemistry* 46, 10392-10404
484 51. Tribet, C. *et al*. (2009) Thermodynamic characterization of the exchange of detergents
485 and amphipols at the surfaces of integral membrane proteins. *Langmuir* 25, 12623-
486 12634
487 52. Gohon, Y. *et al*. (2008) Bacteriorhodopsin/amphipol complexes: structural and
488 functional properties. *Biophys. J*. 94, 3523-3537
489 53. Zoonens, M. *et al*. (2005) NMR study of a membrane protein in detergent-free
490 aqueous solution. *Proc. Natl. Acad. Sci. USA* 102, 8893-8898
491 54. Catoire, L.J. *et al*. (2009) Inter- and intramolecular contacts in a membrane
492 protein/surfactant complex observed by heteronuclear dipole-to-dipole cross-
493 relaxation. *J. Magn. Res.* 197, 91-95
494 55. Catoire, L.J. *et al*. (2010) Structure of a GPCR ligand in its receptor-bound state:
495 leukotriene B4 adopts a highly constrained conformation when associated to human
496 BLT2. *J. Am. Chem. Soc*. 132, 9049-9057
497 56. Charvolin, D. *et al*. (2009) The use of amphipols as universal molecular adapters to
498 immobilize membrane proteins onto solid supports. *Proc. Natl. Acad. Sci. USA* 106,
499 405-410
500 57. Damian, M. *et al*. (2006) Asymmetric conformational changes in a GPCR dimer
501 controlled by G proteins. *EMBO J*. 25, 5693-5702
502 58. Damian, M. *et al*. (2008) G protein activation by the leukotriene B4 receptor dimer.
503 Evidence for an absence of trans-activation. *J. Biol. Chem*. 283, 21084-21092

59. Flötenmeyer, M. *et al.* (2007) The use of amphipathic polymers for cryo-electron microscopy of NADH:ubiquinone oxidoreductase (Complex I). *J. Microsc.* 227, 229-235

60. Martinez, K. L. *et al.* (2002). Allosteric transitions of *Torpedo* acetylcholine receptor in lipids, detergent and amphipols: molecular interactions *vs.* physical constraints. *FEBS Lett.* 528, 251-256

61. Nagy, J. K. *et al*. (2001) Use of amphipathic polymers to deliver a membrane protein to lipid bilayers. *FEBS Lett.* 501, 115-120

62. Tribet, C. *et al*. (1997) Stabilisation of hydrophobic colloidal dispersions in water with amphiphilic polymers: application to integral membrane proteins. *Langmuir* 13, 5570-5576

63. Catoire, L. J. *et al*. (2010) Solution NMR mapping of water-accessible residues in the transmembrane β-barrel of OmpX. *Eur. Biophys. J.* 39, 623-630

64. Raschle, T. *et al.* (2010) Nonmicellar systems for solution NMR spectroscopy of membrane proteins. *Curr. Opin. Struct. Biol*. 20, 1-9

65. Gorzelle, B.M. *et al.* (2002) Amphipols can support the activity of a membrane enzyme. *J. Am. Chem. Soc.* 124, 11594-11595

66. Tribet, C. *et al*. (1998) Scanning transmission electron microscopy study of the molecular mass of amphipol/cytochrome b_6f complexes. *Biochimie* 80, 475-482

Table 1. Level of production of several purified GPCRs following expression in *E. coli*

GPCR[a]	Fusion partner(s)	Site of expression	Yield (mg/L)[b]	Refs.
Neurotensin NTS1	MBP + TRX	Inner membrane	0.13[c]	[20]
Adenosine A2a	MBP	Inner membrane	0.17	[21]
Olfactory OR5	GST	IBs	0.2	[24]
Leukotriene BLT1	T7 tag	IBs	2-3	[25]
Serotonin $5HT_{4A}$	KSI	IBs	0.5	[26]
Cannabinoid CB1	none	IBs	100[c]	[28]
Vasopressin V2, leukotriene BLT2	α_5I	IBs	0.8-1.2	[29]

[a]For each GPCR, the fusion partner(s) and the site of expression in the bacterium, namely the inner membrane or inclusion bodies (IBs), are indicated.

