HHS Public Access

Author manuscript

Methods. Author manuscript; available in PMC 2018 October 13.

Published in final edited form as:

Methods. 2011 December ; 55(4): 293–302. doi:10.1016/j.ymeth.2011.08.004.

Crystallization chaperone strategies for membrane proteins

Raquel L. Lieberman[1,*], **Jeffrey A. Culver**[1], **Kevin C. Entzminger**[2], **Jennifer C. Pai**[2], and **Jennifer A. Maynard**[2,*]

[1]School of Chemistry and Biochemistry and Institute for Bioscience and Bioengineering, Georgia Institute of Technology, 901 Atlantic Drive NW, Atlanta, GA 30332, USA

[2]Department of Chemical Engineering, University of Texas at Austin, MC0400, 1 University Station, Austin TX 78712, USA.

Abstract

From G protein-coupled receptors to ion channels, membrane proteins represent over half of known drug targets. Yet, structure-based drug discovery is hampered by the dearth of available three-dimensional models for this large category of proteins. Other than efforts to improve membrane protein expression and stability, current strategies to improve the ability of membrane proteins to crystallize involve examining many orthologs and DNA constructs, testing the effects of different detergents for purification and crystallization, creating a lipidic environment during crystallization, and cocrystallizing with covalent or non-covalent soluble protein chaperones with an intrinsic high propensity to crystallize. In this review, we focus on this last category, highlighting successes of crystallization chaperones in membrane protein structure determination and recent developments in crystal chaperone engineering, including molecular display to enhance chaperone crystallizability, and end with a novel generic approach in development to target any membrane protein of interest.

Introduction

Integral membrane proteins comprise approximately thirty percent of sequenced genomes [1] and possess functions as diverse as their soluble protein counterparts. G protein-coupled receptors (GPCRs), for example, is the largest family of membrane proteins in mammals [2] and are present in nearly every organ. GPCRs are central to mediating how cells respond to hormones and neurotransmitters as well as the senses of smell, taste, and vision; GPCRs comprise ~50% of all drug targets, as they are involved in numerous human disorders [3]. Similarly, ion channels are critical to the function of nerve and muscle cells, with implications for arrhythmias, diabetes, and epilepsy [4]. Within the cell, membrane proteins play important roles in protein translocation to the endoplasmic reticulum for folding [5], and in pathogenic bacteria, membraneous porins and ATP-binding cassette (ABC) transporters contribute to multidrug resistance [6]. On another end of the spectrum lie, for example, ubiquitous membrane-bound enzymes critical to energy in cells, such as the F_0-F_1 ATP synthase [7] and photosystem II [8].

*Corresponding author: Raquel L. Lieberman: 404-385-3663 (T), 404-894-2295 (F), raquel.lieberman@chemistry.gatech.edu or Jennifer A. Maynard: 512-471-9188 (T), 512-471-7060 (F), maynard@che.utexas.edu.

Despite their involvement throughout biology, integral membrane proteins are severely underrepresented in the Protein Data Bank (PDB; http://www.rcsb.org). Of the nearly 70,000 protein structures housed in the PDB, roughly 300 of these represent unique membrane protein structures solved predominantly by X-ray crystallography, but also by electron crystallography and nuclear magnetic resonance spectroscopy (see http://blanco.biomol.uci.edu/mpstruc/listAll/list). Expression of membrane proteins is one major bottleneck to structure determination due, at least in part, to the naturally low abundance of membrane proteins in their native host, and their potential to be toxic to the heterologous expression host. Recombinant expression, particularly in the case of bacterial membrane proteins [9], in sufficient yield for structure determination experiments, has been achieved using *E. coli* [10]. By contrast, expression of eukaryotic membrane proteins and membrane protein complexes in adequate yield for structural characterization in hosts such as yeast, human embryonic kidney or insect cells, or by cell free expression, remains an ongoing challenge [11, 12]. In general, multiple orthologs, DNA constructs, and expression platforms are explored before a suitable system is found [9]. A second major impediment to membrane protein structure determination is isolation and purification. A membrane protein is typically extracted in a micelle-forming, water-soluble, amphiphilic detergent designed to replace and mimic the phospholipid bilayer. Although hundreds of detergents are commercially available, finding a suitable detergent that retains both structure and function of a membrane protein is an empirical process [10, 13]. Specific to structure determination by X-ray crystallography, a detergent suited to purification is not necessarily suited for crystallization into a three-dimensional lattice. Complicating factors for crystallization include the fact that the detergent itself undergoes phase transitions in the traditional vapor diffusion experiment [13], and residual host lipids may still remain after solubilization with the membrane protein and thus contribute to sample heterogeneity. Moreover, inherent in their adaptation to a hydrophobic lipidic environment, membrane proteins possess a dearth of polar residues necessary for generating stable crystal contacts; these residues must not be occluded by the detergent micelle [14].

Once adequate expression and purification conditions have been identified, there are numerous strategies to increase the likelihood of obtaining crystals of a membrane protein, primarily based on reducing the entropy cost of crystal lattice formation [15] and providing ample residues capable of forming lattice contacts [16]. In this review, we discuss the various methods used to crystallize membrane proteins (summarized in Table 1). We include a brief description of non-chaperone methods that make improvements to the stability of the membrane protein, and focus on non-covalent chaperone technologies in which the membrane protein of interest forms a stable complex with a readily crystallized protein partner to enable lattice formation and subsequent structure determination.

Non-chaperone techniques

The main non-chaperone based methods to improve the solution behavior of a membrane protein to enhance its crystallization potential include making strategic amino acid substitutions to enhance inherent stability [17], removing disordered regions by proteolysis [18], stabilizing the protein of interest by complexation with a known ligand [19] and/or tailoring the surrounding chemical environment to better mimic the hydrophobic lipid

membrane [20]. Although these are highly accessible approaches for any membrane protein, the number of variables introduced systematically necessitates extensive crystallization screening and analysis. Advances in high throughput crystallization screening methodology and instrumentation [21] diminish this disadvantage, however.

Binding Partners.

Cocrystallization of a membrane protein with a known binding partner such as a soluble protein or other high affinity ligand can stabilize the membrane protein and help reduce conformational heterogeneity. This approach does not increase the available lattice contacts and in some cases may interfere with crystallization [22], but in general, an increase in stability and reduction in heterogeneity are expected to enhance crystallization propensity. Just a few of the many examples of membrane protein structures solved in the presence of bound ligands include (1) β-adrenergic receptors in complex with agonists or antagonists designed to restrict the GPCR to a single conformation (PDB code 3QAK, 2Y03, 3PDS, among others) [23–25], (2) a sodium symporter in complex with clinically relevant tricyclic antidepressants and amino acids (PDB code 2QB4, 2Q6H 2Q72, 2QEI) [26], (3) KcsA K^+ channel in the presence of channel-blocking tertbutylammonium ion (PDB code 2HJF, among others) [27–29], (4) glutamate-gated chloride channel in complex with, in turn, an allosteric agonist, an open channel blocker, or the neurotransmitter glutamate (PDB codes 3RHW, 3RI5, 3RIF) [30], (5) *E. coli* ammonia channel in complex with regulator protein GlnK (PDB code 2NUU) [31] and (6) *E. coli* lactose permease in the presence of a lactose homolog, β-D-glactopyrnaosyl-1-thio-β-D-galactopyranoside (PDB code 1PV6) [32].

Detergents.

The empirical process of identifying a detergent suitable to crystallization of a membrane protein, while time consuming, is one that has been met with success for solved integral membrane protein crystal structures to date. Although dodecyl-β-D-maltoside and β-octylglucoside are popular, many membrane proteins are unstable in these detergents [14, 33]. New detergents continue to be synthesized with various properties thought to better mimic the lipid bilayer upon solubilization, and offer additional stability to the membrane protein [34–36]. In addition, facile assays to assess the stability of a protein in a range of detergents using thermal stability [37] or cysteine binding [38] have been developed. These assays can assist in the decision to continue with crystallization trials.

Lipidic Cubic Phases.

An alternative to crystallization in solution containing detergents is crystallization in the presence of a synthetic lipid bilayer system. Crystallization *in cubo* employs a spontaneously forming cubic phase, where single lipid bilayers extend indefinitely in three dimensions, with aqueous channels on either side of each bilayer [39]. Other lower symmetry lipidic phases, sponge phases, can also been employed [40]. The purified membrane protein is incorporated into the lipid phase, and with the addition of a crystallization cocktail, the protein molecules diffuse laterally through the membrane such that crystal nucleation and growth can begin. Since the initial successes, over fifty structures have been solved using this method. Kits for crystallization using LCP are now commercially available (Emerald

BioSystems) and numerous excellent reviews have been written on the subject (see for example [20, 39, 41]).

