Published in final edited form as:

Curr Opin Struct Biol. 2015 June ; 32: 131–138. doi:10.1016/j.sbi.2015.04.005.

Post-expression strategies for structural investigations of membrane proteins

Linda Columbus

Department of Chemistry, University of Virginia, Charlottesville, Virginia, 22902

Abstract

Currently, membrane proteins only comprise 1.5% of the protein data bank and, thus, still remain a challenge for structural biologists. Expression, stabilization in membrane mimics (e.g. detergent), heterogeneity (conformational and chemical), and crystallization in the presence of a membrane mimic are four major bottlenecks encountered. In response, several post-expression protein modifications have been utilized to facilitate structure determination of membrane proteins. This review highlights four approaches: limited proteolysis, deglycosylation, cysteine alkylation, and lysine methylation. Combined these approaches have facilitated the structure determination of more than 40 membrane proteins and, therefore, are a useful addition to the membrane protein structural biologist's toolkit.

Introduction

Membrane proteins are the gatekeepers of cells regulating the flow of information and small molecules across the cell membranes. To better understand membrane proteins, structural biologists determine their high-resolution structure using X-ray crystallography or NMR spectroscopy. The structure provides a starting-point to base hypotheses about how proteins function. Most membrane proteins exist in at least two states – “open or closed” or “on and off” – and move between these states in order to mediate the movement of a signal or molecule across the membrane. A structure potentially provides a snapshot of the protein in one of these states. Obtaining these high-impact structures remains a challenge as indicated by their underrepresentation in the protein data bank.

Membrane protein structural biology has required a multi-tier, highly empirical approach to obtain samples that are amenable to the structural techniques commonly used for soluble proteins. Much attention is paid to the construct (e.g. organism, chimeras, and mutations), expression, and membrane mimic selection [1,2]. However, several structures have required post-expression protein modifications. The post-expression modifications improve structure

columbus@virginia.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest statement

Nothing declared.

determination by a variety of mechanisms such as creating homogenous crystal contacts, trapping a single conformational state, and removing flexible regions. This review highlights the use of limited proteolysis, deglycosylation, reductive methylation, and cysteine alkylation (Figure 1) as post-expression modification approaches to structural investigations of membrane proteins.

Limited proteolysis

Limited proteolysis is the treatment of a protein with a protease such that only exposed and dynamic regions (not folded domains) are cleaved according to the protease selectivity. In addition, for membrane proteins, the transmembrane regions are protected from protease cleavage by the membrane or membrane mimic. The proteolytic product is typically evaluated using sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis and mass spectrometry. The proteolytic cleavage is typically performed during the purification steps (which are mostly done in detergents) and the small cleaved segments are removed from the sample before structure determination.

In some cases, limited proteolysis is used to map out the domains of a protein. The results are then used to guide the molecular cloning of specific domains of interest. For example, the flexible periplasmic domains of the 809-amino acid P pilus usher PapC prevented crystallization and limited proteolysis was used to identify a 55 kD fragment corresponding to the outer membrane translocation domain [3]. The 55 kD fragment was then expressed, purified, and crystallized and the structure was determined with 3.2 Å resolution [3]. In some cases, the new genetic constructs may have lower expression levels, interfere with cellular targeting, or result in instability during purification. Thus, limited proteolysis during the purification may be preferred. Limited proteolysis can be used to remove affinity tags that do not have specific cleavage sites especially in the case of C-terminal tags for which it is desirable to not have additional residues (e.g. using carboxypeptidase for the μ - and δ -opioid receptors [4,5]). Beyond tag removal and domain mapping, limited proteolysis has facilitated nuclear magnetic resonance (NMR) and X-ray crystallography structure determination.

Limited proteolysis: NMR

In order to determine an NMR structure, an assignment of observed resonances needs to be completed. For large proteins or complexes, the assignment process can be challenging. Various assignment strategies exist such as amino acid specific labeling [6,7] and segmental labeling [8]. Another approach is limited proteolysis where loop regions are removed with a protease yet the fragmented protein remains folded. The removal of these flexible regions simplifies the NMR spectrum by removing intense resonances that have spectral overlap with weaker resonances from the folded membrane region. Limited proteolysis has proven to be an effective method for assigning β -barrel membrane protein resonances [9]. The β -barrel membrane proteins are highly stable once folded in detergent as indicated by their resistance to unfolding in SDS (even boiled) and, thus, are likely more amenable to this approach than α -helical membrane proteins.

To date, less than ten β -barrel membrane protein structures have been determined with solution NMR. Most of the proteins investigated do not have significant soluble portions in the periplasmic or extracellular loops. However, in the case of Opa₆₀, the four extracellular loops comprised approximately half of the protein. These loop resonances are more intense than those of the β -barrel and, in many cases, had significant line broadening (Figure 2B). The loop resonance occluded many of the resonances from amino acids within the barrel. β -barrel membrane proteins are typically very stable and remain folded in SDS; thus, the folded and unfolded protein forms migrate differently on an SDS-PAGE gel (Figure 2A). As a result, limited proteolysis of β -barrel membrane proteins in a membrane mimic with a protease may not result in denaturation. For Opa₆₀ [10], Opa₅₀, and OprH [11], the resulting five fragments after trypsin proteolysis maintained the β -barrel fold and the corresponding resonances were nearly superimposable with that of the untreated proteins (Figure 2B) [9]. The ability to conduct the assignments on the proteolysed sample and map the assignment on to the full length was essential to determining the structure of the Opa₆₀ β -barrel [9].

