

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

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

► **To cite this version:**

Jean-Louis Banères, Jean-Luc Popot, Bernard Mouillac. New advances in production and functional folding of G-protein-coupled receptors.. Trends in Biotechnology, Elsevier, 2011, 29 (7), pp.314-22. <10.1016/j.tibtech.2011.03.002>. <inserm-00590416>

HAL Id: inserm-00590416

<http://www.hal.inserm.fr/inserm-00590416>

Submitted on 3 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

2

3 Jean-Louis Banères^{1,2}, Jean-Luc Popot^{3,4} and Bernard Mouillac⁵⁻⁷

4

5 ¹CNRS, UMR-5247, Institut des Biomolécules Max Mousseron, Faculté de Pharmacie, F-
6 34000 Montpellier, France

7

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

9

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

12

13 ⁴Université Paris-7, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, F-
14 75000 Paris, France

15

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

18

19 ⁶INSERM, U661, F-34000 Montpellier, France

20

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

22

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

24

25 **Abstract**

26 G protein-coupled receptors (GPCRs), the largest family of integral membrane proteins,
27 participate in the regulation of many physiological functions and are the targets of around
28 30% of currently marketed drugs. However, knowledge of the structural and molecular bases
29 of GPCR function remains limited, owing to difficulties related to their overexpression,
30 purification and stabilization. The development of new strategies aimed at obtaining large
31 amounts of functional GPCRs is therefore crucial. Here, we review the most recent advances
32 in production and functional folding of GPCRs from *Escherichia coli* inclusion bodies. Major
33 breakthroughs open exciting perspectives for structural and dynamic investigations of
34 GPCRs. In particular, combining targeting to bacterial inclusion bodies with amphipol-
35 assisted folding is emerging as a highly powerful strategy.

36

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

39 Structural information on integral membrane proteins (IMPs) remains limited.
40 According to databases of known three-dimensional (3D) structures of IMPs
41 (http://blanco.biomol.uci.edu/Membrane_proteins_xtal.html; <http://www.mpdb.tcd.ie>;
42 <http://www.drorlist.com/nmr/MPNMR.html>), the crystal or NMR structures of only ~280
43 IMPs have been solved to date, as compared to tens of thousands of soluble proteins. Worse,
44 only a small fraction (~35%) of these proteins are eukaryotic, even though IMPs constitute
45 20-30% of the proteins encoded by human and other eukaryotic genomes. Among IMPs, G
46 protein-coupled receptors (GPCRs), whose transmembrane region is a 7-helix bundle,
47 constitute the largest family [1]. More than 800 GPCRs have been identified, representing 2-
48 3% of the coding sequences in the human genome. They are involved in most essential
49 cellular processes and are the targets of around 30% of current pharmaceutical drugs. It is
50 therefore critical to gain detailed knowledge of their structures and their dynamics in order to
51 understand their functions and/or dysfunctions, as well as to rationally design selective
52 therapeutic compounds. However, except for rhodopsin, whose crystal structure has been
53 solved following its extraction from the retina [2,3], the low natural abundance of GPCRs
54 generally precludes their purification in biochemically relevant amounts. Overexpression is
55 thus a prerequisite to investigating their structure or analyzing their conformational transitions
56 upon interaction with ligands or with signaling proteins like G proteins and arrestins.

57 Overexpressing GPCRs, however, is still problematic, often resulting in low yield,
58 protein aggregation or misfolding, if not cell toxicity. Consequently, although crystal
59 structures of ligand-bound β_1 - and β_2 -adrenergic [4,5], adenosine A_{2A} [6], chemokine CXCR4
60 [7], and dopamine D3 [8] receptors have been recently reported, investigating the structure
61 and dynamics of most GPCRs remains a daunting task. Many overexpression systems have

62 been tested. GPCRs have been successfully produced by cell-free synthesis [9] and by
63 heterologous expression in mammalian [10] and insect [11] cells, in the photoreceptor cells of
64 *Drosophila* [12], *Xenopus* [13] and mouse [14], and in such microbes as yeasts and bacteria
65 [15,16]. As discussed below, expression in *Escherichia coli* holds great promises, not only
66 due to its simplicity, rapidity, safety, scalability or genetic tractability, but also in terms of
67 quantity and homogeneity of the recombinant protein.

