

Review

# Protein Fusion Strategies for Membrane Protein Stabilization and Crystal Structure Determination

Shixuan Liu and Weikai Li \*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA; liu513@wustl.edu

\* Correspondence: weikai@wustl.edu

**Abstract:** Crystal structures of membrane proteins are highly desired for their use in the mechanistic understanding of their functions and the designing of new drugs. However, obtaining the membrane protein structures is difficult. One way to overcome this challenge is with protein fusion methods, which have been successfully used to determine the structures of many membrane proteins, including receptors, enzymes and adhesion molecules. Existing fusion strategies can be categorized into the N or C terminal fusion, the insertion fusion and the termini restraining. The fusions facilitate protein expression, purification, crystallization and phase determination. Successful applications often require further optimization of protein fusion linkers and interactions, whose design can be facilitated by a shared helix strategy and by AlphaFold prediction in the future.

**Keywords:** membrane protein; crystallization; fusion protein; termini restraining; insertion fusion; GPCR; VKOR



**Citation:** Liu, S.; Li, W. Protein Fusion Strategies for Membrane Protein Stabilization and Crystal Structure Determination. *Crystals* **2022**, *12*, 1041. <https://doi.org/10.3390/cryst12081041>

Academic Editors: Albert Guskov, Qun Liu, Dianfan Li, Youzhong Guo and Michael C. Wiener

Received: 1 July 2022

Accepted: 21 July 2022

Published: 27 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Membrane proteins participate in various cellular processes and are major drug targets [1,2]. The structures of these proteins are highly desired for mechanistic understandings of their function and for applications such as drug design [3], engineering for improved or new function [4], and developing biotechnology tools [5] (e.g., nanopore sequencing). However, it is often challenging to obtain crystal structures of membrane proteins.

The difficulties reside in every step of the structure determination process, from protein expression, purification, crystallization to phase determination. Many membrane proteins either have a low expression level or become unstable after detergent extraction, a step usually required for further protein purification [6]. Consequently, it is difficult to obtain membrane proteins in a large quantity and high purity, a prerequisite for protein crystallization. Crystallization is hindered because the membrane regions of these proteins are buried in the detergent micelles and their exposed extramembrane regions are often small or flexible; these regions cannot provide sufficient interactions to form crystal contacts. Consequently, the crystals are either not formed or are of low quality, and high-resolution structures cannot be obtained.

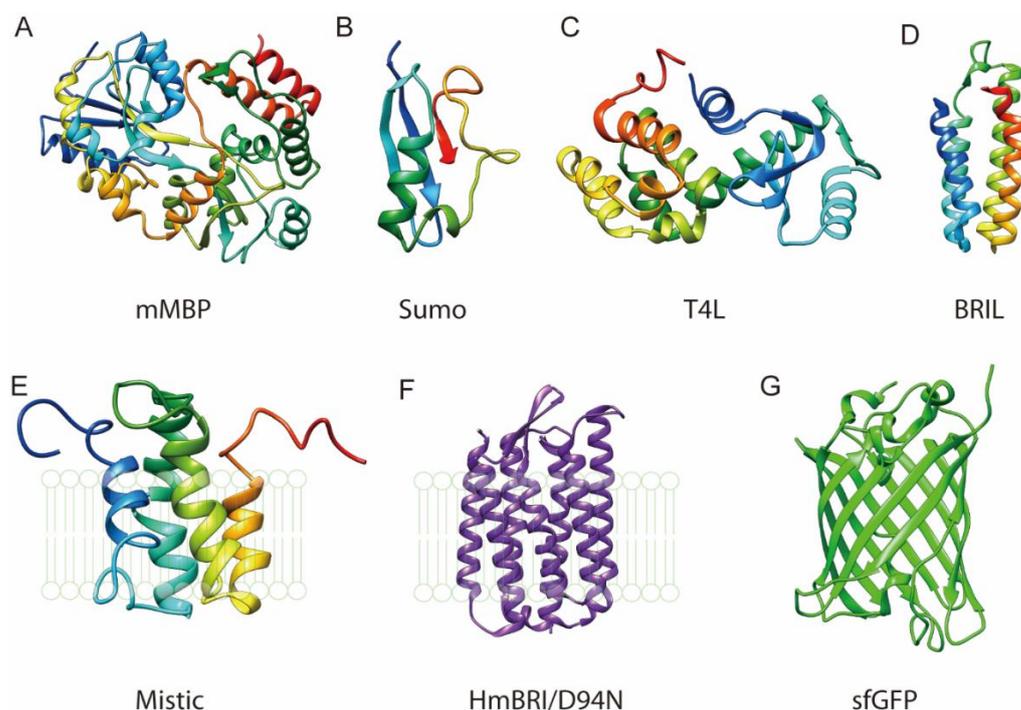
To overcome these difficulties, a highly stable and crystallizable protein is often fused to the membrane protein target to assist its expression, purification and crystallization. Here we will review existing strategies of protein fusion and their advantages, including a recently developed fusion method called termini restraining.

## 2. Using Fusion Protein to Facilitate Membrane Protein Production

One of the bottlenecks for membrane protein crystallization is to obtain proteins that are stable, pure and homogeneous. Membrane proteins with such high qualities are prone to crystallize with good diffraction [7,8]. A traditional way to improve the protein stability and expression level is by systematic mutagenesis and serial truncations of membrane

proteins [9]. In comparison, employing protein fusion is more straightforward and usually requires lower workload.

Relatively large proteins, such as maltose binding protein (MBP) [10] (Figure 1A), has considerable success in improving the expression level and facilitating the purification of both membrane and soluble proteins. Hu et al [11] fused MBP to the N terminus (N-MBP) of 42 membrane proteins from *Escherichia coli* (*E. coli*) and found that most of the N-MBP fused membrane proteins could be expressed in *E. coli*. In contrast, the same proteins with an N terminal His tag showed poor or no expression. MBP has also been used for the expression of G-coupled protein receptors (GPCR) that are of eukaryotic origin. Yeliseev et al. [12] showed that the cannabinoid receptor CB2 (CB2) with N-MBP fusion could be expressed at high level in *E. coli* and with good recovery of the folded protein. Likewise, Shibata et al. [13] showed *E. coli* expression of NTS1 receptors with N-MBP fusion.



**Figure 1.** Various proteins that can be fused to membrane proteins to facilitate their expression and purification. (A) *E. coli* mature Maltose binding protein (MBP), PDB 1ANF. (B) *S. cerevisiae* Ubiquitin-like protein SMT3 (SUMO), PDB 3TIX, Chain A. (C) *E. virus T4* Lysozyme (T4L), PDB 2LZM. (D) *E. coli* B562RIL (BRIL), PDB 1M6T. (E) *B. subtilis* Mystic, PDB 1YGM. (F) *H. marismortui* bacteriorhodopsin I D94N mutant (HmBRI/D94N). (G) *A. victoria* Superfolder GFP (sfGFP), PDB 2B3P. MBP, SUMO, T4L and BRIL are soluble proteins. MBP and SUMO are usually fused to the N terminus of the target membrane protein. T4L and BRIL are usually inserted in a membrane protein between its two TMs (e.g., TM5 and TM6 of a GPCR). Mystic and HmBRI/D94N are membrane proteins. HmBRI/D94N has a purple color and sfGFP has green fluorescence.