Chaperone technologies

Cocrystallization with a macromolecular chaperone is an emerging technique that has been successfully employed for a diverse selection of membrane proteins. Chaperones for this purpose are chosen for their excellent biophysical properties, including proclivity to crystallize. The formation of the chaperone:membrane protein complex not only increases the solubility of a target membrane protein but also increases the hydrophilic surface area available to form crystal contacts (see Figure 1). Hydrophilic surface residues present in a purified membrane protein are few and may be occluded by the detergent micelle; the crystallization chaperone extends beyond the micelle and mediates the majority of contacts within a crystal lattice [16] (Figure 2). In addition, chaperones can mask areas of conformational heterogeneity by reducing local flexibility upon complex formation, thus decreasing the entropy cost of lattice formation. Finally, crystallization chaperones also aid in structure determination. For the molecular replacement phasing strategy, the well-studied chaperone protein can serve as the search model, and can be used to phase the protein:chaperone complex. Alternatively, rather than going through the empirical process of heavy atom derivatization of the membrane protein, which is not likely to habor many metal ion binding sites, defined heavy atom sites can be instead incorporated into the chaperone. A chaperone:membrane protein complex can be generated by fusing the desired soluble protein at the plasmid level or by non-covalent interactions. As detailed below, some of these methods require extensive work with each individual membrane protein, while others can be readily applied to the structural determination of diverse membrane proteins.

Covalent chaperones.

For covalent attachment of the chaperone to the membrane protein of interest, the DNA sequence encoding an area of high conformational flexibility or an otherwise non-conserved hydrophilic region of the protein is replaced with that of a soluble protein with high crystallization propensity. This results in a chimera with increased hydrophilicity and thus has a higher likelihood of forming diffraction-quality crystals. Early attempts of this approach include fusion of cytochrome b_{562} into a cytoplasmic loop of *E. coli* lactose permease [42, 43] and incorporation of protein Z into *E. coli* cytochrome bo_3 ubiquinol oxidase at its C-terminus [44]. Unfortunately, although in both cases crystals were improved, the lattices were still insufficient for structure determination. More recently, however, this method has gained traction [45], as GPCR family members have been crystallized using this technique. Substitution of a highly dynamic intracellular loop with T4 bacteriophage lysozyme in combination with other modifications (see binding partners above and non-covalent chaperones below) enabled the growth of crystals *in cubo* that were suitable for structure determination [46] (Figure 1 b,c, Figure 2a).

Covalent chaperones are attractive because they represent a general approach to membrane protein crystallization; the gene for any crystallizable soluble protein can be readily introduced into a plasmid containing the gene for the membrane protein of interest.

However, the placement of the fusion protein is of critical importance, yet is usually based on limited knowledge of the membrane protein. N- or C- terminal constructs necessitate a peptide linker, which could introduce new, undesirable, heterogeneity that impedes crystal growth [47]. In the β_2-adrenergic receptor (β_2AR) example, several constructs were generated to optimize expression and protein behavior in solution. In addition, it is a challenge not to interfere with membrane protein function [46]. The effects of internal T4 lysozyme insertion in the β_2AR construct ultimately used for structure determination were loss of the ability of G- proteins to bind the GPCR, and higher baseline activity compared to wild type [48]. However, recently, N-terminal fusion of T4 lysozyme to β_2AR enabled a ternary structure of β_2AR in complex with a Gs protein and a GPCR-specific nanobody (described below) to be solved [49]. Nevertheless, overall, this covalent chaperone approach, while elegant in principle, is ultimately cumbersome in practice, as efforts are required both on the molecular biology, protein expression, and protein purification front to generate a stable, active fusion construct, as well as the crystallization front to grow suitable crystals for structure determination.

Non-covalent chaperones.

An increase in the hydrophilic surface area and crystallization propensity of a membrane protein can also be achieved by non-covalent complexation with a highly crystallizable protein partner. Several potential non-covalent binding partners for a single membrane protein of interest can be identified. An important advantage of such a system is that the chaperone and target membrane protein can be optimized independently for their behaviors in solution. The affinity of the two proteins for one another can be measured and modulated as necessary, down to single nanomolar affinity [50]; however, because the high protein concentrations within the crystallization drop are likely significantly greater than the dissociation constant for the complex, an absolute low affinity value is a less stringent requirement compared to other antibody applications. Antibody fragments, especially Fab and Fv, are common choices for membrane protein co-crystallization scaffolds. These fragments have been instrumental in the structure determination of numerous new membrane proteins (see below) and in many cases they form better diffracting crystals that yield a higher resolution structure. They have further assisted visualization of membrane proteins in specific conformations [50, 51] and have been employed for difficult soluble protein [52–54] and RNA [55, 56] structures. However, antibody fragment expression yields can be limited due to folding efficiency, issues with correct disulfide bond formation, and limited solubility of the resulting antibody fragment [57].

Thus, other small crystallizable proteins derived from combinatorial libraries have also been developed [22] (Figure 3). The emergence of non-antibody based scaffolds has further expanded the toolbox for the aspiring crystallographer (see below); these formats not only avoid limitations of the antibody fragments but also provide access to novel structural and chemical properties.

Hybridoma Technology.

One method to obtain a high affinity protein binding partner is to raise monoclonal antibodies using hybridoma technology. After immunization, splenic B cells are fused with

an immortalized myeloma cell line to generate individual hybridoma lines [58, 59]. Recognition of a particular epitope on the protein is difficult to control, but antibodies that bind to key regions of the folded membrane protein can be identified by ELISA screening and corroborated by other techniques [60]. In general, these cell lines will produce an ample supply of a monoclonal antibody for crystallization. After production, the monoclonal antibody is purified by affinity chromatography, and the less flexible, monovalent Fab fragment (Figure 3a) is obtained by proteolysis and removal of the Fc fragment [16]. Alternatively, primers can be used to clone the cDNA of the antibody fragment for recombinant expression [61]; this method is particularly useful if it is desired to express the Fv fragment in *E. coli* [62, 63].

Hybridoma-derived antibody fragments have been used to crystallize numerous membrane proteins. The first successful example of this approach was the 2.7 Å resolution structure of *P. denitrificans* cytochromc *c* oxidase (COX) catalytic subunits (PDB code 1AR1) [64]. The Fv fragment, identified by hybridoma screening, then cloned and expressed in *E. coli* [62], recognizes an epitope on a soluble domain of COX and mediates most of the crystal contacts [16]. Additional contacts are observed between the periplasmic surfaces of adjacent COX proteins [64]. Interestingly, the complex was isolated by incubating solubilized membranes with a strep-tagged Fv fragment and purified using streptavidin affinity resin [62]. Another early example is the 2.3 Å resolution structure of yeast cytochrome *bc1* in complex with an Fv fragment (PDB code 1EZV) [65], which was isolated by gel filtration after mixing the two purified proteins together. The Fv binds to the extrinsic domain of the catalytic subunit (Rieske protein) using a discontinuous epitope [16] and the main crystal contacts are provided by the antibody fragment. This strategy was used again to solve the structure of cytochrome *bc1* in complex with its substrate cytochrome *c* [66], supporting the notion that the Fv fragment does not interfere with cytochrome *bc1* function.

Numerous ion channels have also been solved with the assistance of Fab antibody fragments, and in particular, use of the antibody fragment was found to result in a higher resolution structure. The *A. pernix* KvAP channel structure was solved with and without antibody fragments bound to the voltage sensor elements of the channel. Although no differences in structure were observed, crystals without the Fv fragment diffracted to 8 Å resolution while those with the antibody fragment diffracted to 3.9 Å resolution [67]. Similarly, the first *S. levidans* KcsA potassium ion channel was solved to 3.2 Å resolution (1BL8), whereas a 2.0 Å resolution was obtained when KcsA was solved with a Fab fragment that recognized the functional tetramer of the protein, on its extracellular surface (PDB code 1K4C) (Figure 1a, pink ribbon). In this structure, the Fab fragments not only provided all crystal contacts and a large cavity to accommodate a detergent micelle, but also served as a molecular replacement search model used to solve the structure [68]. After these early successes, Fab-mediated crystallization has enjoyed increasing popularity and success (Table 2), including with GPCRs (Figure 1b). Beyond Fabs, the use of nanobodies, antibodies derived from immunized llamas (see below), is gaining traction. A nanobody was critical to obtaining diffraction-quality crystals of the β_2AR-Gs protein complex [49]. The complex on its own, which itself included an N-terminal T4 lysozyme fusion to β_2AR (see above), diffracted to just 7 Å resolution. The addition of the nanobody formed a ternary

complex in which a conformationally flexible region was stabilized, and provided additional hydrophilic residues for crystal contacts, leading to a 3.2 Å resolution structure.