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

85

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

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

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

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

127 The use as targeting partner of an α_5 integrin fragment (α_5 I, 31 kDa) has allowed
128 many rhodopsin-like GPCRs to be expressed at high levels regardless of their length (from
129 337 to 472 amino acids), their G protein coupling selectivity, or the nature of their
130 endogenous ligands. This efficient and apparently generic procedure has been successfully
131 applied to expressing the β_3 -adrenergic receptor, the vasopressin V2 and V1b and oxytocin
132 OTR receptors, the chemokine CCR5 and CXCR4 and chemokine-like ChemR23 receptors,
133 the ghrelin GHS-1a receptor, the cannabinoid CB1 receptor and the leukotriene BLT1, BLT2,
134 CysLT1 and CysLT2 receptors, without requiring any optimization of either the GPCR
135 coding sequence, the cell culture conditions, or the extraction/purification procedures [29,30].

136 The α_5 I fusion strategy represents an important breakthrough for *in vitro* studies aimed at
137 understanding the molecular bases of GPCR function and structure, and potentially for other
138 membrane protein families.

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

142

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

144 Following expression, IBs are first solubilized under denaturing conditions. Then,
145 following purification (using a metal-affinity chromatography procedure for instance), the
146 fusion partner used for high-level expression has to be removed, usually through proteolytic
147 cleavage. A notable advantage of α_5 I-GPCR fusions regarding this crucial biochemical step is
148 that α_5 I keeps the receptor soluble after dialysis in the absence of denaturing agents, greatly
149 facilitating an efficient proteolytic cleavage of the fusion protein [29]. After another
150 purification step in SDS buffers, folding is then initiated by transfer from SDS to other
151 surfactants (**Figure 2**). Folding efficiency depends on the competition between protein
152 aggregation and 3D structure formation as well as on the ability of the receiving surfactant to
153 stabilize the native 3D state of the folded receptor [31].

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

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

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

177 Efficient folding in detergents has been reported for a limited number of GPCRs such
178 as the leukotriene receptors BLT1 and BLT2 (**Table 2**). BLT1 was folded as a functional
179 protein to ~30% in LDAO [25] whereas the BLT2 receptor was folded as a functional state to
180 ~4% in DPC/HDM mixtures [29]. In both cases, adding lipids (e.g. asolectin) was required for
181 improving the percentage of functional recovery. Other GPCRs have been reported to
182 efficiently fold in detergent micelles. As stated above, the OR5 receptor was first folded in
183 digitonin before insertion in lipid vesicles [24]. Fluorescence-monitored ligand binding assays
184 demonstrated that about 80% of the folded OR5 receptor bound its lily ligand [24]. More
185 recently, refolding of the SDS-solubilized parathyroid hormone receptor 1 (PTH-1R) and of

186 CB1 receptors has been performed by exchanging the SDS for a mixture of the non-ionic
187 detergents DDM and Cymal 6 [28]. In this study, ligand binding assays demonstrated that
188 ~30% of the folded CB1 was functional. The glucagon-like peptide-1 receptor (GLP-1R) has
189 been reported to fold upon transfer from SDS to Brij78 as a functional protein to ~40% [35].
190 In all these cases, i.e. for the PTH1R, CB1 and GLP-1R, the folding process was carried out
191 in the presence of methyl- β -cyclodextrin, used to strip off SDS. Removal of dodecylsulfate
192 can also be achieved by precipitation using K^+ ions (see below) [36].

193 The efficiency of alternate membrane-like environments to fold and stabilize GPCRs
194 recovered from IBs has also been explored. Certain mixtures of long-chain and short-chain
195 phospholipids assemble as bilayer discs, called bicelles, which mimic the membrane
196 environment (**Figure 2**). A limited number of GPCRs, specifically the serotonin 5-HT_{4A} [26]
197 and the neuropeptide Y₂ receptors [37], have been folded to a native-like conformation in
198 DMPC/CHAPS bicelles with folding yields of ~25 % and ~65%, respectively.

199 A few cases of successful folding in lipid vesicles of GPCRs recovered from IBs have
200 also been reported. The first such example was described in the pioneering work of Kiefer's
201 group on the OR5 olfactory receptor, in which the overexpressed receptor was solubilized in
202 the strong, negatively charged detergent *N*-lauroylsarcosine and then folded by transfer to the
203 non-denaturing detergent digitonin [24]. The digitonin-folded receptor was able to bind its
204 ligand (see above), thus providing yet another example of successful folding in detergent. The
205 OR5 receptor was subsequently reconstituted in lipid vesicles by supplementing it with
206 DDM/POPC/POPG mixtures before removing the detergent by adsorption onto hydrophobic
207 beads. Under such conditions, the protein was stabilized in a fully ligand-competent state
208 (~1% of the solubilized and purified material), as assessed by photoaffinity labeling.

209 As in the case of bicelles, successful folding of GPCRs by direct transfer to lipids is
210 limited to a few examples, such as the human neuropeptide Y₁ receptor [38].