A natural chaperonin protein, small ubiquitin-related modifier (SUMO) [14,15] (Figure 1B), has been shown to facilitate the folding and enhance the stability and solubility for both prokaryotic and eukaryotic membrane proteins. Zuo et al. [16] found that SUMO fusion largely enhanced the *E. coli* expression of membrane proteins from severe acute respiratory syndrome coronavirus (SARS-CoV). Other small proteins, T4 lysozyme [17] (T4L) (Figure 1C) and thermostabilized cytochrome b562 with M7W/H102I/R106L mutations (BRIL) [18,19] (Figure 1D), are classical fusion proteins used to improve the stability of various GPCR proteins and facilitate their crystallization.

Fusion with a stable membrane protein has been used to improve the behaviors of other membrane proteins. Membrane-integrating sequence for translation of integral

membrane protein constructs (Mistic) [20–22] (Figure 1E), a small integral membrane protein originated from *Bacillus subtilis*, has been used as a fusion tag for the heterologous expression of membrane proteins in *E. coli* for their folding and membrane integration. A bacteriorhodopsin, HmBRI/D94N [23,24] (Figure 1F), from *Haloarcula marismortui*, which contains 7 transmembrane helices (TMs), has also been used for protein fusion. Hsu et al. [24] tested the expression level of undecaprenyl pyrophosphate phosphatase and carnitine/butyrobetaine antiporter with the N-terminal HmBRI/D94N fusion. The fusion proteins showed significant elevation in their expression level, by 50- and 17-fold (4.9 and 3.4 mg/L), respectively, compared to His tagged proteins. The HmBRI/D94N fusion has another advantage that this fused bacteriorhodopsin shows purple color. Thus, the expression level of target membrane proteins can be readily determined with a spectrometer or even the naked eye. Fusion of a colored protein largely facilitates the purification process, which is known as visible protein purification.

Protein expression and purification can also be ‘visualized’ by the use of fluorescent proteins, such as the green fluorescent protein (GFP) [7,25,26] or a red fluorescent protein, mCherry [27]. Superfolder GFP (sfGFP) [25] (Figure 1G) has been engineered to higher stability than the original GFP, a property highly useful for expressing challenging membrane proteins. With the fused GFP, fluorescence-detection size-exclusion chromatography (FSEC) has been developed for fast screening of the expression level and detergent stability of membrane proteins. FSEC can be applied to cells solubilized in detergent, and with the GFP fluorescence signal, the elution profile of the detergent extracted membrane protein on size-exclusion chromatography can be obtained to judge protein stability and homogeneity. Because this procedure avoids the large effort of protein purification, many protein homologs or constructs can be screened in a relatively short time.

The soluble or membrane proteins used for fusion (named scaffold) are often introduced at the N terminus of the target membrane proteins to improve their expression level, because their N terminal sequence is often a determinant of the protein stability and the fusion may protect degradation [28]. On the other hand, protein visualization tags (e.g., GFP) are often fused to the C terminus of the membrane proteins to minimize the disturbance to their function [7,8]. Moreover, the linker between the scaffold protein and membrane protein target may need to be optimized to maintain the membrane protein folding.

Adding to the toolbox, we recently developed a termini restraining method in which two self-associable protein entities (named coupler) [29,30], such as the two halves of a split sfGFP, are attached to the N and C terminus of membrane proteins with an even number of TMs. If the membrane protein folds correctly, the two halves of sfGFP reassemble into a functional protein that exhibits fluorescence. The coupler restricts the drastic motion between TMs that occurs during protein unfolding, thereby favoring membrane proteins in their stably folded state. On the other hand, the movements are restrained in a relatively minor, reversible manner, allowing protein conformational changes that are often required for their function. The stable, well folded proteins are less prone to be degraded by the cell quality control machineries and better survive the detergent extraction. In addition, the sfGFP allows visible protein purification.

FSEC of the termini restrained membrane proteins showed large improvement in the recovery of folded proteins. Over 100 folds of improvement were observed for human VKOR [29,31] termini restrained by sfGFP, compared to VKOR with a C-terminal sfGFP tag. After purification, the termini restrained VKOR protein showed high yield and monodispersed peak on size exclusion chromatography that suggested high protein homogeneity. The specific activity of restrained VKOR was much higher than the unrestrained protein. Similarly, a VKOR-like protein from *T. rubripes* showed tens of folds of improvement in the yield of purified protein [29,31]. When viewed under confocal fluorescence microscope, these endoplasmic reticulum (ER) membrane proteins located at the ER membrane boundaries, whereas the C-sfGFP tagged proteins tended to aggregate in the ER lumen or vacuoles. FSEC-based thermostability assay showed that the melting temperature ( $T_m$ ) of

the restrained VKOR and VKOR-like increased by 2.2 °C and 8.9 °C, respectively. Taken together, termini restraining improves protein folding and stability, resulting in high yield and quality.

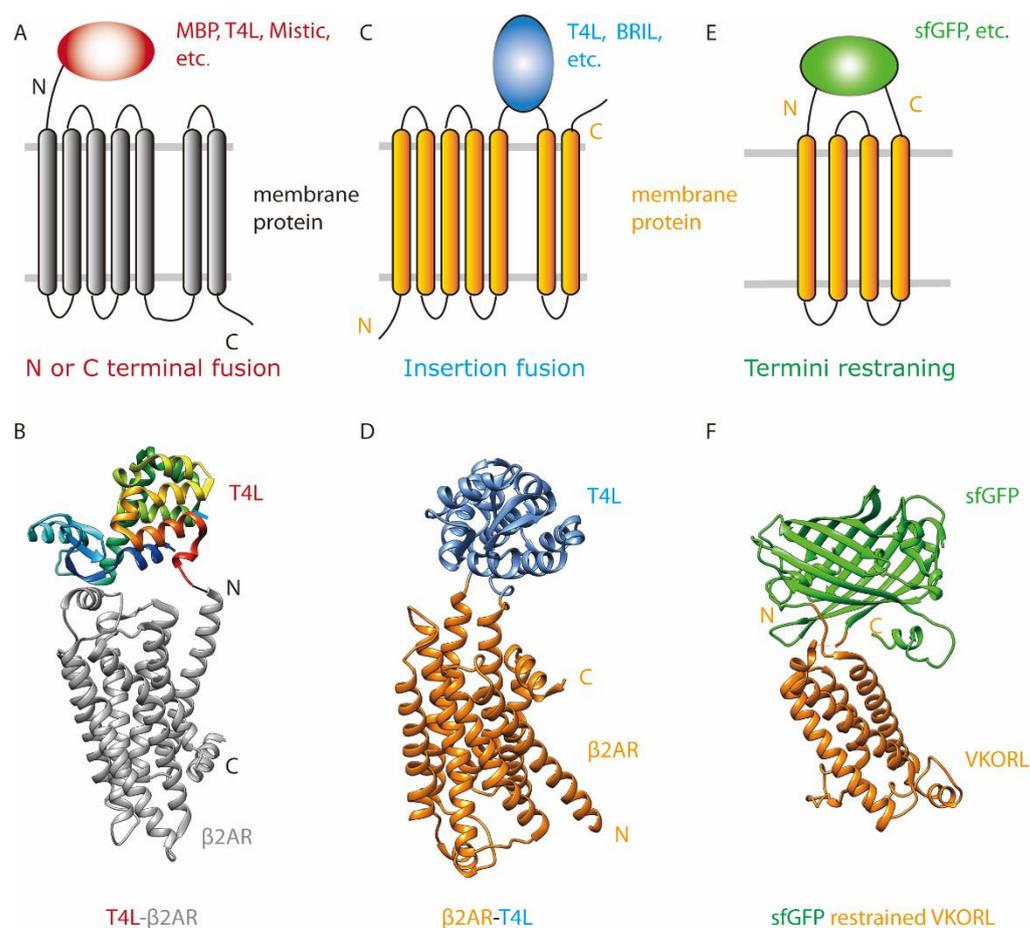
### 3. Fusion Strategies for Membrane Protein Crystallization and Structure Determination