Although the hybridoma approach to obtaining a soluble, crystallizable antibody fragment is relatively robust, straightforward, and a service commercially available, there are some drawbacks. First, the process takes 6–9 months, and not all samples will be sufficiently immunogenic, for example, for cases in which there is a homologous protein in the immunized animal. Even if antibodies are isolated, the affinity of the interaction cannot be modulated, nor are desired favorable biophysical properties such as stability and solubility guaranteed. In addition, because there is little control over the epitope recognized, there is no assurance that the particular antibody fragment will yield crystals suitable for structure determination or the desired conformation [50]. Finally, each protein of interest requires the investment of new time and resources for the immunization procedure.

Directed evolution technologies to discover and optimize non-covalent chaperones.

Molecular display technologies expand the selection of high specificity protein partners for membrane protein crystallization by enabling tailored chaperones to be identified in a higher throughput fashion than the hybridoma approach. Further, display techniques can also be used to optimize the biochemical and biophysical properties of the chaperones, including site-directed mutagenesis to reduce surface entropy [69] and homolog shuffling or random mutagenesis to identify more soluble [70, 71] or more stable [72] variants. Binding partners to a wide variety of ligands are generated using directed mutagenesis of residues directly mediating ligand recognition. Inspired by the successes of antibodies derived from hybridomas, some of the scaffolds used are antibody fragments. These include (a) Fab, which contains a heavy and light chain each from a variable and constant region (Figure 3a) (b) the single-chain variable fragment (scFv), the minimal antibody binding domain in which the variable light and heavy domains are linked by a flexible polypeptide to yield a single gene (Figure 3b) (c) nanobodies, derived from shark and camel antibodies that lack light chains, which are analogous to a single variable heavy chain but employ an exceptionally long complementary determining region loop 3 (CDR3) to mediate high affinity, high specificity interactions [73] (see above, Figure 1d, Figure 3c). Non-immunoglobulin-based formats include affibodies, based on the antibody-binding domain of protein A [74], designed ankyrin repeat proteins (DARPINs), based on the ankyrin repeat domain important in formation of specific protein-protein interactions [75], anticalins, based on a lipocalin protein scaffold endogenously involved in the binding and transport of small molecules [76], and fibronectin type III (FN3) domains, which serve as a minimalist antibody mimic [77] (Figure 3d–g).

All of these formats are adapted to expression in *E. coli*, allowing for inexpensive, facile and rapid production and purification. Many of the alternative formats lack disulfide bonds, enabling cytosolic expression in *E. coli* and much higher yields; for example, DARPINs can yield up to 200 mg/L [75] while an Fv or Fab may express at ~2–5 mg/ L [78]. The availability of a variety of chaperone formats allows for fine-tuning of the size ratio and the geometry of the chaperone-membrane protein complex. If a membrane-protein specific antibody has been identified, that binding specificity can be readily transferred between the

three related antibody-based formats (Fab - 50kDa, Fv - 28kDa and nanobody - 14kDa) [79]. The alternative chaperones are very compact and rigid, characteristics that typically correlate with a propensity to crystallize [69]. Alternate chaperone formats access sizes ranging from 7kDa (affibody), 11kDa (FN3), 17kDa (DARPIN) to 20kDa (anticalin; Figure 3). Finally, this technology enables a chaperone to be selected from an existing library for each new membrane protein of interest, significantly reducing the workload as compared to traditional hybridoma approaches.

Molecular display platforms rely on formation of a stable physical link between genotype and phenotype. Protein variants can then be selected based on a particular characteristic, such as membrane protein specificity, with the corresponding gene recovered and sequenced. A number of display systems have been developed, including ribosome display [80], yeast display [81], bacterial display [82] and phage display [83]. Of these, phage display is readily accessible to most molecular biology labs and requires minimal specialized equipment. Chaperone proteins are expressed as a fusion to the M13 phage minor coat protein, which is then incorporated into the surface of phage particles as a functional protein, while the corresponding DNA is packaged within the phage particle. A typical selection process proceeds with several selection rounds as follows: (1) phage displaying chaperones are allowed to bind an immobilized, detergent solubilized membrane protein, (2) weakly binding phage are removed by washing while (3) strongly bound phage are eluted with an acidic buffer prior to (4) amplification in *E. coli* (Figure 4, Table 3). In this manner, a handful of chaperones can be selected for more thorough characterization from an initial library of 10^7-10^{10} potential chaperones. With this selection process, protein-specific chaperones can be readily identified and optimized. For *de novo* identification, the regions of the chaperone responsible for ligand binding are randomized, generating large libraries of variants in which each variant potentially binds a distinct ligand. For antibody-based chaperones, ligand specificity is determined by the six complementarity determining regions (CDRs), which can each be randomized according to observed variation [84] or a reduced genetic code, in which each position is randomized to tyrosine or serine [85].

The molecular display approach has proven valuable for discovery of chaperones binding rare epitopes or stabilizing active conformations. For example, early structural studies of the KcsA K^+ channel employed truncated variants lacking the flexible C-terminal region [68] (Figure 1a, green ribbon). To crystallize the full-length KcsA channel, Fab molecules specific for the full-length protein were identified from an existing Fab library constructed with expanded diversity in the third heavy chain CDR and restricted or binary code diversity in three other CDRs [86]. Starting with the Herceptin Fab scaffold, which was already engineered for high-level expression and compatibility with phage display, libraries were constructed with limited diversity in four of the six CDRs. After three rounds of selection, four KscA-specific Fabs were selected, of which three bound epitopes in the C-terminal region, and in particular, one led to a full-length structure of KcsA in its closed state [50] (Figure 1a, yellow ribbon). The chaperone recognized a highly charged epitope of the C-terminal domain and was primarily responsible for mediating crystal lattice contacts. Similarly, four Fabs with low nanomolar affinity for the detergent-solubilized citrate carrier *K. pneumonia* CitS were identified from an optimized phage display library that included highly-expressed and stable members [87], although no structure has yet been reported.

Finally, DARPins were selected for high affinity to the *E. coli* AcrB, leading to different lattices and structure quality [88]. Compared to the hybridoma approach, there are fewer examples of membrane proteins crystallized with recombinant chaperones (Table 2). This newer technology, which requires generation of highly diverse, rich libraries requires significant effort and is facilitated by robust robotic screening protocols, is rapidly gaining favor for crystallization of difficult soluble proteins and is expected to enjoy similar successes with membrane proteins in the near future.

Once a protein-specific chaperone gene has been identified, it can be further engineered to enhance crystallization propensity. Again using molecular display technologies, a new library is created, based on variants of a single chaperone. These variants can include site-directed mutations, for instance in the specificity determining residues (as above) or on the surface to adjust hydrophobicity and charge, in order to increase expression levels, solubility and the propensity to form crystal contacts. Alternatively, or to modulate other properties in an unbiased manner, random mutagenesis can be used to create small changes from the native sequence. Optimized chaperones are then selected from the library based on the desired property and further characterized [78].

Multifunctional crystallization chaperones.

All of the techniques described above have demonstrated tremendous success in the crystallization of recalcitrant membrane proteins, but each possesses its own drawbacks, either due to expense, knowledge of specialized techniques, need for prior structural knowledge, lack of control over chaperone properties, or potential loss of function or activity to the protein of interest. In addition, a shortcoming of current crystallization chaperone approaches is that each chaperone is specific for a single target; thus, a new chaperone must be identified for each new membrane protein target (Figure 5a). This can be a slow process, especially when hybridoma approaches are used.

A potentially general solution to these issues involves engineered chaperones with binding specificity for short peptide sequences that can be inserted into a loop of a membrane protein to form a high affinity complex for crystallization (Figure 5b). In principle, the cognate peptide sequence could be imported into any non-conserved, non-functional extramembraneous loop of any membrane protein of interest by site-directed mutagenesis. Even though exact peptide placement may require optimization, the limited extent of the amino acid substitution is likely to be less invasive than protein fusion as with T4 lysozyme. The chaperones will not require re-engineering for each new membrane protein target, allowing limited screening to generate a suitable complex for crystallization, and ensuring compatibility with high-throughput structural genomics programs. As with the other chaperones described herein, complex formation will simultaneously immobilize a flexible loop and provide hydrophilic surface area for crystal contact formation (Figure 1, 2). Ideally, the chaperone will have known crystallization conditions and crystal packing that does not involve specificity-determining regions, allowing for the target protein to be both readily bound by the chaperone and incorporated into the lattice.