Limited proteolysis: Crystallization

Limited proteolysis has also been a useful tool for obtaining x-ray crystal structures. The strategy is different than that used for NMR structure determination. Limited proteolysis is used to (i) remove dynamic soluble regions and/or (ii) to remove soluble domains that may interfere with crystallization. The first KcsA potassium channel crystal structure was the first demonstration of the use of limited proteolysis for structure determination [12]. In this case, chymotrypsin was used to cleave 45 residues from the C-terminus and the cleaved purified product yielded well-diffracting crystals. An interesting note to this protocol was the finding that certain commercial chymotrypsin batches had a trypsin contaminant that yielded a second cleavage product that interfered with crystallization [13]. After demonstrated success of limited proteolysis in the structure determination of KcsA, the method was applied to thirteen other membrane proteins from 1998 to 2014 (Table 1). In all structures determined, either trypsin or chymotrypsin was used as the protease. The cleaved regions tend to be significant portions of the termini. For instance, the chymotrypsin treated LptD:LptE complex resulted in N-terminal residues 22-224 removed from LptD and residues 170-196 from LptE [14]. The method does not seem limited in the membrane protein fold or function investigated as limited proteolysis has been effective for the structure determination of α -helical and β -barrels and a variety of functions (e.g. channels, proteases, and transporters).

Deglycosylation

Membrane proteins isolated from natural sources or expressed in eukaryotic systems can be glycosylated. In some cases, glycosylation may not interfere with crystallization (ref [15,16] for example); however, for other proteins, glycosylation can inhibit crystallization often due to heterogeneous glycosylation. One approach to avoid glycosylation in recombinant systems is to mutate the glycosylated residue so the expressed mutated protein is not glycosylated. This is especially necessary if the modification is not accessible to glycosidases [17]. The mutation approach has been used for many membrane proteins (ref [18] for example); however, for ten membrane protein structures, glycosidases were used

(Table 1). Three different glycosidases, Peptide-N-glycosidase F (PNGase F), endoglycosidase F and endoglycosidase H, were exploited to remove glycosylation modifications from asparagine residues (Figure 1). Endoglycosidase F and H cleave after the first N-acetylglucosamine in high mannose or hybrid glycans. PNGase F cleaves before the first N-acetylglucosamine on the asparagine residue of high mannose, hybrid, and complex oligosaccharides.

The first demonstration of the use of glycosidases for membrane protein crystallography was with AQP1. AQP1 was isolated from bovine red blood cells and both glycosylated and non-glycosylated forms were observed [19]. PNGase F was incubated with AQP1 after solubilization from the membrane for two days and then further purified. Since 2001, nine proteins have been expressed with glycosylation and treated with glycosidases (Table 1). The most recent membrane protein structure obtained using PNGase F was the serotonin 5-HT₃ receptor, which also was subject to trypsin limited proteolysis [20]. In the case of PNGase, the entire modification is removed before crystallization. Therefore, the glycosylation is important for protein production, but hinders crystallization and is removed. In contrast, when EndoH is used an N-acetylglucosamine remains on the modified asparagine residue. The NAMDA (Fig 3) and AMPA receptors are the only demonstration of the use of EndoH and in both cases, N-acetylglucosamine is observed at crystal contacts.

Cysteine Alkylation

Treatment of proteins with iodoacetamide (after the addition of a reducing agent) alkylates any accessible cysteine thereby preventing the cysteine residue from forming disulfide bonds. The importance of this modification was recently illuminated for the β_2 adrenergic receptor [21]. A comparison of the size exclusion chromatogram of the iodoacetamide treated and untreated protein demonstrated that alkylation prevented dimer and higher order oligomers from forming [21]. Since the application of cysteine alkylation for the β_2 adrenergic receptor in 2008, the method has been applied to the structure determination of 21 GPCRs (Table 1). The GPCR structures that did not utilize cysteine alkylation are rhodopsin, NTS1 neurotensin receptor (selected using directed evolution approaches [22] or engineered for thermostability [23]), glucagon receptor [24], corticotropin-releasing factor receptor 1 [25], CXCR4 chemokine receptor [26], CCR5 chemokine receptor [27], and β_1 adrenergic receptor [28] (the four later proteins were engineered for thermostability). Although to date cysteine alkylation has not contributed to the structure determination of other types of membrane proteins, the approach should be considered if a protein has required cysteine residues in flexible regions of the protein that could facilitate disulfide mediated dimers.