Although many fusion proteins have been used to improve the expression and folding of membrane proteins, examples of using fusion proteins to facilitate the crystallization process of membrane proteins are not as many. The fusion strategies for promoting crystallization can be categorized into the N or C terminal fusion, the insertion fusion, and the termini restraining, which introduce the fusion protein in different ways. The fused soluble protein affords a large and stable surface to form crystal contacts, thereby serving as the crystallization ‘scaffold’.

#### 3.1. Fusion to the N Terminus or C Terminus

The most common fusion strategy is adding a stable scaffold protein to one end of the target protein (Figure 2A). Numerous soluble proteins have been crystallized with the N or C terminal fused MBP [10], glutathione S-transferase (GST) [32,33] and other scaffold proteins. There are less reported cases for membrane proteins. For GPCR crystallization, Zou et al. [34] fused T4L to the N terminus of  $\beta_2$  adrenergic receptor (T4L- $\beta_2$ AR) (Figure 2B). They proposed that for the T4L- $\beta_2$ AR fusion protein to crystallize and remain functional, the linker sequence between T4L and  $\beta_2$ AR needs to be short and relatively rigid. Thus, they tested a series of truncations and alanine insertions between the C terminus of T4L and the N terminus of  $\beta_2$ AR. The construct with the highest expression level was obtained by fusing T4L to D29 of  $\beta_2$ AR with a linker of two Ala residues in between. This construct allowed the crystal structure determination of the  $\beta_2$ AR-Gs complex [35]. However, most of the crystal lattice contacts in the structure were mediated by the Gs protein, whereas no contacts were formed between T4L and the extracellular part of its fused  $\beta_2$ AR. The authors speculated that T4L- $\beta_2$ AR fusion was not sufficiently constrained to facilitate crystallization. Thus, they made further modifications to minimize the unstructured region of  $\beta_2$ AR. With the deletion of residues 235 to 263 in the third intracellular loop (ICL3) of  $\beta_2$ AR and the deletion of 366 to end, they obtained the crystal structure of the T4L- $\beta_2$ AR- $\Delta$ ICL3 with an inverse agonist, carazolol. Although the structure was at relatively low resolution (4.0 Å), electron densities were observed for the Ala-Ala linker between T4L and  $\beta_2$ AR. The crystal contacts were formed mainly between T4L molecules, but not between adjacent  $\beta_2$ AR molecules, indicating that the T4L scaffold was essential for the crystallization.

Another example of terminal fusion is the complex structure of human rhodopsin and mouse visual arrestin [36]. In this structure, an arrestin mutant (L374A/V375A/F376A, 3A arrestin) was fused to the C terminus of a rhodopsin mutant (E113Q, M257Y, N2C and N282C) with a 15 AA linker and T4L was fused to the N terminus of the rhodopsin mutant. This sandwich fusion protein, T4L-rhodopsin-arrestin, is monomeric and has a  $T_m$  of 59 °C, suggesting that it is relatively stable and is promising for crystallization. Negative stain electron microscopy confirmed that the rhodopsin mutant formed a stable complex with the C terminal fused 3A arrestin. The sandwich fusion protein was crystallized with the lipid cubic phase (LCP) method. The crystals are of small size (5 to 15  $\mu$ m) and diffracted only to 6–8 Å at a synchrotron beamline. The structure was obtained with the use of serial femtosecond crystallography with an LCP injector. Overall, this sandwich strategy used the 3A arrestin to bind the rhodopsin mutant and stabilize its conformation, along with the T4L that serves as the scaffold for crystal packing.



**Figure 2.** Protein fusion methods for the crystallization of membrane proteins. Membrane proteins are colored in grey or orange and the scaffold proteins in red, blue or green. (A) Scheme of N or C terminal fusion. The scaffold protein is fused to one end of a membrane protein. (B) Example of terminal fusion. T4L is fused to the N-terminus of the  $\beta 2$  adrenergic receptor (T4L- $\beta 2$ AR). PDB: 4GBR. (C) Scheme of insertion fusion. The scaffold protein is fused between the TMs. (D) Example of insertion fusion. T4L is fused between the TM5 and TM6 of human  $\beta 2$  adrenergic receptor ( $\beta 2$ AR-T4L). PDB: 2RH1. (E) Scheme of termini restraining. The scaffold protein is fused to the two termini of the target membrane protein. (F) Example of termini restraining. The sfGFP is fused to the two termini of VKOR-like. PDB: 6WVI.

### 3.2. Insertion Fusion: Scaffold Fusion between Two Transmembrane Helices

Inserting the scaffold protein between two TMs of the target protein is widely used for GPCR crystallization [37] (Figure 2C). The method was first reported for the crystallization of  $\beta 2$ AR [17], in which T4L was inserted between TM5 and TM6, with T4L residues 2–161 replacing ICL3 (231–262) of  $\beta 2$ AR (Figure 2D). Because the ICL3 loop is highly flexible, replacing it with T4L is unlikely to disturb the rest of the  $\beta 2$ AR structure. The  $\beta 2$ AR-T4L fusion protein was crystallized in LCP with a bound ligand, carazolol [38], and the structure was determined to 2.4 Å. The fused T4L and  $\beta 2$ AR molecules form two layers of crystal packing. The T4L scaffold provide one layer of packing, and  $\beta 2$ AR form another layer of type I packing using interactions between TMs. Between the two layers, T4L interacts with the extracellular loops, ECL2 and ECL3, of  $\beta 2$ AR to form crystal contacts.

The successful determination of  $\beta 2$ AR-T4L structure has led to the burst of GPCR structures. To facilitate crystallization of GPCRs, T4L was further engineered to “disulfide stabilized T4L” (dsT4L) [39] and “minimal T4L” (mT4L). Compared to T4L, which contains C54T and C97A mutations, DsT4L introduced four additional cysteines, I3C, T21C, A97C and T142C, which form two disulfide bonds, C3-C97 and C21-C142. These disulfide links

stabilize a relatively closed conformation of T4L and increase the thermal stability of this scaffold protein. As to mT4L, the small and flexible N-terminal lobe of T4L was replaced with a short GGSGG linker.

BRIL is a four-helix bundle protein widely used for GPCR crystallization [19,40]. BRIL replacement of ICL3 improved the protein stability of GPCRs, and the elution profile of GPCR-BRIL fusion on size exclusion chromatography was often better than that of GPCR-T4L fusion. BRIL fusion has been reported also for the structure determination of a copper transporter [41]. Other small proteins can also be used to replace ICL3 for GPCR crystallization. For instance, fusion of a glycogen synthase from *Pyrococcus abyssi* (PGS, 196 AA) enabled crystallization of orexin (also known as hypocretin) [42] when T4L fusion failed.