The first such candidate chaperone, an scFv that recognizes a terminal hexa-histidine tag, was derived from a phage-display library, further optimized by rational site-directed

mutagenesis to confer enhanced solubility, and crystallized in a lattice in which the binding sites faced ~70 Å solvent cavities. This open lattice presents an opportunity to encapsulate and co-crystallize small poly-histidine-tagged proteins [89]. In terms of efforts to apply this strategy to crystallize membrane proteins, a complex between Fab fragments of the commercially available anti-FLAG M2 antibody (Sigma) and the KvPae K+ channel presenting a minimal FLAG peptide in an extramembraneous loop [90], was isolated. Unfortunately, crystallization trials of the complex led to crystals of just the anti-FLAG M2 Fab antibody fragment, the structure of which was solved to 1.8 Å resolution. Recently, we engineered and crystallized scFv chaperones specific for the hexa-histidine and EE peptides (sequence EYMPME), able to bind target proteins presenting C-terminal, intra-loop or intra-domain peptide tags with high affinity [78]. Phage display, in conjunction with directed evolution, allowed us to isolate scFv variants with enhanced expression, solubility and peptide tag affinity, properties that predispose the protein to crystallization. These chaperones form stable complexes with a membrane protein containing the EE peptide in a pre-determined locale (*unpublished results*) and are being converted to Fab and nanobody formats to tailor the chaperone size, and thus the geometry and hydrophillic surface area, of the resulting complex. It is envisioned that in the long term, these crystallization chaperones could provide a diverse toolbox for structural studies of membrane proteins, allowing crystallographers to choose among peptide tags, chaperone size and biophysical characteristics, and the use of a single or multiple chaperones during crystallization trials.

Conclusions

Strategies to increase the likelihood of growing diffraction-quality membrane protein crystals are multifaceted. They include tailoring the construct to achieve a homogenous protein solution, testing different detergents and lipid environments, and introducing soluble proteins either covalently or non-covalently to increase the hydrophilic surface area available to form crystal contacts. Inspired by successes solving membrane protein structures with Fab antibody fragments obtained by hybridoma technology, recombinant crystallization chaperones based on both antibody and alternative scaffolds offer additional versatility and potentially increase throughput for identifying suitable crystallization chaperone candidates for any membrane protein of interest. This expanded toolbox has already yielded several important membrane protein structures (Table 2), and it is expected that novel scaffolds and chaperones with tailored biophysical and crystallization properties will be increasingly important for membrane protein structure determination.

Acknowledgements.

Work in the authors' laboratories on this topic is funded by NIH grants R21DK91357 (RLL) and R01GM095638 (JAM). Figures were prepared using PyMol (www.pymol.org).

References

[1]. Wallin E, von Heijne G, Protein Sci, 7 (1998) 1029–1038. [PubMed: 9568909]
[2]. Blois TM, Bowie JU, Protein Sci, 18 (2009) 1335–1342. [PubMed: 19536805]
[3]. Allen JA, Roth BL, Annu. Rev. Pharmacol. Toxicol, 51 (2011) 117–144. [PubMed: 20868273]
[4]. Yu FH, Catterall WA, Genome Biol, 4 (2003) 207. [PubMed: 12620097]

[5]. Rapoport TA, Nature, 450 (2007) 663–669. [PubMed: 18046402]
[6]. Li XZ, Nikaido H, Drugs, 69 (2009) 1555–1623. [PubMed: 19678712]
[7]. Okuno D, Iino R, Noji H, J Biochem, 149 (2011) 655–664. [PubMed: 21524994]
[8]. Williamson A, Conlan B, Hillier W, Wydrzynski T, Photosynth. Res, 107 (2011) 71–86. [PubMed: 20512415]
[9]. Wang DN, Safferling M, Lemieux MJ, Griffith H, Chen Y, Li XD, Biochim Biophys Acta, 1610 (2003) 23–36. [PubMed: 12586376]
[10]. Newby ZE, O'Connell JD, 3rd, Gruswitz F, Hays FA, Harries WE, Harwood IM, Ho JD, Lee JK, Savage DF, Miercke LJ, Stroud RM, Nat Protoc, 4 (2009) 619–637. [PubMed: 19360018]
[11]. Freigassner M, Pichler H, Glieder A, Microb. Cell Fact, 8 (2009) 69. [PubMed: 20040113]
[12]. Lee JK, Stroud RM, Curr. Opin. Struct. Biol, 20 (2010) 464–470. [PubMed: 20739007]
[13]. Koszelak-Rosenblum M, Krol A, Mozumdar N, Wunsch K, Ferin A, Cook E, Veatch CK, Nagel R, Luft JR, Detitta GT, Malkowski MG, Protein Sci, 18 (2009) 1828–1839. [PubMed: 19554626]
[14]. Prive GG, Methods, 41 (2007) 388–397. [PubMed: 17367711]
[15]. Derewenda ZS, Acta Crystallogr, D67 (2011) 243–248.
[16]. Hunte C, Michel H, Curr Opin Struct Biol, 12 (2002) 503–508. [PubMed: 12163074]
[17]. Argos P, Rossman MG, Grau UM, Zuber H, Frank G, Tratschin JD, Biochemistry, 18 (1979) 5698–5703. [PubMed: 518863]
[18]. Wernimont A, Edwards A, PLoS One, 4 (2009) e5094. [PubMed: 19352432]
[19]. Connelly PR, Curr Opin Biotechnol, 5 (1994) 381–388. [PubMed: 7765170]
[20]. Caffrey M, J Struct Biol, 142 (2003) 108–132. [PubMed: 12718924]
[21]. Stevens RC, Curr Opin Struct Biol, 10 (2000) 558–563. [PubMed: 11042454]
[22]. Koide S, Curr. Opin. Struct. Biol, 19 (2009) 449–457. [PubMed: 19477632]
[23]. Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V, Stevens RC, Science, 332 (2011) 322–327. [PubMed: 21393508]
[24]. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF, Nature, 454 (2008) 486–491. [PubMed: 18594507]
[25]. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK, Nature, 469 (2011) 236–240. [PubMed: 21228876]
[26]. Singh SK, Yamashita A, Gouaux E, Nature, 448 (2007) 952–956. [PubMed: 17687333]
[27]. Faraldo-Gomez JD, Kutluay E, Jogini V, Zhao Y, Heginbotham L, Roux B, J Mol Biol, 365 (2007) 649–662. [PubMed: 17070844]
[28]. Yohannan S, Hu Y, Zhou Y, J Mol Biol, 366 (2007) 806–814. [PubMed: 17196615]
[29]. Lenaeus MJ, Vamvouka M, Focia PJ, Gross A, Nat Struct Mol Biol, 12 (2005) 454–459. [PubMed: 15852022]
[30]. Hibbs RE, Gouaux E, Nature, 474 (2011) 54–60. [PubMed: 21572436]
[31]. Conroy MJ, Durand A, Lupo D, Li XD, Bullough PA, Winkler FK, Merrick M, Proc Natl Acad Sci U S A, 104 (2007) 1213–1218. [PubMed: 17220269]
[32]. Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S, Science, 301 (2003) 610–615. [PubMed: 12893935]
[33]. le Maire M, Champeil P, Moller JV, Biochim. Biophys. Acta, 1508 (2000) 86–111. [PubMed: 11090820]
[34]. Prive GG, Curr. Opin. Struct. Biol, 19 (2009) 379–385. [PubMed: 19682888]
[35]. Howell SC, Mittal R, Huang L, Travis B, Breyer RM, Sanders CR, Biochemistry, 49 (2010) 9572–9583. [PubMed: 20919740]
[36]. Wang X, Corin K, Baaske P, Wienken CJ, Jerabek-Willemsen M, Duhr S, Braun D, Zhang S, Proc Natl Acad Sci U S A, 108 (2011) 9049–9054. [PubMed: 21562213]
[37]. Yeh AP, McMillan A, Stowell MH, Acta Crystallogr D Biol Crystallogr, 62 (2006) 451–457. [PubMed: 16552147]

[38]. Alexandrov AI, Mileni M, Chien EY, Hanson MA, Stevens RC, Structure, 16 (2008) 351–359. [PubMed: 18334210]

[39]. Chiu ML, Nollert P, Loewen MC, Belrhali H, Pebay-Peyroula E, Rosenbusch JP, Landau EM, Acta Crystallogr D Biol Crystallogr, 56 (2000) 781–784. [PubMed: 10818364]

[40]. Cherezov V, Clogston J, Papiz MZ, Caffrey M, J Mol Biol, 357 (2006) 1605–1618. [PubMed: 16490208]

[41]. Caffrey M, Biochem Soc Trans, 39 (2011) 725–732. [PubMed: 21599641]

[42]. Zhuang J, Prive GG, Werner GE, Ringler P, Kaback HR, Engel A, J. Struct. Biol, 125 (1999) 63–75. [PubMed: 10196117]

[43]. Prive GG, Verner GE, Weitzman C, Zen KH, Eisenberg D, Kaback HR, Acta Crystallogr, D50 (1994) 375–379.