[b]In each case, the quantity of pure and functional GPCR that has been obtained is indicated in mg/L of cell suspension. A functional receptor means that it has been overexpressed, separated form its fusion partner (if applicable) by enzyme cleavage, folded and/or purified and its activity checked by ligand binding.

[c]Except for CB1 and NTS1, which were produced in fermentors, all other receptors have been overexpressed in culture flasks.

Table 2. Comparison of GPCR folding yields obtained in different surfactants

GPCR	Folding medium	Average maximum folding yield (%)[a]	Refs.
BLT1	Detergent/lipid mixed micelles (LDAO/asolectin)	30	[25]
	A8-35	50	[42]
	A8-35/asolectin	65	[42]
BLT2	Detergent/lipid mixed micelles (DPC/HDM/asolectin)	4	[29]
	A8-35	50	[42]
	A8-35/asolectin	70	[42]
5-HT_{4A}	DMPC/CHAPS bicelles	25	[26]
	A8-35	30	[42]
	A8-35/asolectin	60	[42]
CB1	Detergent/lipid mixed micelles[b] (Fos-choline-16/asolectin)	0	[42]
	A8-35	30	[42]
	A8-35/asolectin	40	[42]
	Detergent mixed micelles (DDM/Cymal 6)	30	[28]

[a]The folding yield is the amount of functional protein obtained after folding (based on binding of specific ligands) compared to that of protein in starting SDS solution (based on A_{280} measurements).

[b]No extensive detergent screening was carried out to optimize folding of the CB1 receptor in detergent/lipid mixed micelles.

607 **Table 3. Applicability of various approaches to handling and studying GPCR/APol**
608 **complexes**
609

Approach	Applicability	Remarks	Refs.
UV-visible absorption and fluorescence spectroscopy, CD	Yes, but for infra-red studies in the amide absorption bands.	All APols validated to date contain amide bonds.	[41,50,52]
Ligand-binding and functional studies	Yes.	Damping of large-scale transmembrane conformational changes may occur. NAPols to be favored for G protein and arrestin binding studies, as well as for binding studies with cationic amphipathic ligands.	[42,47,52,55, 56,60,65]
Purification	Most purification techniques: ultracentrifugation, size exclusion chromatography, immobilized metal and ligand-based affinity chromatography…	With charged APols, ionic exchange chromatography and isoelectric focusing to be avoided.	[42,50,52]
Electron microscopy, atomic force microscopy	Single particles studied by EM after negative staining and by cryoEM.	No AFM studies reported yet.	[18,52,59,66]
Immobilization onto solid supports	Yes.	Immobilization can be either direct or mediated by an appropriately functionalized APol.	[40,56]
NMR	Yes.	Tested with A8-35, SAPols and NAPols, the latter two giving access to lower pH. Only A8-35 has been deuteriated yet. No solid-state studies reported yet.	[53-55,63]
Mass spectrometry	Yes.	Study in progress.	[54]
X-ray crystallography	Remains to be developed.	Transfer to detergent or lipid cubic phase.	[18]