### 3.3. Termini Restraining

In this new strategy, two parts of a self-associable scaffold protein were fused separately to the N and C termini of a membrane protein [29] (Figure 2E). This method requires the membrane protein to have even number of TMs. If the membrane protein folds in the expected topology, two parts of the scaffold protein reassemble to an integral protein. Unlike T4L insertion, which places the scaffold protein inside the target membrane protein, termini restraining places the entire target membrane protein inside the stable scaffold protein (Figure 2C,E). As a result, termini restraining stabilizes the overall structure of a membrane protein, whereas the insertion fusion stabilizes its local structure.

Based on the known structure of the scaffold protein, a 'split' site for membrane protein insertion is selected. For initial analysis, the full-length membrane protein is generally inserted into this split site, with its native N- and C-termini used as the fusion linker. These termini regions are usually flexible, containing at least 5 AA beyond the terminal TMs. Thus, termini restraining does not necessarily require additional linker sequences, or truncations or mutations of the target membrane protein. Fusion of the entire membrane protein maintains its function and core structure.

Highly stable and crystallizable proteins can be used as the scaffold protein for termini restraining. A demonstrated example is restraining with sfGFP, which dramatically improved the expression level, yield and stability of membrane proteins from distinct functional families. Using sfGFP has an additional advantage that the entire process from protein expression, construct screening, to protein purification and crystallization could be monitored by the fluorescence signal. Furthermore, many surface regions of the sfGFP molecule can participate in crystal packing interactions. This method has resulted in the crystallization of six small membrane proteins, VKOR, VKOR-like, JAGN1, CD53, SPCS1 and DsbB [29]. In all these crystals, the sfGFP and membrane protein molecules form two layers of packing interactions. The sfGFP molecules interact with each other in one layer to afford the crystallization scaffold, and the membrane protein molecules form the second layer. Probably owing to the LCP crystallization condition, the TMs of membrane proteins often form additional interactions, which make the crystals have both type I (interactions between membrane regions) and type II (interactions between soluble regions) properties [43].

Changing the linker length between two termini of the membrane protein and the fused sfGFP is an effective way to improve crystal diffraction. As an example, crystals of the sfGFP-restrained full-length VKOR-like (Figure 2F) diffracted to 4.6 Å. Truncation of 5 AA at the N terminus of VKOR-like protein changed the crystal packing interactions between sfGFP and VKORL molecules and changed the crystal form from P<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> to C2. The crystal diffraction was improved to 2.4 Å after the truncation. With the sfGFP restraining and the minor optimization of termini length, 11 crystal structures [31] of VKOR and VKOR-like were determined, which captured almost their entire catalytic cycle and revealed their inhibition mechanism by oral anticoagulants targeting VKOR.

Structure determination of DsbB, an *E. coli* oxidoreductase, further demonstrates the advantages of the termini restraining strategy, the use of which readily generated a

2.9 Å structure. In comparison, many years of attempts with traditional crystallization approaches only resulted in 3.6 Å resolution at the best. Large efforts were taken to obtain the low-resolution structure by co-crystallization, such as identifying a suitable Fab [44] for DsbB or forming disulfide linked complex with its partner protein, DsbA [45]. Termini restraining bypassed these demanding processes and generated the high-resolution structure that showed this thiol oxidoreductase uses a catalytic triad similar to that observed in serine or cysteine proteases, thereby affording new insights into the catalytic mechanism of DsbB [46,47].

#### 4. Selection of Scaffold Proteins

There are several criteria to consider for the selection of the fusion scaffold. First, the proteins should be stable and highly crystallizable to serve as a scaffold. Second, high-resolution structures of the scaffold should be available in the structure database, which indicates that this protein alone can form well diffracted crystals. After the fusion constructs are generated, cautions should be taken to confirm that the function or activity of the target membrane protein is unaffected.

For N or C terminal fusion, the optimal scaffold proteins can be those allowing a helical linker with the target membrane protein. For instance, the use of MBP, a largely helical protein, allows its connection with a membrane protein via a rigid and continuous helix [10,33]. The length of this fused helix can be varied to promote contacting interactions between the scaffold and the membrane protein. The size of the scaffold protein or linker length is generally not restricted for N or C terminal fusion.

For insertion fusion, the size of the scaffold protein can be restricted. The scaffold proteins used for insertion are relatively small: both T4 lysozyme and BRIL are around 10 KDa, and the largest scaffold reported to date is the 22 KDa PGS [42]. Because the scaffold protein inserts between TMs of the target membrane protein, the large scaffold may interfere with folding interactions between the TMs. Moreover, the distance between rigid structure features (e.g.,  $\alpha$ -helix or  $\beta$ -strand) at the termini of the scaffold protein should roughly match the size of the insertion site. For instance, the distance between the ends of terminal helices in T4 lysozyme is about 10 Å [17]. Consistently, the distance between TM5 and TM6 of GPCR is approximately 6–14 Å. Conversely, a mismatch in the distances may increase flexibility of the connecting regions and/or interfere with the membrane protein folding.

Similarly, termini restraining may require a rough distance match between the two terminal TMs of the membrane protein and the structural gap at the split site of the scaffold. A large mismatch may increase the flexibility between the fused scaffold and membrane protein molecules, which may in turn hinder crystallization. Despite this concern, the sfGFP scaffold can at least tolerate the distances of 8–21 Å between terminal TMs for the crystallization of small membrane proteins and for obtaining their high-resolution structures [29].