Lysine Methylation

Accessible lysine residues and amino termini can be mono- and dimethylated using sodium borohydride and ^{13}C labeled formaldehyde as the methyl donor. The incorporated ^{13}C -labeled methyl can be detected by NMR and, coupled with a high resolution structure, ligand-binding and ligand-induced conformational changes can be investigated. The resulting ^{13}C -methylated receptor can be investigated with NMR, specifically using ^1H , ^{13}C -

HSQC (heteronuclear single-quantum coherence) or saturation transfer differencing-filtered HMQC (heteronuclear multiple-quantum coherence) spectra. Each resonance of the ^1H , ^{13}C correlation spectrum can be assigned using lysine to arginine mutations and monomethylated resonances are distinct from dimethylated.

This approach was used to investigate the conformational equilibrium of the extracellular loops of the β_2 adrenergic receptor upon ligand binding [29]. For the unliganded receptor, two methyl peaks are significantly shifted upfield from the other observed methyl resonances (indicating a different chemical environment) and were assigned to the dimethylated Lys305. In the crystal structure, Lys305 (at the transition from TM7 and extracellular loop 3, which connects transmembrane helices 6 and 7) forms a salt bridge with Asp192 (in extracellular loops 2, which connects transmembrane helices 4 and 5). Based on mutagenesis and several other experiments, the chemical shift perturbation of Lys305 was concluded to be due to the formation of the salt bridge. These peaks were monitored with different ligands to investigate the conformational equilibria of the salt bridge. When agonist carazolol is bound, the two methyl peaks shift downfield and the chemical shift difference between the two methyl peaks decreases. Thus, binding of carazolol alters the chemical environment surrounding the Lys305 methyls. Binding of neutral antagonists, such as alprenolol, do not cause chemical shift changes in the observed Lys305 methyl resonances with the peaks remaining unchanged compared to unliganded receptor. Agonists, such as formoterol, also modulate the Lys305 methyl resonances; however, rather than chemical shift changes, the peaks are no longer observed indicating a loss of the salt bridge. The combined results led to the hypothesis that there are three distinct conformations with respect to TM6 and TM7 that are modulated by the Lys305 – Asp192 salt bridge, which in turn is coupled to ligand binding. These results complement the static crystal structures by providing conformational dynamics and defining the energetic landscape modulated by diffusible ligands.

Beyond NMR investigations, it is interesting to note that reductive methylation has improved soluble protein crystallization [30-35] and the approach may be useful in membrane protein crystallization. Reductive methylation in soluble proteins improves resolution and decreases B factors compared to non-methylated native structures [35]. A more ordered crystal packing appears to be the result of an increase in the interaction radius of the lysine residue upon methylation [35]. Reductive methylation and alkylation is an inexpensive approach to modifying protein surface properties that could be used in conjunction or as an alternative to mutagenesis required for membrane protein crystallization [18,28,36].

Other post-expression considerations

There are a few other post-expression modifications and treatments that are worth mentioning in this review: engineered disulfides or crosslinking and ligand complexes. At least five membrane protein structures were determined with engineered cysteine residues for disulfide formation [37-40] or cross linking with mercury [41,42] (Table 1). The engineered cysteine residues are used to improve crystal diffraction quality and/or stabilize alternative protein conformations. In the case of β_2 adrenergic receptor, a disulfide crosslink

was formed between the protein and a low affinity agonist in order to trap the agonist-bound protein conformation [40]. The cysteine residues cross-linked with mercury stabilized a conformation of the glutamate transporter distinct from the outward facing conformation observed without the disulfide [41]. For the NMDA structure, the disulfide improved the diffraction quality of the crystals (compared to the non-crosslinked structure) by decreasing the conformational flexibility of the extracellular domains (Fig 3) [37].

Although not a covalent modification, protein-protein and protein-ligand complexes are often formed after protein expression. These complexes can be difficult to form and can require 24 hour incubation periods. For instance, to obtain the SecA-SecYEG complex with ADP:BeFx ligand bound, the complex was dialyzed overnight with buffer containing ADP and BeFx [43]. Simply adding the ligands directly and setting up crystallization trays did not yield complex formation. To form the β_2 adrenergic receptor-G protein complex, over 50 agonists were screened for complex stabilization and BI-167107 was determined to provide significant stabilization of the G-protein complex with a half-time of ~30 hours. To form the complex, the two proteins were incubated for 3 hours in the presence of agonist. The complex was further stabilized with a nanobody, which was added to the agonist stabilized complex and incubated for 1 hour before crystallization. These two examples highlight the difficulties in forming membrane protein complexes and different approaches used to obtain structures of these complexes.

Summary

In the soluble structure determination field, many techniques have been developed for recalcitrant proteins. All membrane proteins fall into this challenging category and almost every stage of preparation requires screening of conditions. Although daunting in the number of avenues to try, post-expression modifications provide additional options for the membrane protein structural biologists. The approaches described in this review are not specific to membrane proteins; however, their use for membrane proteins has been demonstrated and should be added to the toolkit of membrane protein structure determination. Additional modifications specific to the properties of membrane proteins (e.g. amino acid modifications that increase solubility) are uncharted territory and may lead to significant improvement in the quality and number of membrane protein structures.