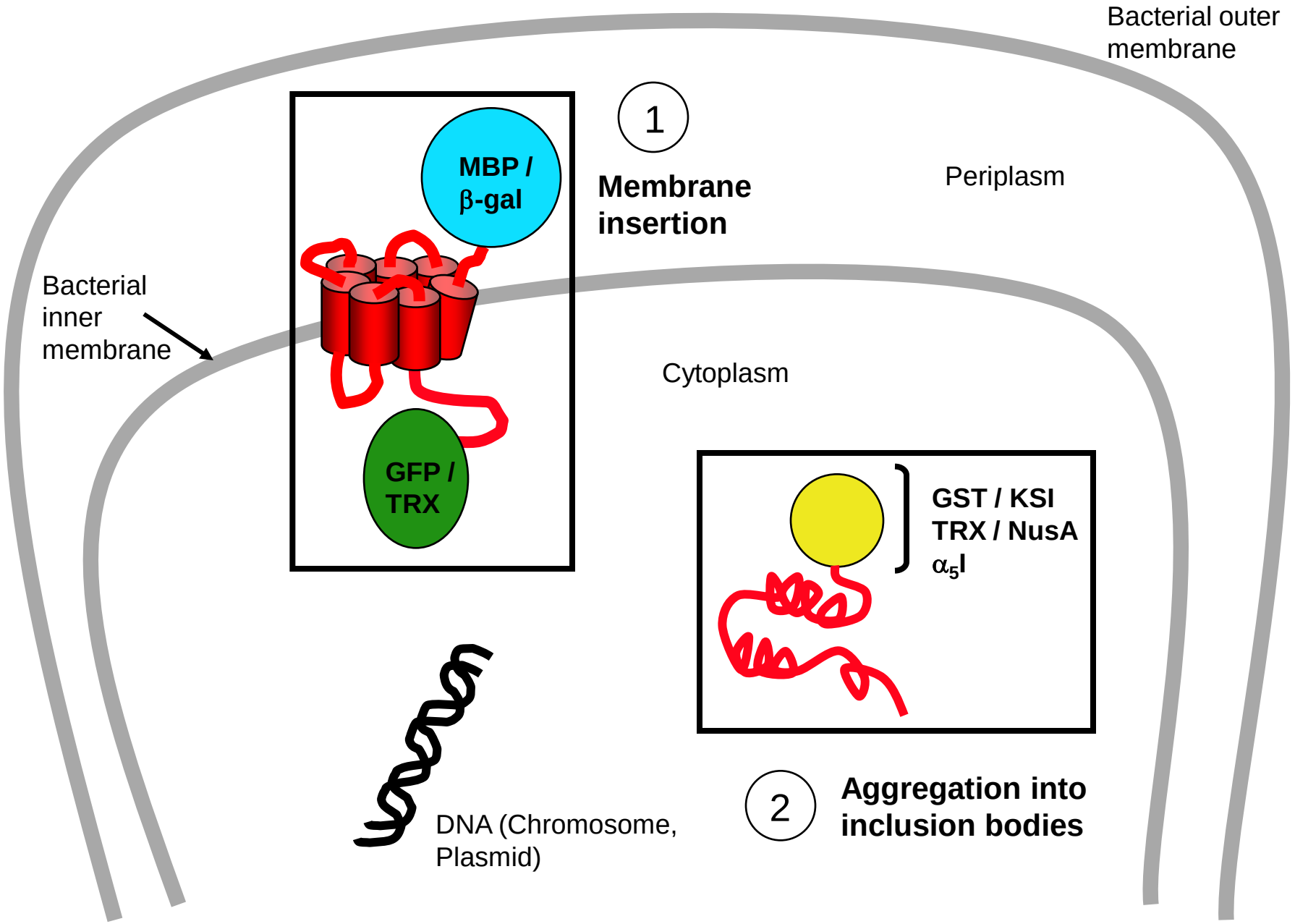
610
611

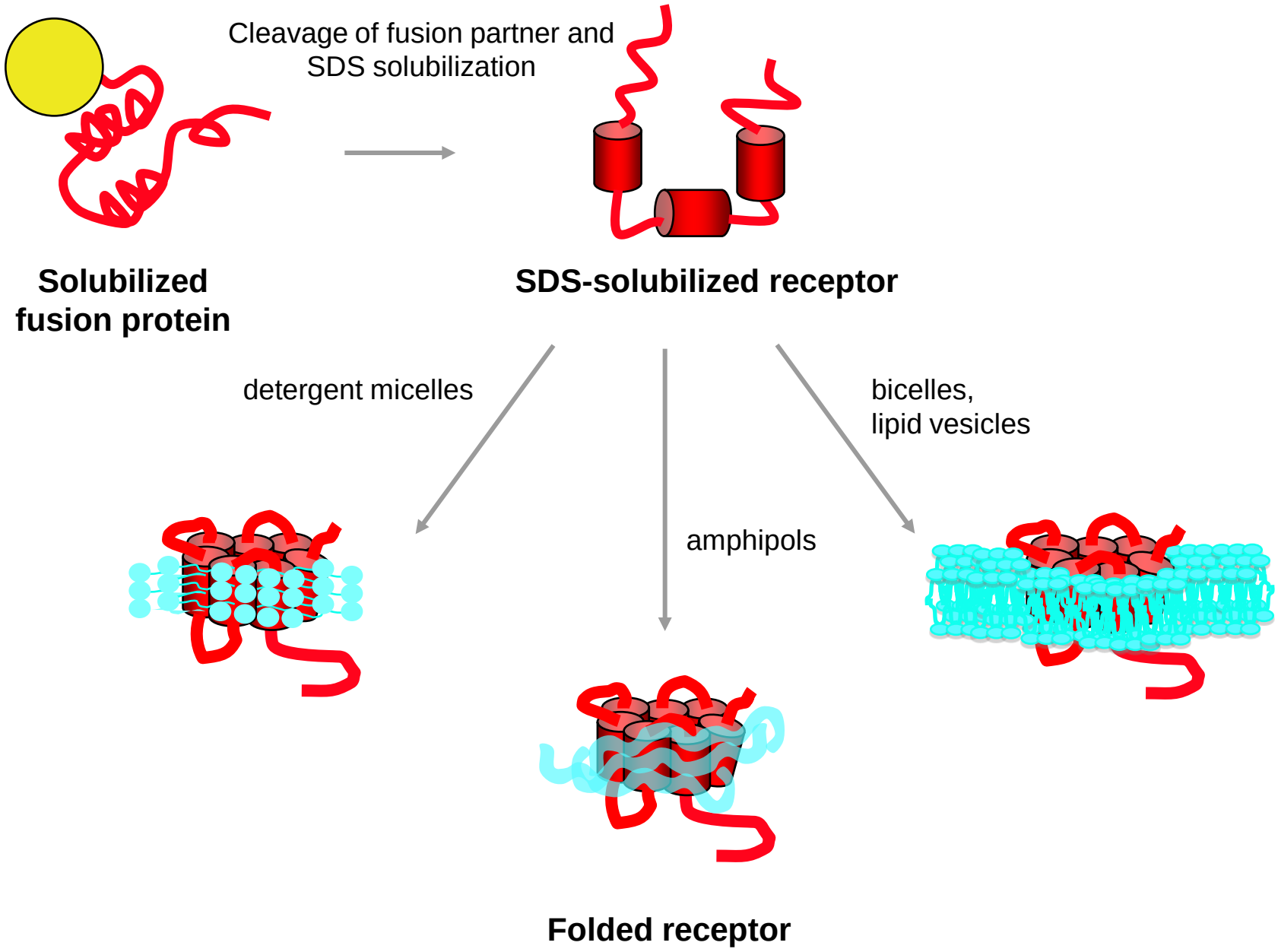
Figure legends

Figure 1. Strategies to overexpress GPCRs in *E. coli*. GPCRs can be produced in bacteria either by insertion into the inner membrane ① or by accumulation into inclusion bodies ②. In the first case, a protein partner is coupled at the N-terminus (MBP, β-gal) or C-terminus (GFP, TRX) of the receptor, or at both extremities (for instance MBP and TRX), so as to target the recombinant protein to the membrane. In the second case, targeting to IBs is favored by coupling to the N-terminus another fusion partner, such as GST, KSI, TRX, NusA or α_5I. GPCRs accumulated as inclusion bodies are not adequately folded and have to be solubilized in a harsh detergent before folding.

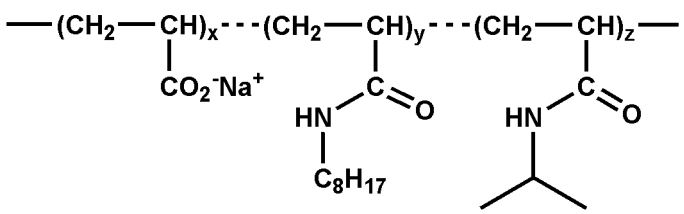
Figure 2. Strategies to fold GPCRs from IBs. Before folding, the recombinant GPCR accumulated in IBs has to be solubilized in denaturing buffers (for instance a mix of urea and SDS), purified by immobilized-metal affinity chromatography (IMAC) and dialyzed in an aqueous buffer to eliminate most SDS and allow removal of the fusion partner, for instance via a thrombin cleavage. After elimination of the cleaved fusion partner through a second IMAC purification in denaturing conditions, the GPCR is kept soluble in SDS solution in which it displays a significant content of helicity. Folding can be achieved by transfer from SDS to mild detergents or amphipols, supplemented or not with lipids. The functional fraction of the GPCR preparation can be evaluated using pharmacological assays (ligand binding, activation of purified G proteins or arrestins).