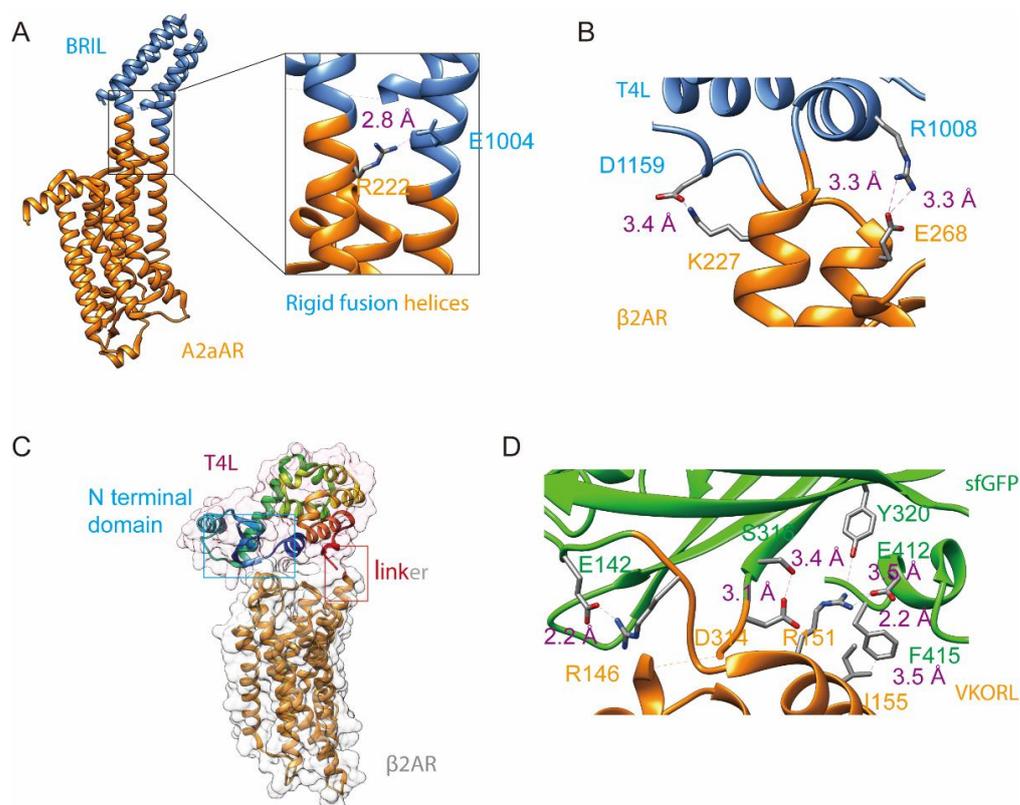
#### 5. Design of Fusion Linkers

For the fusion protein crystallization, restricting the relative movement between the target protein and the scaffold protein may increase the chance of obtaining well diffracted crystals. With relative short linkers, additional contacting interactions may be formed between the two fused proteins.

One way to create a relatively rigid linker between the target protein and the scaffold protein is by joining their helical regions to form an uninterrupted helix [48]. This method has been used successfully for soluble proteins with MBP fusion. For example, a human hPS1-card domain was fused to the C terminus of MBP via a connection helix [49]. MBP contains a C-terminal helix that ends at residue T366. If the target protein has an N-terminal helix, a helical connection can be generated by fusing the two helices and, if necessary, adding helix-forming linkers such as AAAAF [50]. The helical linker can be further optimized by extension or truncation, with each residue change brings a rotation of

approximately 100 degrees between the two fused protein molecules. A sufficiently long helical linker may help maintain the native conformation and function of the target protein.

Uninterrupted helical connections have been observed for insertion fusion in structures of GPCR [40,51–54] and Ctr1 [41] (a copper transporter) that are fused to T4L or BRIL by insertion. In the structure of human A (2A) adenosine receptor (PDB: 4EIY) [51] (Figure 3A), BRIL (1-106 AA) replaced ECL3 between residues L208 and E219. The N terminal helix of BRIL connects to TM5 of the receptor via a long helix that is slightly bent. Similarly, the C terminal helix of BRIL connects to TM6 via a bent helix. These helical connections lower the flexible motions between the fusion proteins, resulting in a 1.8 Å structure. The structure is further stabilized by interactions formed between residues R222 from the receptor and E1004 from BRIL.



**Figure 3.** Apart from the linker connection, various interactions can form to stabilize the conformation of the fusion protein. (A) Structure of insertion fused BRIL with human A (2A) adenosine receptor (A2aAR-BRIL, PDB 4EIY). Helical connections are formed at the two linker regions between BRIL and A2aAR. Inset, salt bridge is formed between the E1004 from BRIL and R222 from the A2aAR. (B) Structure of insertion fused  $\beta$ 2AR-T4L (PDB 2RH1). Designed helical connection is not formed. Salt bridges are formed between T4L R1008 and  $\beta$ 2AR E268, and between T4L D1159 and  $\beta$ 2AR K227. (C) Structure of T4L fused to the N-terminus of  $\beta$ 2AR (PDB 4GBR). The surface of T4L is shown in light pink and that of  $\beta$ 2AR in grey. (D) Structure of termini restrained sfGFP-VKOR-like (PDB 6WVI). Many interactions are formed between the fused sfGFP and VKORL (see description in text).

For insertion fusion of  $\beta$ 2AR with T4L ( $\beta$ 2AR-T4L, PDB:2RH1) [17] (Figure 3B), however, helical connection was not observed. On the other hand, T4L interacts with  $\beta$ 2AR via salt bridges, T4L D1159 with  $\beta$ 2AR K227 and T4L-R1008 with  $\beta$ 2AR E268. Approximately 400 Å<sup>2</sup> surface is buried between T4L and  $\beta$ 2AR that stabilizes their interaction and facilitates crystal formation. For terminal fusion in the T4L- $\beta$ 2AR structure [34], the C terminal helix of T4L was fused to the N-terminal helix of  $\beta$ 2AR with a short Ala-Ala linker that is prone to form  $\alpha$ -helix. The designed helical connection, however, was not formed. Only a 4 Å structure was determined, in which  $\beta$ 2AR interacts with the N terminal domain of T4L

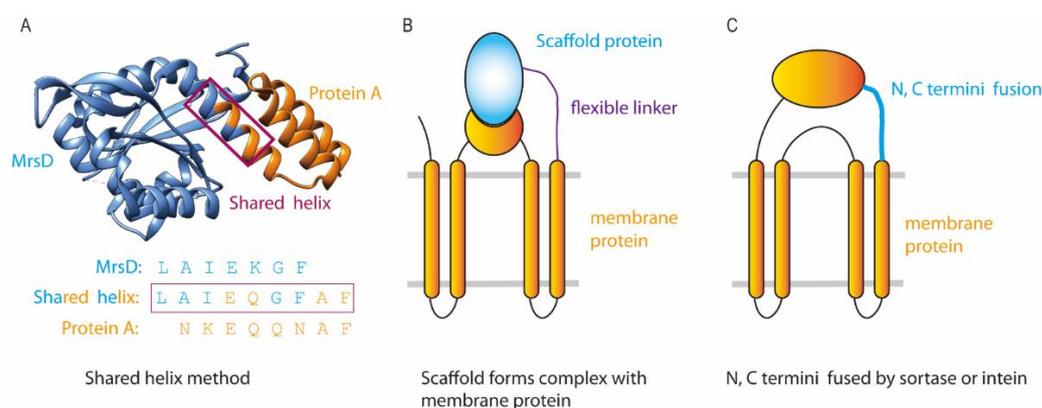
to stabilize their relative conformation (Figure 3C). Taken together, the helix connection can be tricky to achieve by protein design.

Even if a helix connection cannot be created, the scaffold protein may form stabilization interactions with the membrane protein. The interactions can be salt bridges (most common), hydrogen bonds and hydrophobic interactions. Such interactions have been observed for the terminal-fused T4L- $\beta$ 2AR, insertion fused  $\beta$ 2AR-T4L, and termini restrained VKOR and VKORL. A short linker is expected to favor these interactions. The two linking points used in insertion fusion and termini restraining have the additional advantage of restricting the flexible movement between the fusion proteins, and shortening of the two linkers may further favor stabilization interactions.

As an example, the linkers between the split site of sfGFP (a  $\beta$ -strand region) and the termini-restrained human VKOR or VKOR-like are flexible loops [31]. Shortening one of the linkers by 5 AA improved the structural resolution of VKOR-like from 4.6 Å to 2.4 Å [29] (Figure 3D). The short linker promoted salt bridge formation between VKORL R7 and sfGFP E142, and between VKORL R12 and sfGFP E412 (introduced from the cloning). In addition, VKORL D175 forms a hydrogen bond with sfGFP S147, and VKORL I16 interacts hydrophobically with F415. All these interactions, together with the backbone connection, lower the flexibility of the fused membrane proteins and facilitate its crystallization.