Acknowledgements

The research in my laboratory is funded by National Institutes of Health (RO1GM087828), Jeffress Memorial Trust, Research Corporation for Scientific Advancement, and an NSF CAREER award (MCB 0845668).

References and recommended readings

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H. Overcoming barriers to membrane protein structure determination. *Nat Biotechnol.* 2011; 29:335–340. [PubMed: 21478852]
2. Moraes I, Evans G, Sanchez-Weatherby J, Newstead S, Stewart PD. Membrane protein structure determination - the next generation. *Biochim Biophys Acta.* 2014; 1838:78–87. [PubMed: 23860256]
3. Remaut H, Tang C, Henderson NS, Pinkner JS, Wang T, Hultgren SJ, Thanassi DG, Waksman G, Li H. Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell.* 2008; 133:640–652. [PubMed: 18485872]
4. Granier S, Manglik A, Kruse AC, Kobilka TS, Thian FS, Weis WI, Kobilka BK. Structure of the delta- opioid receptor bound to naltrindole. *Nature.* 2012; 485:400–404. [PubMed: 22596164]
5. Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S. Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature.* 2012; 485:321–326. [PubMed: 22437502]
6. Weigelt J, van Dongen M, Uppenberg J, Schultz J, Wikstrom M. Site-selective screening by NMR spectroscopy with labeled amino acid pairs. *J Am Chem Soc.* 2002; 124:2446–2447. [PubMed: 11890787]
7. Whittaker JW. Selective isotopic labeling of recombinant proteins using amino acid auxotroph strains. *Methods Mol Biol.* 2007; 389:175–188. [PubMed: 17951643]
8. Muona M, Aranko AS, Raulinaitis V, Iwai H. Segmental isotopic labeling of multi-domain and fusion proteins by protein trans-splicing in vivo and in vitro. *Nat Protoc.* 2010; 5:574–587. [PubMed: 20203672]
9. Fox DA, Columbus L. Solution NMR resonance assignment strategies for beta-barrel membrane proteins. *Protein Sci.* 2013; 22:1133–1140. [PubMed: 23754333]
10. Fox DA, Larsson P, Lo RH, Kroncke BM, Kasson PM, Columbus L. Structure of the Neisserial outer membrane protein Opa₆₀: loop flexibility essential to receptor recognition and bacterial engulfment. *J Am Chem Soc.* 2014; 136:9938–9946. •• First demonstration of limited proteolysis used to facilitate NMR assignments and membrane protein structure determination. [PubMed: 24813921]
11. Edrington TC, Kintz E, Goldberg JB, Tamm LK. Structural basis for the interaction of lipopolysaccharide with outer membrane protein H (OprH) from *Pseudomonas aeruginosa*. *J Biol Chem.* 2011; 286:39211–39223. [PubMed: 21865172]
12. Doyle DA, Morais Cabral J, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT, MacKinnon R. The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity. *Science.* 1998; 280:69–77. •• First demonstration of the use of limited proteolysis for X-ray crystallography structure determination of membrane proteins. [PubMed: 9525859]
13. Cohen SL, Chait BT. Mass spectrometry as a tool for protein crystallography. *Annu Rev Biophys Biomol Struct.* 2001; 30:67–85. [PubMed: 11340052]
14. Dong H, Xiang Q, Gu Y, Wang Z, Paterson NG, Stansfeld PJ, He C, Zhang Y, Wang W, Dong C. Structural basis for outer membrane lipopolysaccharide insertion. *Nature.* 2014; 511:52–56. [PubMed: 24990744]
15. Dreyfus C, Laursen NS, Kwaks T, Zuijdgheest D, Khayat R, Ekiert DC, Lee JH, Metlagel Z, Bujny MV, Jongeneelen M, et al. Highly conserved protective epitopes on influenza B viruses. *Science.* 2012; 337:1343–1348. [PubMed: 22878502]
16. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, et al. New world bats harbor diverse influenza A viruses. *PLoS Pathog.* 2013; 9:e1003657. [PubMed: 24130481]
17. Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, et al. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature.* 2011; 477:549–555. [PubMed: 21772288]
18. Sobolevsky AI, Rosconi MP, Gouaux E. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature.* 2009; 462:745–756. [PubMed: 19946266]