211

212 **Amphipol-assisted folding of GPCRs: a generic approach?**

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

219 APols were initially designed, synthesized and validated as mild alternatives to classi-
220 cal detergents [39]. They are defined as “amphipathic polymers that are able to keep indivi-
221 dual IMPs soluble under the form of small complexes” [18]. What is of interest here is that :
222 (i) most IMPs are more (and generally much more) stable in APols than they are in detergent
223 solutions [18,40], and (ii) APols have proven an efficient medium in which to fold IMPs to
224 their native state [41,42]. APols are relatively short polymers (their mass is typically in the 8-
225 20-kDa range) that carry a high density both of hydrophobic chains and of highly hydrophilic
226 groups. The prototypal APol, named A8-35 (**Figure 3A**) [40,41], remains by far the most
227 thoroughly studied and most widely used APol [18,43,44]. The high solubility of A8-35 in
228 water is due to the presence of carboxylates. As a consequence, A8-35 becomes insoluble in
229 acidic solutions [45,46], a limitation that has prompted the development of alternative
230 chemical structures such as sulfonated APols (SAPols; **Figure 3B**) [47] or glucose-based,
231 non-ionic APols (NAPols; **Figure 3C**) [48,49], both of which are insensitive to pH. In aque-
232 ous solutions, APols form small, micelle-like particles, each of which comprises only a few
233 APol molecules (~4 of them in the case of A8-35) [46].

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

236 formation of ternary complexes [50,51]. The detergent is then removed, yielding small,
237 compact IMP/APol complexes [52], in which the APol forms a thin layer covering the
238 hydrophobic transmembrane surface of the protein [53,54]). Alternatively, a denatured IMP in
239 SDS and/or urea is transferred to APols, during which process it recovers or adopts its native
240 3D structure (**Figure 2**, and see below). As a rule, APol-trapped IMPs are much more stable
241 than their detergent-solubilized counterparts [18,40]. The underlying mechanisms are several.
242 The most important factor is that APols are less efficient than detergents at disrupting the pro-
243 tein/protein and protein/lipid interactions that determine and stabilize the 3D structure of
244 IMPs [40,47]. This led to the suggestion that, in addition to being less aggressive towards
245 properly folded, native IMPs, APols might provide a favorable medium for the formation or
246 reformation of native-like interactions starting from a denatured protein.

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

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

261 general. This led to testing it on GPCRs recovered under denaturing conditions (namely in the
262 presence of SDS) from *E. coli* IBs [42]. Conditions initially established to refold BR were
263 applied essentially without any changes to folding six GPCRs, namely the leukotriene B₄
264 receptors BLT1 and BLT2, the serotonin receptor 5-HT_{4A}, the cannabinoid CB1 receptor [42]
265 and, more recently, the ghrelin GHSR-1a and the vasopressin V2 receptors (J.-L. Banères and
266 B. Mouillac, unpublished). Folding yields between 30 and 50% were systematically achieved,
267 depending on the receptor considered (these determinations were based on ligand binding
268 studies). They rose up to 60-70% in the presence of lipids (**Table 2**). It has been observed that
269 the presence of lipids increases the stability of APol-trapped GPCRs [42]. One possibility is
270 that they do so by binding to sites that form when the transmembrane surface achieves its
271 native state. Thereby, they would contribute to driving folding towards the latter. As observed
272 for most APol-trapped MPs, GPCRs folded in A8-35 are significantly more stable than those
273 kept in lipid/detergent mixtures [42], which is of great interest for subsequent investigations
274 [55]. The BLT1 and GHSR-1a receptors have also been folded in NAPols, with yields similar
275 to those achieved in A8-35 (J.-L. Banères, unpublished). Although less thoroughly studied
276 than A8-35, NAPols can be of interest when the purified proteins under study must be
277 handled or studied at acidic pH, when their ligands tend to interact with polyanions such as
278 A8-35, as is the case with ghrelin and vasopressin, or when studying the kinetics of
279 interaction of G proteins with activated GPCRs, which is slowed down in the presence of
280 A8-35 (J.-L. Banères, unpublished).