[44]. Byrne B, Abramson J, Jansson M, Holmgren E, Iwata S, Biochim. Biophys. Acta, 1459 (2000) 449–455. [PubMed: 11004462]

[45]. Xie H, Guo XM, Chen H, Mol. Biotechnol, 42 (2009) 135–145. [PubMed: 19199085]

[46]. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK, Science, 318 (2007) 1266–1273. [PubMed: 17962519]

[47]. Smyth DR, Mrozkiewicz MK, McGrath WJ, Listwan P, Kobe B, Protein Sci, 12 (2003) 1313–1322. [PubMed: 12824478]

[48]. Rosenbaum DM, Rasmussen SG, Kobilka BK, Nature, 459 (2009) 356–363. [PubMed: 19458711]

[49]. Uysal S, Vasquez V, Tereshko V, Esaki K, Fellouse FA, Sidhu SS, Koide S, Perozo E, Kossiakoff A, Proc. Natl. Acad. Sci. U S A, 106 (2009) 6644–6649. [PubMed: 19346472]

[50]. Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grutter MG, PLoS Biol, 5 (2007) e7. [PubMed: 17194213]

[51]. Lam AY, Pardon E, Korotkov KV, Hol WG, Steyaert J, J Struct Biol, 166 (2009) 8–15. [PubMed: 19118632]

[52]. Domanska K, Vanderhaegen S, Srinivasan V, Pardon E, Dupeux F, Marquez JA, Giorgetti S, Stoppini M, Wyns L, Bellotti V, Steyaert J, Proc Natl Acad Sci U S A, 108 1314–1319.

[53]. Prongay AJ, Smith TJ, Rossmann MG, Ehrlich LS, Carter CA, McClure J, Proc Natl Acad Sci U S A, 87 (1990) 9980–9984. [PubMed: 2124709]

[54]. Koldobskaya Y, Duguid EM, Shechner DM, Suslov NB, Ye J, Sidhu SS, Bartel DP, Koide S, Kossiakoff AA, Piccirilli JA, Nat Struct Mol Biol, 18 100–106.

[55]. Ye JD, Tereshko V, Frederiksen JK, Koide A, Fellouse FA, Sidhu SS, Koide S, Kossiakoff AA, Piccirilli JA, Proc Natl Acad Sci U S A, 105 (2008) 82–87. [PubMed: 18162543]

[56]. Arbabi-Ghahroudi M, Tanha J, MacKenzie R, Cancer Metastasis Rev, 24 (2005) 501–519. [PubMed: 16408159]

[57]. Harlow E, Lane D, Antibodies : a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988.

[58]. Kohler G, Milstein C, Nature, 256 (1975) 495–497. [PubMed: 1172191]

[59]. Day PW, Rasmussen SG, Parnot C, Fung JJ, Masood A, Kobilka TS, Yao XJ, Choi HJ, Weis WI, Rohrer DK, Kobilka BK, Nat Methods, 4 (2007) 927–929. [PubMed: 17952087]

[60]. Krebber A, Bornhauser S, Burmester J, Honegger A, Willuda J, Bosshard HR, Pluckthun A, J Immunol Methods, 201 (1997) 35–55. [PubMed: 9032408]

[61]. Kleymann G, Ostermeier C, Ludwig B, Skerra A, Michel H, Biotechnology (N Y), 13 (1995) 155–160. [PubMed: 9634756]

[62]. Ostermeier C, Essen LO, Michel H, Proteins, 21 (1995) 74–77. [PubMed: 7716172]

[63]. Ostermeier C, Harrenga A, Ermler U, Michel H, Proc Natl Acad Sci U S A, 94 (1997) 10547–10553. [PubMed: 9380672]

[64]. Hunte C, Koepke J, Lange C, Rossmanith T, Michel H, Structure, 8 (2000) 669–684. [PubMed: 10873857]

[65]. Lange C, Hunte C, Proc Natl Acad Sci U S A, 99 (2002) 2800–2805. [PubMed: 11880631]

[66]. Lee SY, Lee A, Chen J, MacKinnon R, Proc Natl Acad Sci U S A, 102 (2005) 15441–15446. [PubMed: 16223877]

[67]. Zhou YF, Morais-Cabral JH, Kaufman A, MacKinnon R, Nature, 414 (2001) 43–48. [PubMed: 11689936]

[68]. Derewenda ZS, Acta Crystallogr D Biol Crystallogr, 66 (2010) 604–615. [PubMed: 20445236]

[69]. Keenan RJ, Siehl DL, Gorton R, Castle LA, Proc Natl Acad Sci U S A, 102 (2005) 8887–8892. [PubMed: 15951425]

[70]. Maynard J, Petersson K, Wilson DH, Adams EJ, Blondelle SE, Boulanger MJ, Wilson DM, Garcia KC, Immunity, 22 (2005) 81–92. [PubMed: 15664161]

[71]. Jespers L, Schon O, Famm K, Winter G, Nat Biotechnol, 22 (2004) 1161–1165. [PubMed: 15300256]

[72]. Tereshko V, Uysal S, Koide A, Margalef K, Koide S, Kossiakoff AA, Protein Sci, 17 (2008) 1175–1187. [PubMed: 18445622]

[73]. Eigenbrot C, Ultsch M, Dubnovitsky A, Abrahmsen L, Hard T, Proc Natl Acad Sci U S A, 107 (2010) 15039–15044. [PubMed: 20696930]

[74]. Sennhauser G, Grutter MG, Structure, 16 (2008) 1443–1453. [PubMed: 18940601]

[75]. Schonfeld D, Matschiner G, Chatwell L, Trentmann S, Gille H, Hulsmeyer M, Brown N, Kaye PM, Schlehuber S, Hohlbaum AM, Skerra A, Proc Natl Acad Sci U S A, 106 (2009) 8198–8203. [PubMed: 19416843]

[76]. Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S, Proc Natl Acad Sci U S A, 104 (2007) 6632–6637. [PubMed: 17420456]

[77]. Pai JC, Culver JA, Drury JE, Motani RS, Lieberman RL, Maynard JA, Protein Eng Des Sel, 24 (2011) 419–428. [PubMed: 21217145]

[78]. Jung S, Pluckthun A, Protein Eng, 10 (1997) 959–966. [PubMed: 9415446]

[79]. Zahnd C, Amstutz P, Pluckthun A, Nat Methods, 4 (2007) 269–279. [PubMed: 17327848]

[80]. Boder ET, Wittrup KD, Nat Biotechnol, 15 (1997) 553–557. [PubMed: 9181578]

[81]. Daugherty PS, Chen G, Olsen MJ, Iverson BL, Georgiou G, Protein Eng, 11 (1998) 825–832. [PubMed: 9796833]

[82]. Smith GP, Petrenko VA, Chem Rev, 97 (1997) 391–410. [PubMed: 11848876]

[83]. Cobaugh CW, Almagro JC, Pogson M, Iverson B, Georgiou G, J. Mol. Biol, 378 (2008) 622–633. [PubMed: 18384812]

[84]. Sidhu SS, Kossiakoff AA, Curr. Opin. Chem. Biol, 11 (2007) 347–354. [PubMed: 17500026]

[85]. Fellouse FA, Esaki K, Birtalan S, Raptis D, Cancasci VJ, Koide A, Jhurani P, Vasser M, Weismann C, Kossiakoff AA, Koide S, Sidhu SS, J. Mol. Biol, 373 (2007) 924–940. [PubMed: 17825836]

[86]. Röthlisberger D, Pos KM, Pluckthun A, FEBS J, 264 (2004) 340–348.

[87]. Monroe N, Sennhauser G, Seeger MA, Briand C, Grutter MG, J. Struct. Biol, 174 (2011) 269–281. [PubMed: 21296164]

[88]. Kaufmann M, Lindner P, Honegger A, Blank K, Tschopp M, Capitani G, Pluckthun A, Grutter MG, J Mol Biol, 318 (2002) 135–147. [PubMed: 12054774]

[89]. Roosild TP, Castronovo S, Choe S, Acta Crystallogr. F Struct. Biol. Cryst. Comun, 62 (2006) 835–839.