Taken together, new interactions may form between the fused proteins and changing the linkers may favor such interactions. The membrane protein may even bind tightly to the scaffold protein, such as the fusion of T4L–rhodopsin mutant to arrestin. In order to not interfere with such binding interactions, a flexible fusion linker [50], such as GGGGS, can be used.

To ensure helical connection, a shared helix method [55] (Figure 4A) has been recently developed. The principle of this method is that helix formation in proteins does not usually occur in isolation but is stabilized by interactions with adjacent structural regions. Thus, these interactions should be maintained when designing a helical connection. Specifically, the two helices from the two fusion proteins are first superimposed on a model helix to visualize the structural arrangement between the fused proteins and to avoid their structural clashes. Next, residues from each helix were selected to maintain the native interactions in each protein. Additional connections use residues prone to form helices, such as A and F. The shared helix method has succeeded in the crystallization of soluble fusion proteins [55]. In addition, Jeong et al. [56] used a chemical crosslinker to stabilize the connecting helix. AlphaFold [57] prediction may also be helpful for designing the helical linkers, saving benchwork efforts.



**Figure 4.** Protein fusion with shared helix, interaction partner and circular connection. Scaffold proteins are colored in blue and the target proteins in orange. (A) The shared helix method generates rigid linker connection. In the two connecting helices, residues forming native interactions in each fused protein are kept in the sequence of the shared helix (blue from NrsD and orange from protein A). (B) Interaction partners of the target membrane protein are used as the scaffold protein. (C) Membrane protein with a soluble domain is termini connected to generate a circular protein.

## 6. Crystallization and Phase Determination

Membrane proteins with the fusions are often crystallized in meso, probably because the LCP method usually gives better structural resolution owing to type I crystal formation. However, the fusion method is also compatible with vapor diffusion methods, such as crystallization of the copper transporter Ctr1 with BRIL fusion.

Presence of the scaffold protein overcomes the difficulties of solving the phase problem in crystallography. The scaffold affords a model for molecular replacement, allowing initial phases to be obtained. Furthermore, the structure and diffraction data from scaffold-alone crystals can be used for cross-crystal averaging with the crystals of the fusion protein, a density modification method that dramatically improves the phases and generates high-quality electron density maps [58].

## 7. Discussion

Membrane proteins have been difficult targets for crystallization. As a prominent example, the journey for the structure determination of GPCRs took several decades. Owing to the development of the insertion fusion methods, the first GPCR structure was determined by the T4L fusion [17], resulting in a Nobel prize award in 2012. Since then, many other GPCR structures were determined to a high resolution. There are, however, limited reports of applying insertion fusion to other families of membrane proteins, because the selection of the proper insertion site can be difficult for proteins with various structural folds and may require large effort of trial-and-error. In particular, insertion of a large scaffold into small membrane proteins may interfere with their native conformation.

Fusion by termini restraining has largely overcome the difficulties with crystallizing small membrane proteins. The workloads are limited because generally the entire membrane protein is inserted into the scaffold protein, and only the fusion linkers may need to be adjusted to reach high structural resolution. Even without any linker optimization, the structure of human CD53 [59] and JAGN1 [29] was determined to a high resolution. Termini restraining is highly promising for the crystallization of many small membrane proteins with unknown structures. This method, however, is so far limited to membrane proteins with an even number of TMs. Additionally, caution should be taken for membrane proteins containing termini regions essential for their function. For difficult membrane proteins, all three categories of the fusion methods are worth trying.

Overall, the fusion strategies have several advantages for the crystallographic studies of membrane proteins. The primary purpose of fusing a scaffold is to introduce a large hydrophilic surface that enables crystal packing. Presence of the fused scaffold in the crystals also facilitates phase determination. The fusion often improves the stability, expression level and effective yield of membrane proteins. Designing the linker region(s) between the fused proteins may reduce the linker flexibility and promote interactions between fused proteins, and the shared helix method or AlphaFold prediction can assist the fusion design.

Fusion technologies can be further developed to fulfill the need for stabilizing and crystallizing membrane proteins. For instance, interaction partners of membrane proteins can be used as the scaffold protein, using N or C terminal fusion (Figure 4B). The two ends of membrane proteins (especially when an N or C-terminal soluble domain is present) can be connected by sortase [60] or intein [61] to generate a circular protein (Figure 4C) for increased thermostability [62].