Figure 3. Chemical structures of three families of APols. A) A8-35, the prototypal APol. B) Sulfonated APol (SAPol). C) Glucose-based non-ionic APol (NAPol).

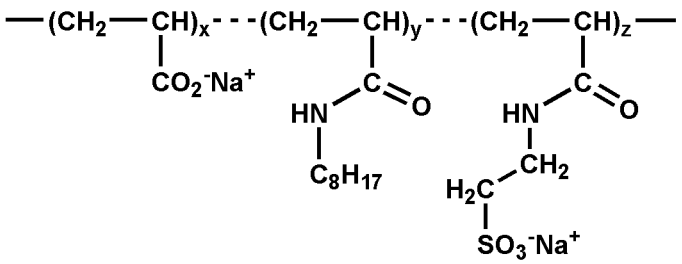

Bacterial outer membrane
MBP / β-gal
1
Membrane insertion
Periplasm
Bacterial inner membrane
Cytoplasm
GFP / TRX
GST / KSI
TRX / NusA
α5I
DNA (Chromosome, Plasmid)
2
Aggregation into inclusion bodies


Cleavage of fusion partner and
SDS solubilization
Solubilized
fusion protein
SDS-solubilized receptor
detergent micelles
amphipols
bicelles,
lipid vesicles
Folded receptor

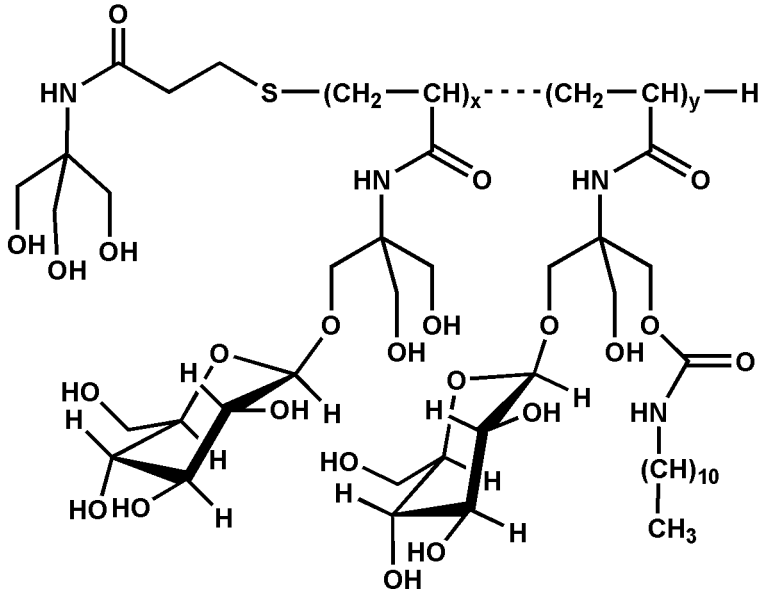
A


(CH2—CH)x
(CH2—CH)y
(CH2—CH)z
CO2-Na+
HN
C
O
C8H17
HN
C
O

B


(CH2—CH)x
(CH2—CH)y
(CH2—CH)z
CO2-Na+
HN
C
O
C8H17
HN
C
O
CH2
H2C
SO3-Na+

C


O
HN
S
(CH2—CH)x
(CH2—CH)y
H
HN
O
HN
O
OH
OH
OH
O
OH
O
OH
O
O
HO
H
O
H
OH
H
H
HO
HO
H
O
H
H
OH
HO
H
H
HO
OH
HN
(CH)10
CH3