[90]. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC, Science, 322 (2008) 1211–1217. [PubMed: 18832607]

[91]. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC, Science, 318 (2007) 1258–1265. [PubMed: 17962520]

[92]. Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC, Science, 330 (2010) 1066–1071. [PubMed: 20929726]

[93]. Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC, Science, 330 (2010) 1091–1095. [PubMed: 21097933]

[94]. Fang Y, Jayaram H, Shane T, Kolmakova-Partensky L, Wu F, Williams C, Xiong Y, Miller C, Nature, 460 (2009) 1040–1043. [PubMed: 19578361]

[95]. Shaffer PL, Goehring A, Shankaranarayanan A, Gouaux E, Science, 325 (2009) 1010–1014. [PubMed: 19608859]

[96]. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK, Nature, 450 (2007) 383–387. [PubMed: 17952055]

[97]. Inaba K, Murakami S, Nakagawa A, Iida H, Kinjo M, Ito K, Suzuki M, EMBO J, 28 (2009) 779–791. [PubMed: 19214188]

[98]. Jiang Y, Lee A, Chen J, Ruta V, Cadene M, Chait BT, MacKinnon R, Nature, 423 (2003) 33–41. [PubMed: 12721618]

[99]. Hino T, Matsumoto Y, Nagano S, Sugimoto H, Fukumori Y, Murata T, Iwata S, Shiro Y, Science, 330 (2010) 1666–1670. [PubMed: 21109633]

[100]. Tsukazaki T, Mori H, Fukai S, Ishitani R, Mori T, Dohmae N, Perederina A, Sugita Y, Vassylyev DG, Ito K, Nureki O, Nature, 455 (2008) 988–991. [PubMed: 18923527]

[101]. Doyle DA, Morais Cabral J, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT, MacKinnon R, Science, 280 (1998) 69–77. [PubMed: 9525859]

[102]. Rasmussen SG, Choi HJ, Fung JJ, Pardon E, Casarosa P, Chae PS, Devree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK, Nature, 469 (2011) 175–180. [PubMed: 21228869]

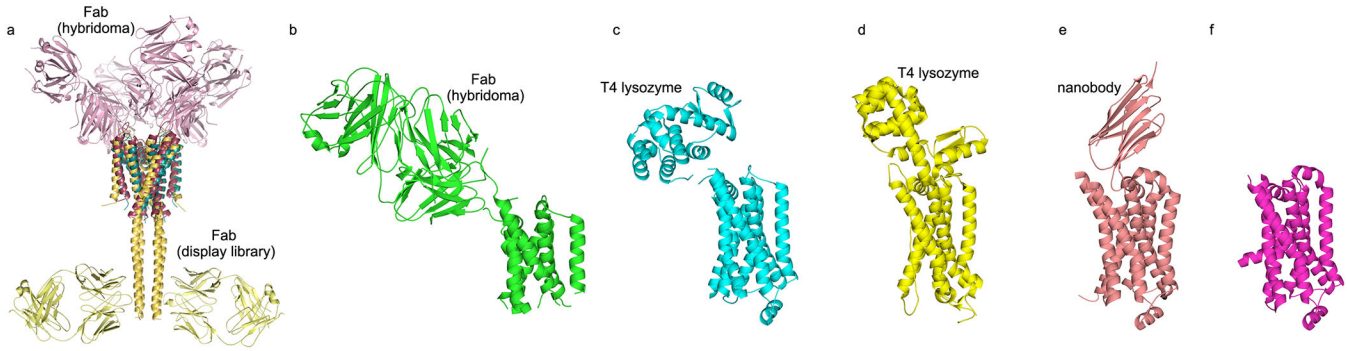

Figure 1. Comparison of selected membrane protein structures solved with and without chaperones.

(a) Superposition of KcsA structures. Green, structure solved without chaperones (PDB code 1BL8) [103]; pink, structure solved in high potassium salt, in complex with Fab obtained via hybridoma screening to recognize the functional KcsA tetramer (PDB code 1K4C)[68]; yellow, full-length KcsA structure solved in complex with Fab identified by molecular display technologies (PDB code 3EFF)[50]. (b) β_2AR GPCR structure solved in complex with Fab fragment derived from hybridoma (PDB code 2R4R)[98]. (c-d) β_2AR (PDB code 2RH1) [92] and A_2AR (PDB code 3EML) [91] structures solved with covalent chaperone T4 lysozyme, respectively (e) β_2AR structure solved in complex with nanobody (PDB code 3P0G) [104] (f) β_1AR structure solved without chaperone assistance but including various modifications (PDB code 2VT4) [24]

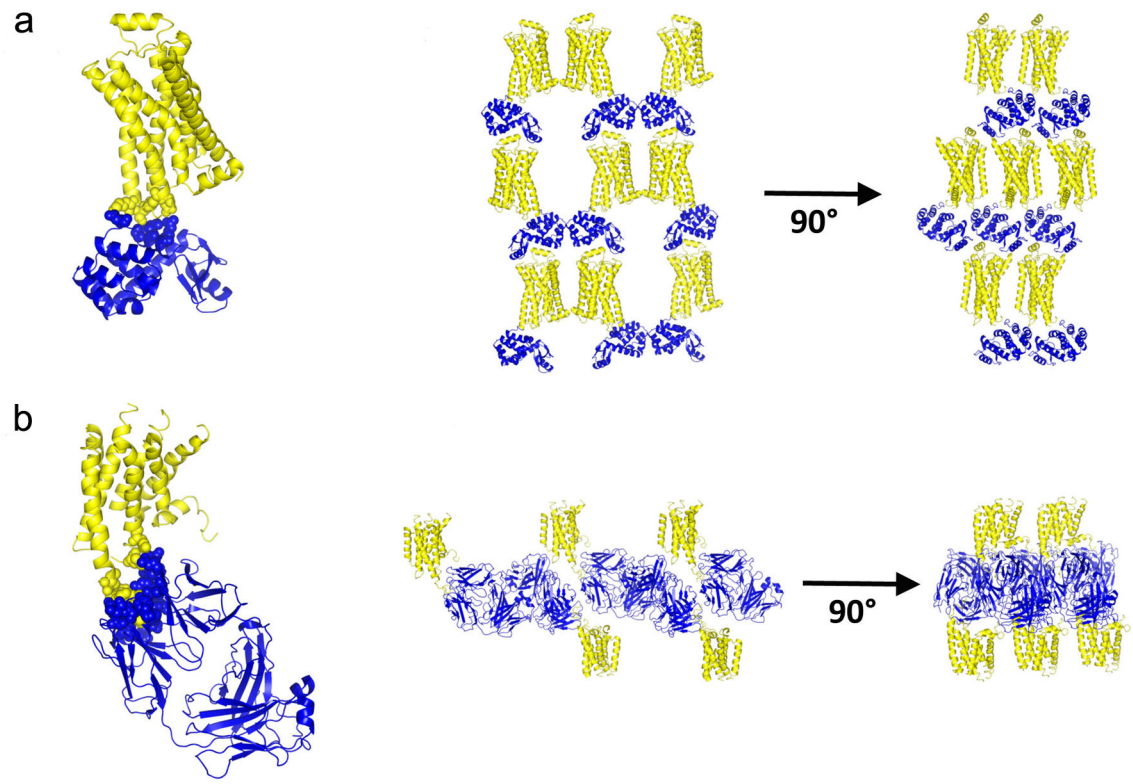

Figure 2. Crystal packing of β_2AR (yellow) with crystallization chaperones (blue).

(a) Covalent chaperone (PDB code 2RH1) [92]. T4 lysozyme is inserted into the third intracellular loop (left). Four major crystal contacts occur in T4 lysozyme-β_2AR lattice, three of which involve T4 lysozyme (right). (b) Non-covalent chaperone (PDB code 2R4R) [98], utilizing a Fab fragment specific for the third intracellular loop (left). The Fab fragment mediates nearly all crystal contacts within the lattice containing the Fab-β_2AR complex (right).