19. Sui H, Walian PJ, Tang G, Oh A, Jap BK. Crystallization and preliminary X-ray crystallographic analysis of water channel AQP1. *Acta Crystallogr D Biol Crystallogr*. 2000; 56:1198–1200. [PubMed: 10957645]
20. Hassaine G, Deluz C, Grasso L, Wyss R, Tol MB, Hovius R, Graff A, Stahlberg H, Tomizaki T, Desmyter A, et al. X-ray structure of the mouse serotonin 5-HT₃ receptor. *Nature*. 2014; 512:276–281. •• A tour de force of approaches (including limited proteolysis and deglycosylation) to obtain the X-ray crystallography structure of the serotonin 5-HT₃ receptor. [PubMed: 25119048]
21. Mathiasen S, Christensen SM, Fung JJ, Rasmussen SG, Fay JF, Jorgensen SK, Veshaguri S, Farrens DL, Kiskowski M, Kobilka B, et al. Nanoscale high-content analysis using compositional heterogeneities of single proteoliposomes. *Nat Methods*. 2014; 11:931–934. [PubMed: 25086504]
22. Egloff P, Hillenbrand M, Klenk C, Batyuk A, Heine P, Balada S, Schlinkmann KM, Scott DJ, Schutz M, Pluckthun A. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proc Natl Acad Sci U S A*. 2014; 111:E655–662. [PubMed: 24453215]
23. White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, et al. Structure of the agonist-bound neurotensin receptor. *Nature*. 2012; 490:508–513. [PubMed: 23051748]
24. Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, et al. Structure of the human glucagon class B G-protein-coupled receptor. *Nature*. 2013; 499:444–449. [PubMed: 23863937]
25. Hollenstein K, Kean J, Bortolato A, Cheng RK, Dore AS, Jazayeri A, Cooke RM, Weir M, Marshall FH. Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature*. 2013; 499:438–443. [PubMed: 23863939]
26. Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. 2010; 330:1066–1071. [PubMed: 20929726]
27. Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Li J, et al. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*. 2013; 341:1387–1390. [PubMed: 24030490]
28. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzyanov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*. 2008; 454:486–491. [PubMed: 18594507]
29. Bokoch MP, Zou Y, Rasmussen SG, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS, et al. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature*. 2010; 463:108–112. •• Demonstration of lysine methylation for the use of investigating conformational states of membrane proteins. [PubMed: 20054398]
30. Rayment I. Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods Enzymol*. 1997; 276:171–179. [PubMed: 9048376]
31. Schubot FD, Waugh DS. A pivotal role for reductive methylation in the de novo crystallization of a ternary complex composed of *Yersinia pestis* virulence factors YopN, SycN and YscB. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:1981–1986. [PubMed: 15502305]
32. Sledz P, Zheng H, Murzyn K, Chruszcz M, Zimmerman MD, Chordia MD, Joachimiak A, Minor W. New surface contacts formed upon reductive lysine methylation: improving the probability of protein crystallization. *Protein Sci*. 2010; 19:1395–1404. [PubMed: 20506323]
33. Tan K, Kim Y, Hatzos-Skintges C, Chang C, Cuff M, Chhor G, Osipiuk J, Michalska K, Nocek B, An H, et al. Salvage of failed protein targets by reductive alkylation. *Methods Mol Biol*. 2014; 1140:189–200. [PubMed: 24590719]
34. Walter TS, Meier C, Assenberg R, Au KF, Ren J, Verma A, Nettleship JE, Owens RJ, Stuart DI, Grimes JM. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure*. 2006; 14:1617–1622. [PubMed: 17098187]
35. Kim Y, Quartey P, Li H, Volkart L, Hatzos C, Chang C, Nocek B, Cuff M, Osipiuk J, Tan K, et al. Large-scale evaluation of protein reductive methylation for improving protein crystallization. *Nat Methods*. 2008; 5:853–854. [PubMed: 18825126]