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

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

290 From a more practical perspective, conditions initially established to refold BR have
291 been applied without much change to refolding of six distinct GPCRs, with functional yields
292 between 30 and 70%. Should this approach turn out to be sufficiently general and easy to
293 implement, as these data suggest, it would represent an important breakthrough for *in vitro*
294 studies aimed at understanding the molecular bases of the function of rhodopsin-like GPCRs
295 and, possibly, of many other IMPs. It is to be noted in this context that all of those IMPs that
296 have been refolded to date using APols, although they display different length (for instance
297 262 amino acids for BR *versus* 472 amino acids for CB1), have relatively simple structures.
298 In particular, none of them displays extended, complex extramembrane domains like GPCRs
299 from classes B (secretin receptor-like) and C (glutamate receptor-like) or from adhesion and
300 frizzled families. It is currently an open question whether APols would favor or interfere with
301 the folding of such large N-terminal extracellular structures. It would be of interest to express
302 and fold some reference GPCRs from classes B or C following the α_5 I-amphipol expression-
303 folding strategy described above to assess whether the presence of their large soluble N-
304 termini may influence expression and functional folding. As indicated before, the class B
305 GPCR GLP-1R has been accumulated in bacterial IBs and functionally folded by transfer
306 from SDS to Brij78 in the presence of methyl- β -cyclodextrin [35]. However the N-terminal
307 domain of GLP-1R is rather short (predicted to be 122 amino acids). In addition, trials to
308 overexpress metabotropic glutamate receptors from the GPCR class C in *E. coli* IBs were not
309 successful [27], but the α_5 I fusion strategy has not been applied to these targets so far.
310 Scrambling of disulfide bridges has not been a problem until now, but will undoubtedly be in

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

313

314 **What to do next?**

315 Once a GPCR has been folded using APols, it can be studied in this environment, or
316 transferred to another one. Most biochemical and biophysical techniques can be applied to
317 APol-trapped MPs (**Table 3**) [18]. There are reasons to believe that studying protein-protein
318 interactions like oligomerization of GPCRs or recruitment of non-membrane protein partners
319 from signaling complexes can be performed in APols. First, the binding of large soluble
320 toxins and of antibodies to APol-trapped IMPs has already been described [40,56], and these
321 data are promising considering GPCRs and their signaling proteins. Second, GPCRs
322 expressed in IBs and subsequently folded in detergent:lipid mixed micelles can be isolated as
323 dimers that can be used for characterizing molecular events that occur upon activation [29,
324 57,58]. In the same way, trapping with APols does not prevent GPCRs from assembling into
325 dimers [42]. Purified GPCRs can also interact functionally with signaling proteins. This is
326 true for G proteins with receptors folded in mixed micelles [29,57,58] but also applies to
327 receptors folded in APols. Indeed, both G proteins (Gs/Gq) and arrestins can bind to APols-
328 trapped vasopressin V2 and ghrelin GHSR-1a GPCRs, respectively (J.-L. Banères and B.
329 Mouillac, unpublished). Cryo-electron microscopy can be applied to APol-trapped IMP
330 complexes [59] and supercomplexes (T. Althoff, PhD thesis, University of Frankfurt-am-
331 Main, 2011), whose structure can then be solved by single-particle image analysis. This
332 approach could conceivably be applied to studying the arrangement of GPCRs associated into
333 dimers (oligomers) and/or interacting with their associated signaling proteins. Since
334 complexation by APols is compatible with ligand binding studies [18,42,55,56,60], trapping

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

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

343 Crystallization of APol-trapped IMPs is still in its infancy [18]. Crystallization of a
344 GPCR that has been folded using APols should probably best be attempted after transferring
345 the receptor either to a detergent solution, to bicelles, to a lipidic cubic phase or sponge phase.
346 NMR, on the other hand, appears as a particularly promising route to studying the structure of
347 ligands bound to APol-trapped GPCRs, ligand-induced conformational transitions and,
348 possibly, at least some aspects of GPCR structure. Several small IMPs in complexes with
349 APols have been studied to date by solution NMR (T. Dahmane, PhD thesis, University of
350 Paris 7, 2007; P. Bazzacco, PhD thesis, Université of Paris 7, 2009) [53,54,63,64].
351 GPCR/APol complexes, because of their large size, remain difficult to study *in toto*.
352 However, transferred nuclear Overhauser effects (trNOEs) can be exploited to determine the
353 structure of GPCR-bound ligands. In a recent study, deuterated BLT2 receptor was folded and
354 stabilized using a partially deuterated version of the amphipol A8-35. One of its ligands, the
355 leukotriene LTB₄, was then added in its hydrogenated form, and its receptor-bound structure
356 determined from the magnitude of 89 trNOE signals [55]. The range of NMR studies
357 applicable to APol-trapped GPCRs ought to be extended by developments in the chemistry of
358 APols, such as the availability of a perdeuterated version of A8-35 (F. Giusti, unpublished) or
359 of pH-insensitive APols such as SAPols (**Figure 3B**; T. Dahmane, PhD thesis, Université