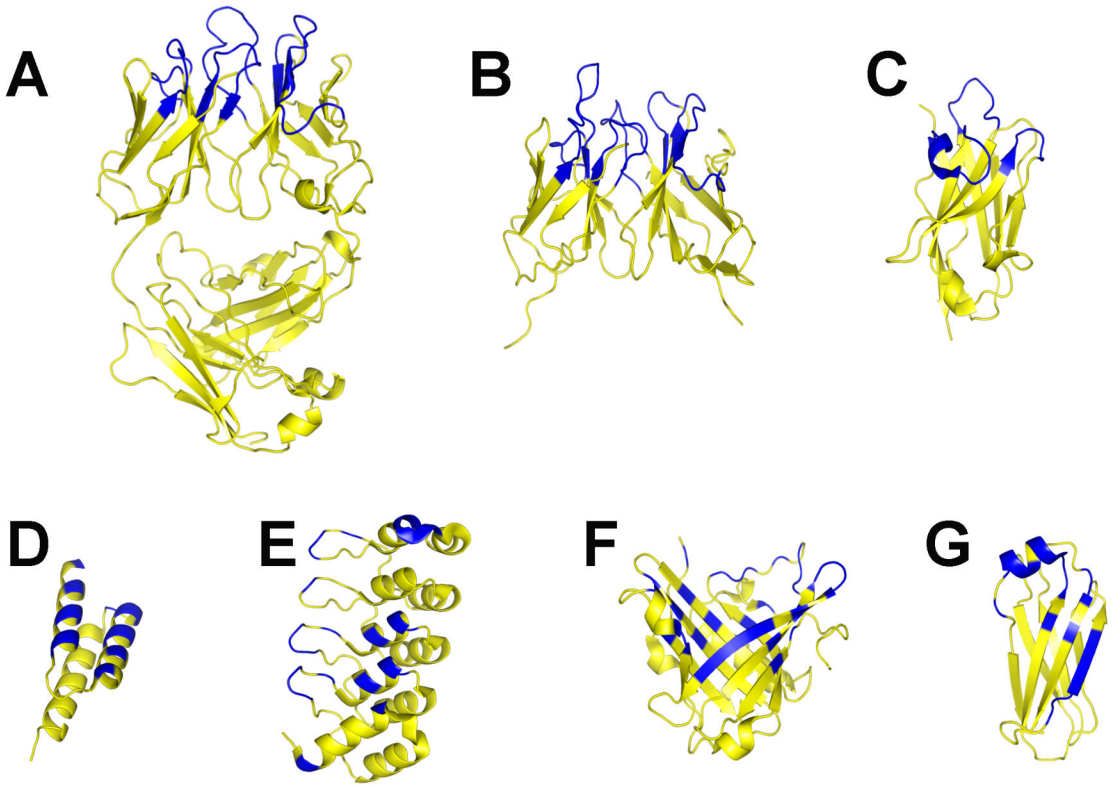

Figure 3. Chaperone platforms for membrane protein co-crystallization (yellow), depicting relative sizes and binding surfaces (blue).

(a) Fab fragment (PDB code 3EFF) [50]. (b) scFv (PDB code 3NN8) [78]. (c) Camelid nanobody (PDB code 2P42) [73]. (d) Affibody (PDB code 3MZW) [74]. (e) DARPin (PDB code 2J8S) [51]. (f) Anticalin (PDB code 3BX7) [76]. (g) FN3 domain (PDB code 2OCF).

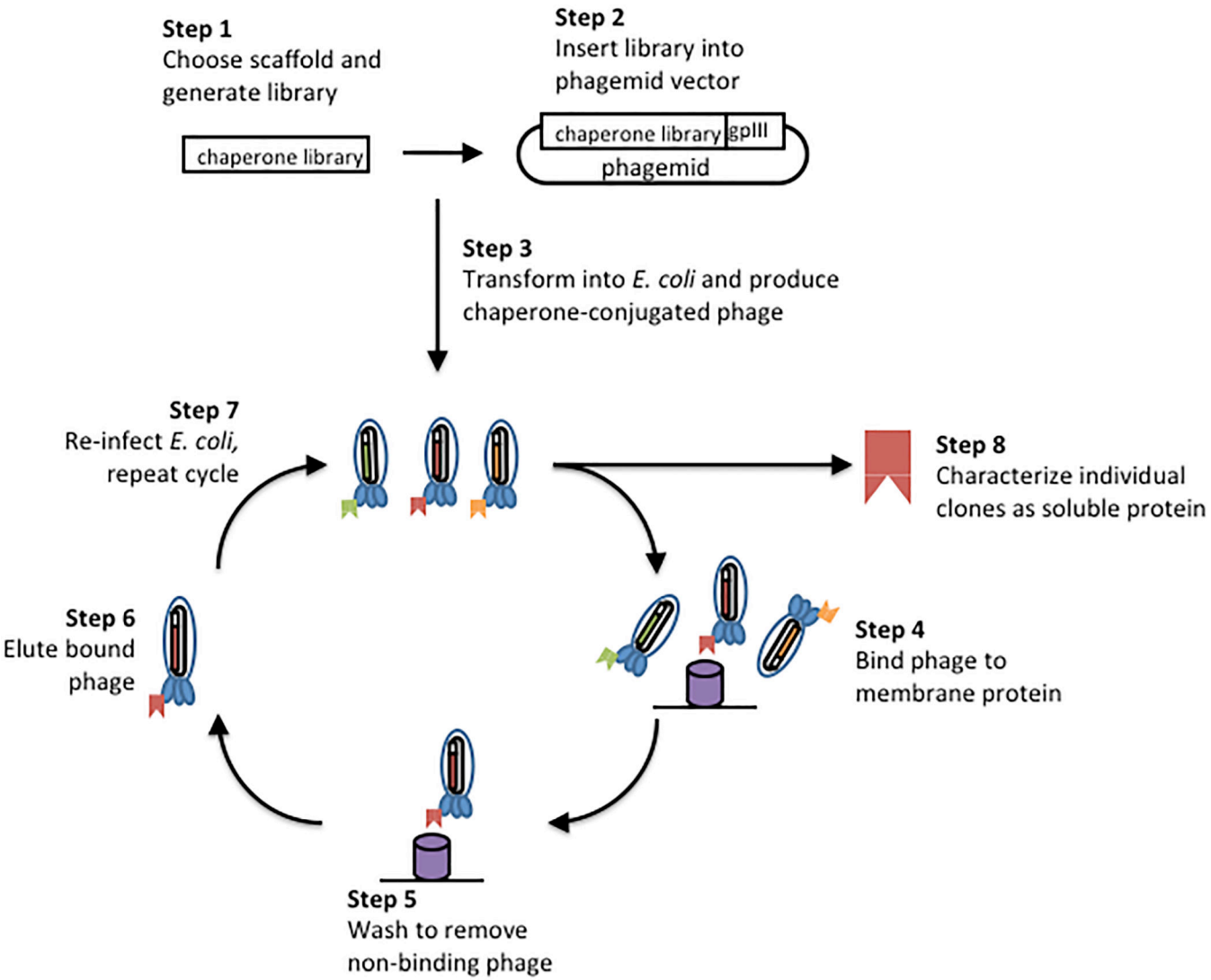

Figure 4. Phage display procedure for crystallization chaperone selection.

After selecting a chaperone and generating a library to engineer novel specificity, the chaperone library is inserted into a phagemid vector, which displays the chaperone as a fusion construct with a phage surface protein, such as the coat protein pIII. Phage expressed in *E. coli* contain the phagemid that encodes for the chaperone variant on the surface of the phage. The library of chaperone-conjugated phage is panned over the target membrane protein, and iterative rounds of binding, washing, elution and re-infection enrich the library for variants that bind the membrane protein. Hundreds of clones can be easily screened after multiple rounds of selection, and clones can be further optimized using additional rounds of phage display as desired.