36. Warne T, Moukhametzianov R, Baker JG, Nehme R, Edwards PC, Leslie AG, Schertler GF, Tate CG. The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature*. 2011; 469:241–244. [PubMed: 21228877]
37. Lee CH, Lu W, Michel JC, Goehring A, Du J, Song X, Gouaux E. NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature*. 2014; 511:191–197. •• Use of disulfide crosslinking to reduce conformational flexibility to improve diffraction quality and EndoH to produce homogenous N-acetylglucosamine modified asparagine residues for crystal contacts. [PubMed: 25008524]
38. Hohl M, Hurlimann LM, Bohm S, Schoppe J, Grutter MG, Bordignon E, Seeger MA. Structural basis for allosteric cross-talk between the asymmetric nucleotide binding sites of a heterodimeric ABC exporter. *Proc Natl Acad Sci U S A*. 2014; 111:11025–11030. [PubMed: 25030449]
39. Korkhov VM, Mireku SA, Locher KP. Structure of AMP-PNP-bound vitamin B12 transporter BtuCD-F. *Nature*. 2012; 490:367–372. [PubMed: 23000901]
40. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, et al. Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature*. 2011; 469:236–240. [PubMed: 21228876]
41. Reyes N, Ginter C, Boudker O. Transport mechanism of a bacterial homologue of glutamate transporters. *Nature*. 2009; 462:880–885. •• Use of mercury crosslinking to trap an inward facing conformation. [PubMed: 19924125]
42. Verdon G, Boudker O. Crystal structure of an asymmetric trimer of a bacterial glutamate transporter homolog. *Nat Struct Mol Biol*. 2012; 19:355–357. [PubMed: 22343718]
43. Zimmer J, Nam Y, Rapoport TA. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature*. 2008; 455:936–943. [PubMed: 18923516]
44. Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*. 2000; 405:914–919. [PubMed: 10879525]
45. Wang Y, Zhang Y, Ha Y. Crystal structure of a rhomboid family intramembrane protease. *Nature*. 2006; 444:179–180. [PubMed: 17051161]
46. Chandran V, Fronzes R, Duquerroy S, Cronin N, Navaza J, Waksman G. Structure of the outer membrane complex of a type IV secretion system. *Nature*. 2009; 462:1011–1015. [PubMed: 19946264]
47. Ho JD, Yeh R, Sandstrom A, Chorny I, Harries WE, Robbins RA, Miercke LJ, Stroud RM. Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance. *Proc Natl Acad Sci U S A*. 2009; 106:7437–7442. [PubMed: 19383790]
48. Phan G, Remaut H, Wang T, Allen WJ, Pirker KF, Lebedev A, Henderson NS, Geibel S, Volkan E, Yan J, et al. Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature*. 2011; 474:49–53. [PubMed: 21637253]
49. Santos JS, Asmar-Rovira GA, Han GW, Liu W, Syeda R, Cherezov V, Baker KA, Stevens RC, Montal M. Crystal structure of a voltage-gated K⁺ channel pore module in a closed state in lipid membranes. *J Biol Chem*. 2012; 287:43063–43070. [PubMed: 23095758]
50. Levin EJ, Cao Y, Enkavi G, Quick M, Pan Y, Tajkhorshid E, Zhou M. Structure and permeation mechanism of a mammalian urea transporter. *Proc Natl Acad Sci U S A*. 2012; 109:11194–11199. [PubMed: 22733730]
51. Geibel S, Procko E, Hultgren SJ, Baker D, Waksman G. Structural and energetic basis of folded-protein transport by the FimD usher. *Nature*. 2013; 496:243–246. [PubMed: 23579681]
52. Cao B, Zhao Y, Kou Y, Ni D, Zhang XC, Huang Y. Structure of the nonameric bacterial amyloid secretion channel. *Proc Natl Acad Sci U S A*. 2014; 111:E5439–5444. [PubMed: 25453093]
53. Zhao Y, Mao G, Liu M, Zhang L, Wang X, Zhang XC. Crystal structure of the E. coli peptide transporter YbgH. *Structure*. 2014; 22:1152–1160. [PubMed: 25066136]
54. Kodan A, Yamaguchi T, Nakatsu T, Sakiyama K, Hipolito CJ, Fujioka A, Hirokane R, Ikeguchi K, Watanabe B, Hiratake J, et al. Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proc Natl Acad Sci U S A*. 2014; 111:4049–4054. [PubMed: 24591620]

55. Sui H, Han BG, Lee JK, Walian P, Jap BK. Structural basis of water-specific transport through the AQP1 water channel. *Nature*. 2001; 414:872–878. • First use of deglycosylation to improve crystallization of a membrane protein. [PubMed: 11780053]
56. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, et al. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*. 2007; 450:383–387. •• First crystal structure of a GPCR activated by diffusible ligands. [PubMed: 17952055]
57. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science*. 2008; 322:1211–1217. [PubMed: 18832607]
58. Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science*. 2010; 330:1091–1095. [PubMed: 21097933]
59. Srivastava A, Yano J, Hirozane Y, Kefala G, Gruswitz F, Snell G, Lane W, Ivetac A, Aertgeerts K, Nguyen J, et al. High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. *Nature*. 2014; 513:124–127. [PubMed: 25043059]
60. Zhang K, Zhang J, Gao ZG, Zhang D, Zhu L, Han GW, Moss SM, Paoletta S, Kiselev E, Lu W, et al. Structure of the human P2Y12 receptor in complex with an antithrombotic drug. *Nature*. 2014; 509:115–118. [PubMed: 24670650]
61. Dürr KL, Chen L, Stein RA, De Zorzi R, Folea IM, Walz T, McHaourab HS, Gouaux E. Structure and dynamics of AMPA receptor GluA2 in resting, pre-open, and desensitized states. *Cell*. 2014; 158:778–792. [PubMed: 25109876]
62. Yin J, Mobarec JC, Kolb P, Rosenbaum DM. Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. *Nature*. 2015; 519:247–250. [PubMed: 25533960]
63. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*. 2007; 318:1258–1265. [PubMed: 17962520]
64. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, et al. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science*. 2007; 318:1266–1273. •• First use of cysteine alkylation to improve yields and homogeneity of protein preparation. [PubMed: 17962519]
65. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, et al. Structure of the human histamine H1 receptor complex with doxepin. *Nature*. 2011; 475:65–70. [PubMed: 21697825]
66. Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, Reinhart G, Desale H, Clemons B, Cahalan SM, Schuerer SC, et al. Crystal structure of a lipid G protein-coupled receptor. *Science*. 2012; 335:851–855. [PubMed: 22344443]
67. Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang XP, Carroll FI, et al. Structure of the human kappa-opioid receptor in complex with JDTic. *Nature*. 2012; 485:327–332. [PubMed: 22437504]
68. Thompson AA, Liu W, Chun E, Katritch V, Wu H, Vardy E, Huang XP, Trapella C, Guerrini R, Calo G, et al. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature*. 2012; 485:395–399. [PubMed: 22596163]
69. Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer D, Zheng Y, Green HF, Pandey A, Dror RO, Shaw DE, et al. High-resolution crystal structure of human protease-activated receptor 1. *Nature*. 2012; 492:387–392. [PubMed: 23222541]
70. Kruse AC, Hu J, Pan AC, Arlow DH, Rosenbaum DM, Rosemond E, Green HF, Liu T, Chae PS, Dror RO, et al. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature*. 2012; 482:552–556. [PubMed: 22358844]
71. Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, Hubner H, Pardon E, Valant C, Sexton PM, et al. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*. 2013; 504:101–106. [PubMed: 24256733]

72. Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, Han GW, Liu W, Huang XP, Vardy E, et al. Structural basis for molecular recognition at serotonin receptors. *Science*. 2013; 340:610–614. [PubMed: 23519210]
73. Wacker D, Wang C, Katritch V, Han GW, Huang XP, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, et al. Structural features for functional selectivity at serotonin receptors. *Science*. 2013; 340:615–619. [PubMed: 23519215]
74. Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC. Structure of the human smoothened receptor bound to an antitumour agent. *Nature*. 2013; 497:338–343. [PubMed: 23636324]
75. Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, Niswender CM, Katritch V, Meiler J, Cherezov V, et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science*. 2014; 344:58–64. [PubMed: 24603153]
76. Dore AS, Okrasa K, Patel JC, Serrano-Vega M, Bennett K, Cooke RM, Errey JC, Jazayeri A, Khan S, Tehan B, et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature*. 2014; 511:557–562. [PubMed: 25042998]
77. Burg JS, Ingram JR, Venkatakrishnan AJ, Jude KM, Dukkupati A, Feinberg EN, Angelini A, Waghay D, Dror RO, Ploegh HL, et al. Structural biology. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science*. 2015; 347:1113–1117. [PubMed: 25745166]
78. Qin L, Kufareva I, Holden LG, Wang C, Zheng Y, Zhao C, Fenalti G, Wu H, Han GW, Cherezov V, et al. Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science*. 2015; 347:1117–1122. [PubMed: 25612609]

Highlights

Post-expression modifications have enabled membrane protein structure determination.

Limited proteolysis removes flexible regions.

Deglycosylation removes heterogeneous post-translational modifications.

Cysteine alkylation prevents disulfide dimers and higher order oligomers.

Other modifications allow the structure determination of alternative conformations.

Figure 1.

Post-expression modifications that have contributed to membrane protein structural investigations. Limited proteolysis cleaves exposed and dynamic regions of membrane proteins and has been applied to membrane proteins for x-ray crystallography structure determination and solution NMR assignments. The scissor icon represents a protease, which is commonly chymotrypsin or trypsin. Glycosidases are used to remove post-translational glycosylations (deglycosylation) to facilitate structure determination. The scissor icon represents the site of cleavage for the two types of enzymes. Reductive methylation and cysteine alkylation are covalent modifications that are used to facilitate conformational equilibria or crystallization, respectively.

Figure 2. Limited proteolysis facilitate NMR assignments. (A) SDS-PAGE of Opa₆₀ (Lanes 1–4), Opa₅₀ (Lanes 5–8), and OprH (Lanes 9–12) 8 M urea (unfolded), folded, cleaved with trypsin, and boiled after trypsin cleavage. (B) Full-length (red) and trypsin-treated (blue) ¹⁵N, ¹H-HSQC spectra of ²H, ¹⁵N-labeled Opa₆₀.

Figure 3.

The structure and crystal packing of the NMDA receptor. For one hetero-terramer, GluN1 is colored orange and GluN2B subunit is colored blue. N-acetylglucosamine and modified asparagine residues are rendered with spheres and colored according to atom type. The disulfide crosslink between GluN2B is shown as red spheres. A symmetry mate is colored gray to highlight the N-acetylglucosamine moiety at the crystal contact.