**Author Contributions:** All authors have contributed to the writing of this manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** W.L. is supported by National Heart, Lung, and Blood Institute (R01 HL121718), W. M. Keck Foundation (Forefront of Science Award), American Heart Association (Collaborative Sciences Award), Children's Discovery Institute (MCII 2020-854), National Eye Institute (R21 EY028705), National Institute of General Medical Sciences (R01 GM131008), and National Institute of Allergy and Infectious Diseases (R01 AI158500).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are openly available in Pubmed Central.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yin, H.; Flynn, A.D. Drugging Membrane Protein Interactions. *Annu. Rev. Biomed. Eng.* **2016**, *18*, 51–76. [[CrossRef](#)] [[PubMed](#)]
2. Gong, J.; Chen, Y.; Pu, F.; Sun, P.; He, F.; Zhang, L.; Li, Y.; Ma, Z.; Wang, H. Understanding Membrane Protein Drug Targets in Computational Perspective. *Curr. Drug Targets* **2019**, *20*, 551–564. [[CrossRef](#)] [[PubMed](#)]
3. Lee, Y.; Basith, S.; Choi, S. Recent Advances in Structure-Based Drug Design Targeting Class A G Protein-Coupled Receptors Utilizing Crystal Structures and Computational Simulations. *J. Med. Chem.* **2018**, *61*, 1–46. [[CrossRef](#)] [[PubMed](#)]
4. Sowlati-Hashjin, S.; Gandhi, A.; Garton, M. Dawn of a New Era for Membrane Protein Design. *BioDes. Res.* **2022**, *2022*, 9791435. [[CrossRef](#)]
5. Ayub, M.; Bayley, H. Engineered transmembrane pores. *Curr. Opin. Chem. Biol.* **2016**, *34*, 117–126. [[CrossRef](#)] [[PubMed](#)]
6. Kang, H.J.; Lee, C.; Drew, D. Breaking the barriers in membrane protein crystallography. *Int. J. Biochem. Cell Biol.* **2013**, *45*, 636–644. [[CrossRef](#)] [[PubMed](#)]
7. Kawate, T.; Gouaux, E. Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **2006**, *14*, 673–681. [[CrossRef](#)] [[PubMed](#)]
8. Hattori, M.; Hibbs, R.E.; Gouaux, E. A Fluorescence-Detection Size-Exclusion Chromatography-Based Thermostability Assay for Membrane Protein Precrystallization Screening. *Structure* **2012**, *20*, 1293–1299. [[CrossRef](#)] [[PubMed](#)]
9. Magnani, F.; Serrano-Vega, M.J.; Shibata, Y.; Abdul-Hussein, S.; Lebon, G.; Miller-Gallacher, J.; Singhal, A.; Strege, A.; Thomas, J.A.; Tate, C.G. A mutagenesis and screening strategy to generate optimally thermostabilized membrane proteins for structural studies. *Nat. Protoc.* **2016**, *11*, 1554–1571. [[CrossRef](#)]
10. Waugh, D.S. Crystal structures of MBP fusion proteins. *Protein Sci.* **2016**, *25*, 559–571. [[CrossRef](#)]
11. Hu, J.; Qin, H.; Gao, F.P.; Cross, T.A. A systematic assessment of mature MBP in membrane protein production: Overexpression, membrane targeting and purification. *Protein Expr. Purif.* **2011**, *80*, 34–40. [[CrossRef](#)] [[PubMed](#)]
12. Yeliseev, A.A.; Wong, K.K.; Soubias, O.; Gawrisch, K. Expression of human peripheral cannabinoid receptor for structural studies. *Protein Sci.* **2005**, *14*, 2638–2653. [[CrossRef](#)] [[PubMed](#)]
13. Shibata, Y.; White, J.F.; Serrano-Vega, M.J.; Magnani, F.; Aloia, A.L.; Grishammer, R.; Tate, C.G. Thermostabilization of the Neurotensin Receptor NTS1. *J. Mol. Biol.* **2009**, *390*, 262–277. [[CrossRef](#)]
14. Schalch, T.; Job, G.; Shanker, S.; Partridge, J.F.; Joshua-Tor, L. The Chp1-Tas3 core is a multifunctional platform critical for gene silencing by RITS. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1351–1357. [[CrossRef](#)] [[PubMed](#)]
15. Panavas, T.; Sanders, C.; Butt, T.R. SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. *Methods Mol. Biol.* **2009**, *497*, 303–317. [[PubMed](#)]
16. Zuo, X.; Li, S.; Hall, J.; Mattern, M.R.; Tran, H.; Shoo, J.; Tan, R.; Weiss, S.R.; Butt, T.R. Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. *J. Struct. Funct. Genom.* **2005**, *6*, 103–111. [[CrossRef](#)] [[PubMed](#)]
17. Cherezov, V.; Rosenbaum, D.M.; Hanson, M.A.; Rasmussen, S.G.; Thian, F.S.; Kobilka, T.S.; Choi, H.J.; Kuhn, P.; Weis, W.I.; Kobilka, B.K.; et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258–1265. [[CrossRef](#)] [[PubMed](#)]
18. Chu, R.; Takei, J.; Knowlton, J.R.; Andrykovitch, M.; Pei, W.; Kajava, A.V.; Steinbach, P.J.; Ji, X.; Bai, Y. Redesign of a Four-helix Bundle Protein by Phage Display Coupled with Proteolysis and Structural Characterization by NMR and X-ray Crystallography. *J. Mol. Biol.* **2002**, *323*, 253–262. [[CrossRef](#)]
19. Chun, E.; Thompson, A.A.; Liu, W.; Roth, C.B.; Griffith, M.T.; Katritch, V.; Kunken, J.; Xu, F.; Cherezov, V.; Hanson, M.A.; et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **2012**, *20*, 967–976. [[CrossRef](#)]
20. Xu, Y.; Kong, J.; Kong, W.T. Improved membrane protein expression in *Lactococcus lactis* by fusion to Mistic. *Microbiol.-Sgm.* **2013**, *159*, 1002–1009. [[CrossRef](#)] [[PubMed](#)]
21. Deniaud, A.; Bernaudat, F.; Frelet-Barrand, A.; Juillan-Binard, C.; Vernet, T.; Rolland, N.; Pebay-Peyroula, E. Expression of a chloroplast ATP/ADP transporter in *E. coli* membranes: Behind the Mistic strategy. *Bba-Biomembranes* **2011**, *1808*, 2059–2066. [[CrossRef](#)] [[PubMed](#)]
22. Chowdhury, A.; Feng, R.T.; Tong, Q.; Zhang, Y.X.; Xie, X.Q. Mistic and TarCF as fusion protein partners for functional expression of the cannabinoid receptor 2 in *Escherichia coli*. *Protein Expr. Purif.* **2012**, *83*, 128–134. [[CrossRef](#)]
23. Pao, P.J.; Hsu, M.F.; Chiang, M.H.; Chen, C.T.; Lee, C.C.; Wang, A.H. Structural basis of an epitope tagging system derived from *Haloarcula marismortui* bacteriorhodopsin I D94N and its monoclonal antibody GD-26. *FEBS J.* **2022**, *289*, 730–747. [[CrossRef](#)] [[PubMed](#)]
24. Hsu, M.F.; Yu, T.F.; Chou, C.C.; Fu, H.Y.; Yang, C.S.; Wang, A.H. Using *Haloarcula marismortui* bacteriorhodopsin as a fusion tag for enhancing and visible expression of integral membrane proteins in *Escherichia coli*. *PLoS ONE* **2013**, *8*, e56363. [[CrossRef](#)] [[PubMed](#)]