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

362

363 **Acknowledgements**

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

366

367 **References**

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

- 383 8. Chien, E.Y.T. *et al.* (2010) Structure of the human dopamine D3 receptor in complex
384 with a D2/D3 selective antagonist. *Science* DOI: 10.1126/science.1197410
385 (<http://www.science.org>)
- 386 9. Klammt, C. *et al.* (2007) Cell-free production of G protein-coupled receptors for
387 functional and structural studies. *J. Struct. Biol.* 158, 482-493
- 388 10. Reeves, P.J. *et al.* (2002) Structure and function in rhodopsin : a tetracycline inducible
389 system in stable mammalian cell lines for high-level expression of opsin mutants.
390 *Proc. Natl. Acad. Sci. U.S.A.* 99, 13413-13418
- 391 11. Mouillac, B. *et al.* (1992) Agonist-modulated palmitoylation of β_2 -adrenergic receptor
392 in Sf9 cells. *J. Biol. Chem.* 267, 21733-21737
- 393 12. Panneels, V. and Sinning, I. (2010) Membrane protein expression in the eyes of
394 transgenic flies. *Methods Mol. Biol.* 601, 135-147
- 395 13. Zhang, L. *et al.* (2005) Expression of functional G protein-coupled receptors in
396 photoreceptors of transgenic *Xenopus laevis*. *Biochemistry* 44, 14509-14518
- 397 14. Li, N. *et al.* (2007) Heterologous expression of the adenosine A1 receptor in
398 transgenic mouse retina. *Biochemistry* 46, 8350-8359
- 399 15. Sarramegna, V. *et al.* (2005) Solubilization, purification and mass spectrometry
400 analysis of the human mu-opioid receptor expressed in *Pichia pastoris*. *Protein Expr.*
401 *Purif.* 43, 85-93
- 402 16. Lundstrom, K. (2005) Structural genomics of GPCRs. *Trends Biotechnol.* 23, 103-108
- 403 17. Bokoch, M.P. *et al.* (2010) Ligand-induced regulation of the extracellular surface of a
404 G protein-coupled receptor. *Nature* 463, 108-112
- 405 18. Popot, J-L. (2010) Amphipols, nanodiscs, and fluorinated surfactants: three
406 nonconventional approaches to studying membrane proteins in aqueous solutions.
407 *Annu. Rev. Biochem.* 79, 13.1-13.39

- 408 19. Marullo, S. *et al.* (1988) Human β 2-adrenergic receptors expressed in *Escherichia coli*
409 membranes retain their pharmacological properties. *Proc. Natl. Acad. Sci. U.S.A.* 85,
410 7551-7555
- 411 20. Tucker, J. and Grisshammer, R. (1996) Purification of a rat neurotensin receptor
412 expressed in *Escherichia coli*. *Biochem. J.* 317, 891-899
- 413 21. Weiss, H.M. and Grisshammer, R. (2002) Purification and characterization of the
414 human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur. J.*
415 *Biochem.* 269, 82-92
- 416 22. Skretas, G. and Georgiou, G. (2009) genetic analysis of G protein-coupled receptor
417 expression in *Escherichia coli*. *Biotechnol. Bioeng.* 102, 357-367
- 418 23. White, J.F. *et al.* (2004) Automated large-scale purification of a G protein-coupled
419 receptor for neurotensin. *FEBS Lett.* 564, 289-293
- 420 24. Kiefer, H. *et al.* (1996) Expression of an olfactory receptor in *Escherichia coli*:
421 purification, reconstitution and ligand binding. *Biochemistry* 35, 16077-16084
- 422 25. Banères, J-L. *et al.* (2003) Structure-based analysis of GPCR function: conformational
423 adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant
424 BLT1. *J. Mol. Biol.* 329, 801-814
- 425 26. Banères, J-L. *et al.* (2005) Molecular characterization of a purified 5-HT4 receptor. *J.*
426 *Biol. Chem.* 280, 20253-20260
- 427 27. Michalke, K. *et al.* (2009) Mammalian G protein-coupled receptor expression in
428 *Escherichia coli*: I. High throughput large-scale production as inclusion bodies. *Anal.*
429 *Biochem.* 386, 147-155
- 430 28. Michalke, K. *et al.* (2009) Mammalian G protein-coupled receptor expression in
431 *Escherichia coli*: II. Refolding and biophysical characterization of mouse cannabinoid
432 receptor 1 and human parathyroid hormone receptor 1. *Anal. Biochem.* 401, 74-80