Table 1

X-ray crystal structures determined with post-expression modification *

Protein	Organism of origin	Expression system	PDB	Reference, year
Limited proteolysis				
KcsA	<i>Streptomyces lividans</i>	<i>Escherichia coli</i>	1BL8	[12], 1998
TolC	<i>E. coli</i>	<i>E. coli</i>	1EK0	[44], 2000
GlpG	<i>E. coli</i>	<i>E. coli</i>	2IC8	[45], 2006
VirV7:VirB9:VirB10(T4S)	<i>E. coli</i>	<i>E. coli</i>	3JQO	[46], 2009
AQP4	<i>Homo sapiens</i>	<i>Pichia pastoris</i>	3DG8	[47], 2009
FimD-FimC	<i>E. coli</i>	<i>E. coli</i>	3RFZ	[48], 2011
KvLm PM	<i>Listeria monocytogenes</i>	<i>E. coli</i>	4H33	[49], 2012
UT-B urea transporter	<i>Bos taurus</i>	<i>Spodoptera frugiperda</i>	4EZC	[50], 2012
FimD-FimC-FimF-FimG-FimH	<i>E. coli</i>	<i>E. coli</i>	4J30	[51], 2013
CsgG	<i>E. coli</i>	<i>E. coli</i>	4Q79	[52], 2014
LptD:LptE	<i>Salmonella enterica</i>	<i>E. coli</i>	4N4R	[14], 2014
Serotonin 5-HT ₃	<i>Mus musculus</i>	HEK293F cells	4PIR	[20], 2014
YbgH peptide transporter	<i>E. coli</i>	<i>E. coli</i>	4Q65	[53], 2014
P-glycoprotein	<i>Cyanidioschyzon merolae</i>	<i>Komagataella pastoris</i>	3WME	[54], 2014
Deglycosylation				
AQP1 [‡]	<i>B. taurus</i>	<i>B. taurus</i>	1J4N	[55], 2001
β2 adrenergic receptor [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	2R4R	[56], 2007
A _{2A} adenosine receptor [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	3EML	[57], 2008
Dopamine D3 receptor [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	3PBL	[58], 2010
CXCR4 chemokine receptor [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	30DU	[26], 2010
GPR40 free fatty-acid receptor 1 [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	4PHU	[59], 2014
P2Y ₁₂ [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	4NTJ	[60], 2014
NMDA receptor [‡]	<i>Xenopus laevis</i>	HEK293S <i>GnTI</i> -cells	4TLL	[37], 2014
GluA2 Glutamate receptor (AMPA) [‡]	<i>Rattus norvegicus</i>	HEK293S cells	4U2P	[61], 2014
Serotonin 5-HT ₃ [‡]	<i>M. musculus</i>	HEK293F cells	4PIR	[20], 2014
OX ₂ orexin receptor [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	4S0V	[62], 2015
Cysteine alkylation				
β2 adrenergic receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	2RH1	[63,64], 2007
A _{2A} adenosine receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	3EML	[57], 2008
Dopamine D3 receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	3PBL	[58], 2010

Protein	Organism of origin	Expression system	PDB	Reference, year
Histamine H1 receptor	<i>H. sapiens</i>	<i>P. pastoris</i>	3RZE	[65], 2011
Sphingosine 1-phosphate receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	3V2W	[66], 2012
κ -opioid receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4DJH	[67], 2012
μ -opioid receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4DKL	[5], 2012
δ -opioid receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4EJ4	[4], 2012
Nociceptin/orphanin FQ (N/OFQ) receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4EA3	[68], 2012
Protease-activated receptor 1	<i>H. sapiens</i>	<i>S. frugiperda</i>	3VW7	[69], 2012
M3 muscarinic acetylcholine receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4DAJ	[70], 2012
M2 muscarinic acetylcholine receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4MQS	[71], 2013
5-HT _{1B} serotonin receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4IAR	[72], 2013
5-HT _{2B} serotonin receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4IB4	[73], 2013
Smoothed receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4JKV	[74], 2013
P2Y ₁₂	<i>H. sapiens</i>	<i>S. frugiperda</i>	4NTJ	[60], 2014
Metabotropic Glutamate Receptor 1	<i>H. sapiens</i>	<i>S. frugiperda</i>	4OR2	[75], 2014
Metabotropic Glutamate Receptor 5	<i>H. sapiens</i>	<i>S. frugiperda</i>	4009	[76], 2014
GPR40 free fatty-acid receptor 1 [‡]	<i>H. sapiens</i>	<i>S. frugiperda</i>	4PHU	[59], 2014
US28 chemokine receptor	Human cytomegalovirus	HEK293S <i>GnTI</i> -cells	4XT1	[77], 2015
OX ₂ orexin receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4S0V	[62], 2015
Disulfide crosslinking				
Glutamate transporter (Glt _{ph})	<i>Pyrococcus horikoshii</i>	<i>E. coli</i>	3KBC, 3V8F	[41], 2009 [42], 2012
β 2 adrenergic receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	3PDS	[40], 2011
ABC transporter BtuCD	<i>E. coli</i>	<i>E. coli</i>	4F13	[39], 2012
NMDA receptor	<i>Xenopus laevis</i>	HEK293S <i>GnTI</i> -cells	4TLL	[37], 2014
Heterodimeric ABC exporter TM287/288	<i>Tierrmotoga maritima</i>	<i>E. coli</i>	4Q4J	[38], 2014
CXCR4 chemokine receptor	<i>H. sapiens</i>	<i>S. frugiperda</i>	4RWS	[78], 2015

* For proteins with multiple structures determined, the first report is provided.

[‡] Endoglycosidase H was used,

[‡] PNGaseF was used.