25. Pedelacq, J.D.; Cabantous, S.; Tran, T.; Terwilliger, T.C.; Waldo, G.S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **2006**, *24*, 79–88. [[CrossRef](#)] [[PubMed](#)]
26. Baird, G.S.; Zacharias, D.A.; Tsien, R.Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11241–11246. [[CrossRef](#)] [[PubMed](#)]
27. Fan, J.Y.; Cui, Z.Q.; Wei, H.P.; Zhang, Z.P.; Zhou, Y.F.; Wang, Y.P.; Zhang, X.E. Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. *Biochem. Biophys. Res. Commun.* **2008**, *367*, 47–53. [[CrossRef](#)] [[PubMed](#)]
28. Varshavsky, A. The N-end rule pathway and regulation by proteolysis. *Protein Sci.* **2011**, *20*, 1298–1345. [[CrossRef](#)]
29. Liu, S.; Li, S.; Yang, Y.; Li, W. Termini restraining of small membrane proteins enables structure determination at near-atomic resolution. *Sci. Adv.* **2020**, *6*, eabc3717. [[CrossRef](#)] [[PubMed](#)]
30. Liu, S.; Li, S.; Krezel, A.M.; Li, W. Stabilization and structure determination of integral membrane proteins by termini restraining. *Nat. Protoc.* **2022**, *17*, 540–565. [[CrossRef](#)]
31. Liu, S.; Li, S.; Shen, G.; Sukumar, N.; Krezel, A.M.; Li, W. Structural basis of antagonizing the vitamin K catalytic cycle for anticoagulation. *Science* **2021**, *371*, eabc5667. [[CrossRef](#)] [[PubMed](#)]
32. Ye, Q.; Singh, V.K.; Blonde, J.D.; Jia, Z. Crystallization and preliminary X-ray analysis of the GST-fused human Bri3 N-terminal domain. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2005**, *61*, 62–64. [[CrossRef](#)]
33. Smyth, D.R.; Mrozkiewicz, M.K.; McGrath, W.J.; Listwan, P.; Kobe, B. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **2003**, *12*, 1313–1322. [[CrossRef](#)] [[PubMed](#)]
34. Zou, Y.; Weis, W.I.; Kobilka, B.K. N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. *PLoS ONE* **2012**, *7*, e46039. [[CrossRef](#)] [[PubMed](#)]
35. Rasmussen, S.G.; DeVree, B.T.; Zou, Y.; Kruse, A.C.; Chung, K.Y.; Kobilka, T.S.; Thian, F.S.; Chae, P.S.; Pardon, E.; Calinski, D.; et al. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **2011**, *477*, 549–555. [[CrossRef](#)] [[PubMed](#)]
36. Kang, Y.; Zhou, X.E.; Gao, X.; He, Y.; Liu, W.; Ishchenko, A.; Barty, A.; White, T.A.; Yefanov, O.; Han, G.W.; et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* **2015**, *523*, 561–567. [[CrossRef](#)]
37. Lu, M.; Wu, B. Structural studies of G protein-coupled receptors. *IUBMB Life* **2016**, *68*, 894–903. [[CrossRef](#)] [[PubMed](#)]
38. Caffrey, M. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallogr. F Struct. Biol. Commun.* **2015**, *71*, 3–18. [[CrossRef](#)]
39. Thorsen, T.S.; Matt, R.; Weis, W.I.; Kobilka, B.K. Modified T4 Lysozyme Fusion Proteins Facilitate G Protein-Coupled Receptor Crystallization. *Structure* **2014**, *22*, 1657–1664. [[CrossRef](#)] [[PubMed](#)]
40. Miyagi, H.; Asada, H.; Suzuki, M.; Takahashi, Y.; Yasunaga, M.; Suno, C.; Iwata, S.; Saito, J.I. The discovery of a new antibody for BRIL-fused GPCR structure determination. *Sci. Rep.* **2020**, *10*, 11669. [[CrossRef](#)] [[PubMed](#)]
41. Ren, F.; Logeman, B.L.; Zhang, X.; Liu, Y.; Thiele, D.J.; Yuan, P. X-ray structures of the high-affinity copper transporter Ctr1. *Nat. Commun.* **2019**, *10*, 1386. [[CrossRef](#)]
42. Yin, J.; Mobarec, J.C.; Kolb, P.; Rosenbaum, D.M. Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. *Nature* **2015**, *519*, 247–250. [[CrossRef](#)]
43. Birch, J.; Axford, D.; Foadi, J.; Meyer, A.; Eckhardt, A.; Thielmann, Y.; Moraes, I. The fine art of integral membrane protein crystallisation. *Methods* **2018**, *147*, 150–162. [[CrossRef](#)]
44. Inaba, K.; Murakami, S.; Nakagawa, A.; Iida, H.; Kinjo, M.; Ito, K.; Suzuki, M. Dynamic nature of disulphide bond formation catalysts revealed by crystal structures of DsbB. *EMBO J.* **2009**, *28*, 779–791. [[CrossRef](#)]
45. Inaba, K.; Ito, K. Structure and mechanisms of the DsbB-DsbA disulfide bond generation machine. *Biochim. Biophys. Acta* **2008**, *1783*, 520–529. [[CrossRef](#)]
46. Phan, J.; Zdanov, A.; Evdokimov, A.G.; Tropea, J.E.; Peters, H.K., 3rd; Kapust, R.B.; Li, M.; Wlodawer, A.; Waugh, D.S. Structural basis for the substrate specificity of tobacco etch virus protease. *J. Biol. Chem.* **2002**, *277*, 50564–50572. [[CrossRef](#)]
47. Matthews, D.A.; Dragovich, P.S.; Webber, S.E.; Fuhrman, S.A.; Patick, A.K.; Zalman, L.S.; Hendrickson, T.F.; Love, R.A.; Prins, T.J.; Marakovits, J.T.; et al. Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11000–11007. [[CrossRef](#)]
48. Kwon, N.Y.; Kim, Y.; Lee, J.O. The application of helix fusion methods in structural biology. *Curr. Opin. Struct. Biol.* **2020**, *60*, 110–116. [[CrossRef](#)]
49. Jin, T.; Chuenchor, W.; Jiang, J.; Cheng, J.; Li, Y.; Fang, K.; Huang, M.; Smith, P.; Xiao, T.S. Design of an expression system to enhance MBP-mediated crystallization. *Sci. Rep.* **2017**, *7*, 40991. [[CrossRef](#)]
50. Chen, X.; Zaro, J.L.; Shen, W.C. Fusion protein linkers: Property, design and functionality. *Adv. Drug Deliv. Rev.* **2013**, *65*, 1357–1369. [[CrossRef](#)]
51. Liu, W.; Chun, E.; Thompson, A.A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G.W.; Roth, C.B.; Heitman, L.H.; AP, I.J.; et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **2012**, *337*, 232–236. [[CrossRef](#)]
52. McCorvy, J.D.; Wacker, D.; Wang, S.; Agegnehu, B.; Liu, J.; Lansu, K.; Tribo, A.R.; Olsen, R.H.J.; Che, T.; Jin, J.; et al. Structural determinants of 5-HT2B receptor activation and biased agonism. *Nat. Struct. Mol. Biol.* **2018**, *25*, 787–796. [[CrossRef](#)]
53. White, K.L.; Eddy, M.T.; Gao, Z.G.; Han, G.W.; Lian, T.; Deary, A.; Patel, N.; Jacobson, K.A.; Katritch, V.; Stevens, R.C. Structural Connection between Activation Microswitch and Allosteric Sodium Site in GPCR Signaling. *Structure* **2018**, *26*, 259–269.e5. [[CrossRef](#)]

54. Vuckovic, Z.; Gentry, P.R.; Berizzi, A.E.; Hirata, K.; Varghese, S.; Thompson, G.; van der Westhuizen, E.T.; Burger, W.A.C.; Rahmani, R.; Valant, C.; et al. Crystal structure of the M5 muscarinic acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 26001–26007. [[CrossRef](#)]
55. Youn, S.J.; Kwon, N.Y.; Lee, J.H.; Kim, J.H.; Choi, J.; Lee, H.; Lee, J.O. Construction of novel repeat proteins with rigid and predictable structures using a shared helix method. *Sci. Rep.* **2017**, *7*, 2595. [[CrossRef](#)]
56. Jeong, W.H.; Lee, H.; Song, D.H.; Eom, J.H.; Kim, S.C.; Lee, H.S.; Lee, H.; Lee, J.O. Connecting two proteins using a fusion alpha helix stabilized by a chemical cross linker. *Nat. Commun.* **2016**, *7*, 11031. [[CrossRef](#)]
57. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
58. Li, W.; Li, F. Cross-crystal averaging with search models to improve molecular replacement phases. *Structure* **2011**, *19*, 155–161. [[CrossRef](#)]
59. Yang, Y.; Liu, X.R.; Greenberg, Z.J.; Zhou, F.; He, P.; Fan, L.; Liu, S.; Shen, G.; Egawa, T.; Gross, M.L.; et al. Open conformation of tetraspanins shapes interaction partner networks on cell membranes. *EMBO J.* **2020**, *39*, e105246. [[CrossRef](#)]
60. Chen, I.; Dorr, B.M.; Liu, D.R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11399–11404. [[CrossRef](#)]
61. Pinto, F.; Thornton, E.L.; Wang, B. An expanded library of orthogonal split inteins enables modular multi-peptide assemblies. *Nat. Commun.* **2020**, *11*, 1529. [[CrossRef](#)]
62. Conlan, B.F.; Gillon, A.D.; Craik, D.J.; Anderson, M.A. Circular Proteins and Mechanisms of Cyclization. *Biopolymers* **2010**, *94*, 573–583. [[CrossRef](#)]