- 433 29. Arcemisbèhère, L. *et al.* (2010) Leukotriene BLT2 receptor monomers activate the G₁₂
434 GTP-binding protein more efficiently than dimers. *J. Biol. Chem.* 285, 6337-6347
- 435 30. Mouillac, B. and Banères, J-L. (2010) Mammalian membrane receptors expression as
436 inclusion bodies in *Escherichia coli*. *Methods Mol. Biol.* 601, 39-48
- 437 31. Kiefer, H. (2003) In vitro folding of alpha-helical membrane proteins. *Biochim.*
438 *Biophys. Acta.* 1610, 57-62
- 439 32. Huang, K.-S. *et al.* (1981) Refolding of an integral membrane protein. Denaturation,
440 renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic
441 fragments. *J. Biol. Chem.* 256, 3802-3809
- 442 33. Muller, I. *et al.* (2008) The full-length mu-opioid receptor: a conformational study by
443 circular dichroism in trifluoroethanol and membrane-mimetic environments. *J.*
444 *Membr. Biol.* 223, 49-57
- 445 34. Miller, D. *et al.* (2009) In vitro unfolding and refolding of the small multidrug
446 transporter EmrE. *J. Mol. Biol.* 393, 815-832
- 447 35. Schröder-Tittmann, K. *et al.* (2010) Recombinant expression, *in vitro* refolding, and
448 biophysical characterization of the human glucagon-like peptide-1 receptor.
449 *Biochemistry* 49, 7956-7965
- 450 36. Popot, J-L. *et al.* (1987) Refolding of bacteriorhodopsin in lipid bilayers. A
451 thermodynamically controlled two-stage process. *J. Mol. Biol.* 198, 655-676
- 452 37. Schmidt, P. *et al.* (2010) A reconstitution protocol for the in vitro folded human G
453 protein-coupled Y2 receptor into lipid environment. *Biophys. Chem.* 150, 29-36
- 454 38. Schimmer, S. *et al.* (2010) Functional characterization of the in vitro folded human
455 y(1) receptor in lipid environment. *Protein Pept. Lett.* 17, 605-609
- 456 39. Tribet, C. *et al.* (1996) Amphipols: polymers that keep membrane proteins soluble in
457 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 Ca²⁺-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

- 504 59. Flötenmeyer, M. *et al.* (2007) The use of amphipathic polymers for cryo-electron
505 microscopy of NADH:ubiquinone oxidoreductase (Complex I). *J. Microsc.* 227, 229-
506 235
- 507 60. Martinez, K. L. *et al.* (2002). Allosteric transitions of *Torpedo* acetylcholine receptor
508 in lipids, detergent and amphipols: molecular interactions *vs.* physical constraints.
509 *FEBS Lett.* 528, 251-256
- 510 61. Nagy, J. K. *et al.* (2001) Use of amphipathic polymers to deliver a membrane protein
511 to lipid bilayers. *FEBS Lett.* 501, 115-120
- 512 62. Tribet, C. *et al.* (1997) Stabilisation of hydrophobic colloidal dispersions in water with
513 amphiphilic polymers: application to integral membrane proteins. *Langmuir* 13, 5570-
514 5576
- 515 63. Catoire, L. J. *et al.* (2010) Solution NMR mapping of water-accessible residues in the
516 transmembrane β -barrel of OmpX. *Eur. Biophys. J.* 39, 623-630
- 517 64. Raschle, T. *et al.* (2010) Nonmicellar systems for solution NMR spectroscopy of
518 membrane proteins. *Curr. Opin. Struct. Biol.* 20, 1-9
- 519 65. Gorzelle, B.M. *et al.* (2002) Amphipols can support the activity of a membrane
520 enzyme. *J. Am. Chem. Soc.* 124, 11594-11595
- 521 66. Tribet, C. *et al.* (1998) Scanning transmission electron microscopy study of the
522 molecular mass of amphipol/cytochrome b_6f complexes. *Biochimie* 80, 475-482
523
524
525

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_5 I	IBs	0.8-1.2	[29]

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

^bIn 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 from its fusion partner (if applicable) by enzyme cleavage, folded and/or purified and its activity checked by ligand binding.

^cExcept for CB1 and NTS1, which were produced in fermentors, all other receptors have been overexpressed in culture flasks.

562 **Table 2. Comparison of GPCR folding yields obtained in different surfactants**
 563

564 GPCR	565 Folding medium	566 Average maximum 567 folding yield (%) ^a	568 Refs.
569 BLT1	570 Detergent/lipid mixed micelles 571 (LDAO/asolectin)	572 30	573 [25]
	574 A8-35	575 50	576 [42]
	577 A8-35/asolectin	578 65	579 [42]
580 BLT2	581 Detergent/lipid mixed micelles 582 (DPC/HDM/asolectin)	583 4	584 [29]
	585 A8-35	586 50	587 [42]
	588 A8-35/asolectin	589 70	590 [42]
591 5-HT _{4A}	592 DMPC/CHAPS bicelles	593 25	594 [26]
	595 A8-35	596 30	597 [42]
	598 A8-35/asolectin	599 60	600 [42]
601 CB1	602 Detergent/lipid mixed micelles ^b 603 (Fos-choline-16/asolectin)	604 0	605 [42]
	606 A8-35	30	[42]
	A8-35/asolectin	40	[42]
	Detergent mixed micelles (DDM/Cymal 6)	30	[28]

601 ^aThe folding yield is the amount of functional protein obtained after folding (based on binding
 602 of specific ligands) compared to that of protein in starting SDS solution (based on A₂₈₀
 603 measurements).