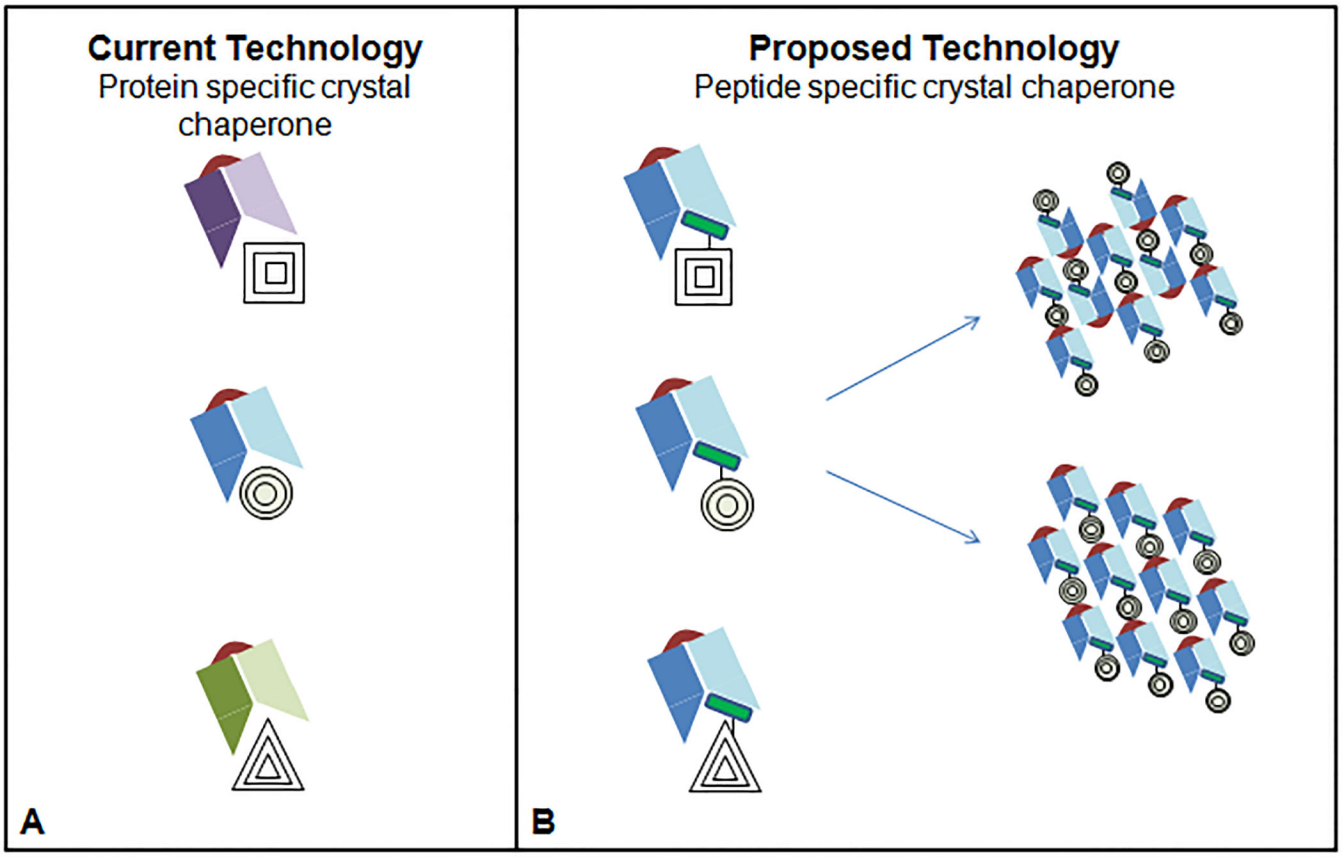

Figure 5. Comparison of the current methods to engineer non-covalent crystallization chaperones.

(a) Current hybridoma or molecular display methods are used to engineer a single chaperone to bind a single membrane protein. The chaperone cannot easily be adapted to alternative proteins, limiting the use of this method for high-throughput structural biology applications, as a new chaperone must be engineered for each additional protein. (b) An alternative is to engineer a chaperone to bind a peptide, which can be readily introduced into a construct. This would enable the use of a single chaperone for co-crystallization with many different membrane proteins.

Table 1.

Summary of strategies used to crystallize membrane proteins at each stage of the pipeline.

Stage in Pipeline	Strategy	Rationale	Caveats/Comments
Molecular Biology and Protein Expression	Bacterial orthologs	Bacterial proteins express better in *E. coli*, can be stable in a range of conditions, are not glycosylated, and structures are likely very similar	Bacterial proteins can differ somewhat in structure, especially if modular domains are missing/present; function may differ enough from human homolog to limit drug discovery
	Stabilizing mutations	Less conformational heterogeneity will promote crystal growth	Non-native structure will be solved, flexibility that may be relevant to function will be masked
	Expression platform	Yield and stability can be vastly different for expression platforms, both of which are important parameters for future success in structure determination	Human proteins expressed in *E. coli* will lack glycosylation, which may be functionally relevant; other expression platforms may yield insufficient quantities of protein even for low volume crystallization screening
Membrane Protein Purification	Detergent selection for membrane solubilization	Yield is important for adequate crystallization trials, and maintaining function in detergent is also a good prognosticator	Many detergents exist and certain detergents have been used successfully for several membrane proteins, finding a suitable detergent for any protein of interest is unpredictable. In addition, a functional assay may not be available but proxy biophysical properties such as stability and circular dichroism spectra can be substituted
	Detergent selection/ combination for crystallization	Optimized detergent for solubilization may not be appropriate for crystallization but detergents can be exchanged by immobilizing the protein on a column and eluting with a new buffer	Each detergent examined for crystallization must be rescreened to identify conditions
Membrane Protein Crystallization	Lipid phases	Crystallization in a membrane environment is likely to be more stabilizing and facilitate two-dimensional ordering of the membrane protein lattice	Some specialized equipment (commercially available, relatively inexpensive) is needed to reliably generate cubic phases
	Ligand or other additive	Less conformational heterogeneity and increase in stability facilitate crystallization	In many cases, a good ligand is unknown
	Covalent: gene for chaperone inserted into membrane protein construct	Exquisite control over placement, can replace a disordered loop with more compact three dimensional soluble protein	Limited knowledge of the membrane protein renders this process trial-and-error; limited number of successful different examples of proteins utilizing this approach
	Non-covalent: high affinity complex obtained by hybridoma or library screening (antibody, non-antibody platforms)	Highly crystallizable fragment provides lattice contacts, library screening methods offer medium to high-throughput	Hybridoma technology is expensive and low-throughput but newer methods are increasingly accessible and generalizable

Table 2.

List of membrane protein structures utilizing a crystallization chaperone.

Membrane Protein	Chaperone	PDB code	Resolution (Å)	Reference
Covalent Chaperones				
A_{2A} adenosine receptor (A_2AR)	T4 Lysozyme	3EML	2.6	[91]
A_2AR	T4 Lysozyme	3QAK	2.7	[23]
β_2 adrenergic receptor (β_2AR)	T4 Lysozyme	2RH1	2.4	[92]
β_2AR-Gs protein complex	T4 Lysozyme	3SN6	3.2	[49]
	Nanobody			
CXCR4 chemokine receptor	T4 Lysozyme	3ODU	2.5	[93]
Dopamine D3 receptor	T4 Lysozyme	3PBL	2.9	[94]
Histamine H_1 receptor	T4 Lysozyme	3RZE	3.1	[95]
Noncovalent Chaperones				
AcrB multidrug exporter	DARPin*	2J8S	2.5	[51]
AdiC arginine:agmatine antiporter	Fab	3NCY	3.2	[96]
ApcT amino acid transporter	Fab	3GI9	2.5	[97]
β_2AR	Fab	2R4R	3.4	[98]
β_2AR-Gs protein complex	T4 Lysozyme	3SN6	3.2	[49]
	Nanobody			
Cytochrome bc_1 complex	Fv*	1EZV	2.3	[65]
Cytochrome *c* oxidase	Fv*	1AR1	2.7	[64]
Disulfide bond formation protein B (DsbB)	Fab	2ZUQ	3.3	[99]
KcsA K^+ channel (open)	Fab	1K4C	2.0	[68]
(closed)	Fab*	3EFD	3.8	[50]
KvAP voltage-gated potassium channel	Fab	1ORQ	3.2	[100]
	Fv*	2A0L	3.9	[67]
Nitric oxide reductase	Fab	3O0R	2.7	[101]
SecYE protein-conducting channel	Fab	2ZJS	3.2	[102]

*Chaperone obtained by recombinant methods

Table 3.

Outline of chaperone development using phage display for an experienced user.

Step	Considerations
Week 1. Generation of Chaperone Library	
1. Select a chaperone scaffold format	Chaperone expression level & host, size, compatibility with different selection platforms, biophysical characteristics.
2. Clone the chaperone into a phagemid vector	Phagemid vector should be chosen based on desired monovalent or multivalent display
3. Generate a library of mutations based on the chaperone	Site-directed random mutagenesis of the binding site is an efficient method to engineer novel specificity and affinity for a chaperone. Library diversity can be limited so that the theoretical diversity can be efficiently sampled during panning.
4. Express & purify solubilized, native membrane protein for selection	Membrane protein can be locked in a particular state by complexation with agonist or antagonist to stabilize this state during chaperone selection and ultimately, crystallization.
Week 2–3. Clone Selection by Phage Panning	
5. Transform phagemid library into *E. coli*	Maximize library size to be able to efficiently sample library diversity
6. Express chaperone-conjugated phage, incubate with membrane protein	The membrane protein can be immobilized in a plastic ELISA well for solid-phase selection or biotinylated for solution-phase selection followed by recovery using streptavidin-coated beads. Wash to remove non- and weakly-bound phage.
7. Elute bound phage, infect *E. coli* to repeat selection cycle	Repeat selection for additional rounds as necessary (often 3–5). Selection can be monitored by phage titer and/ or DNA sequencing. After several rounds, the recovered titer will typically increase and sequencing will reveal the presence of a few highly represented clones.
Week 4–6. Characterization of Chaperone Candidate	
8. Screen selected clones from phage panning	Hundreds of clones can be efficiently screened using high-throughput methods, such as growth and binding assays in 96-well plates.
9. Express selected variants as soluble proteins	A few chaperone candidates are sequenced and sub-cloned for expression as a soluble chaperone.
10. Measure affinity and specificity for membrane protein	Candidate chaperones are characterized.
11. Optimize chaperone as necessary	Random mutagenesis of a lead chaperone can be used to improve both affinity and biophysical characteristics [77].
Week 7-?. Optimization and Screening of Co-crystallization Conditions	