604 ^bNo extensive detergent screening was carried out to optimize folding of the CB1 receptor in
 605 detergent/lipid mixed micelles.
 606

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

612 **Figure legends**

613 **Figure 1. Strategies to overexpress GPCRs in *E. coli*.** GPCRs can be produced in bacteria
614 either by insertion into the inner membrane ① or by accumulation into inclusion bodies ②. In
615 the first case, a protein partner is coupled at the N-terminus (MBP, β -gal) or C-terminus
616 (GFP, TRX) of the receptor, or at both extremities (for instance MBP and TRX), so as to
617 target the recombinant protein to the membrane. In the second case, targeting to IBs is
618 favored by coupling to the N-terminus another fusion partner, such as GST, KSI, TRX, NusA
619 or α_5 I. GPCRs accumulated as inclusion bodies are not adequately folded and have to be
620 solubilized in a harsh detergent before folding.

621
622 **Figure 2. Strategies to fold GPCRs from IBs.** Before folding, the recombinant GPCR
623 accumulated in IBs has to be solubilized in denaturing buffers (for instance a mix of urea and
624 SDS), purified by immobilized-metal affinity chromatography (IMAC) and dialyzed in an
625 aqueous buffer to eliminate most SDS and allow removal of the fusion partner, for instance
626 via a thrombin cleavage. After elimination of the cleaved fusion partner through a second
627 IMAC purification in denaturing conditions, the GPCR is kept soluble in SDS solution in
628 which it displays a significant content of helicity. Folding can be achieved by transfer from
629 SDS to mild detergents or amphipols, supplemented or not with lipids. The functional fraction
630 of the GPCR preparation can be evaluated using pharmacological assays (ligand binding,
631 activation of purified G proteins or arrestins).

632
633 **Figure 3. Chemical structures of three families of APols.** A) A8-35, the prototypal APol.
634 B) Sulfonated APol (SAPol). C) Glucose-based non-ionic APol (NAPol).

Bacterial outer membrane

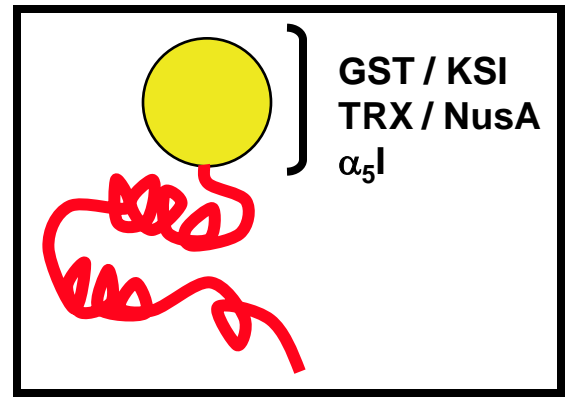
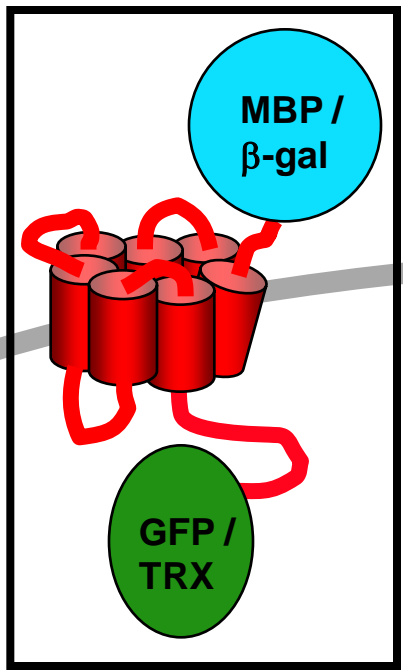
1

Membrane insertion

Periplasm

Cytoplasm


Bacterial inner membrane

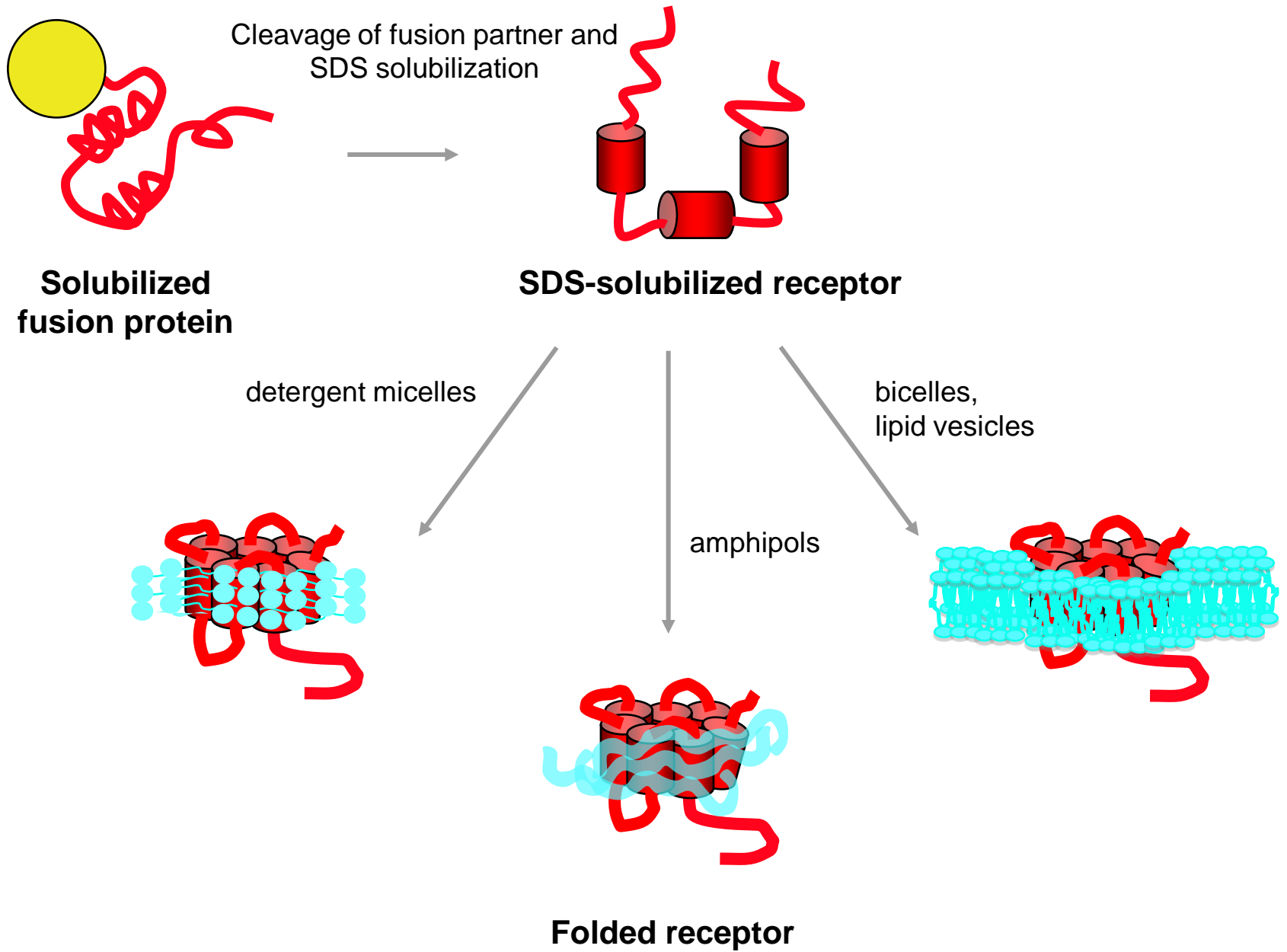


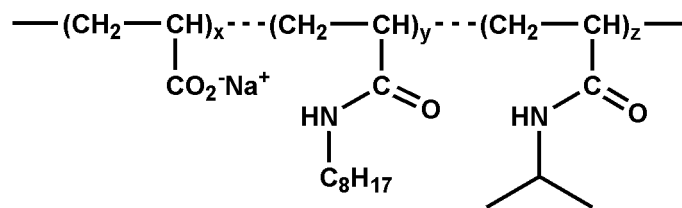
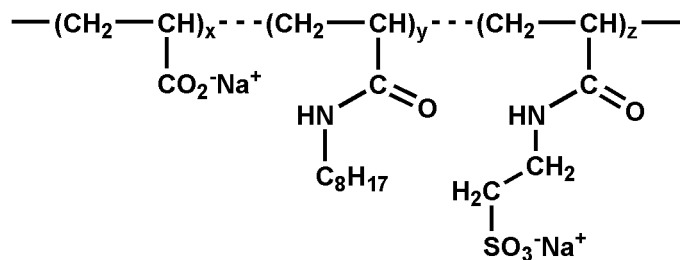
2

Aggregation into inclusion bodies

DNA (Chromosome, Plasmid)





A**B****C**