Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

# Post-expression strategies for structural investigations of membrane proteins

Linda Columbus

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.

**Address**
Department of Chemistry, University of Virginia, Charlottesville, VA 22902, United States

Corresponding author: Columbus, Linda (columbus@virginia.edu)

**Current Opinion in Structural Biology** 2015, **32**:131–138

This review comes from a themed issue on **New constructs and expressions of proteins**

Edited by **Imre Berger** and **Roslyn M Bill**

For a complete overview see the Issue and the Editorial

Available online 16th May 2015

http://dx.doi.org/10.1016/j.sbi.2015.04.005



## 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.

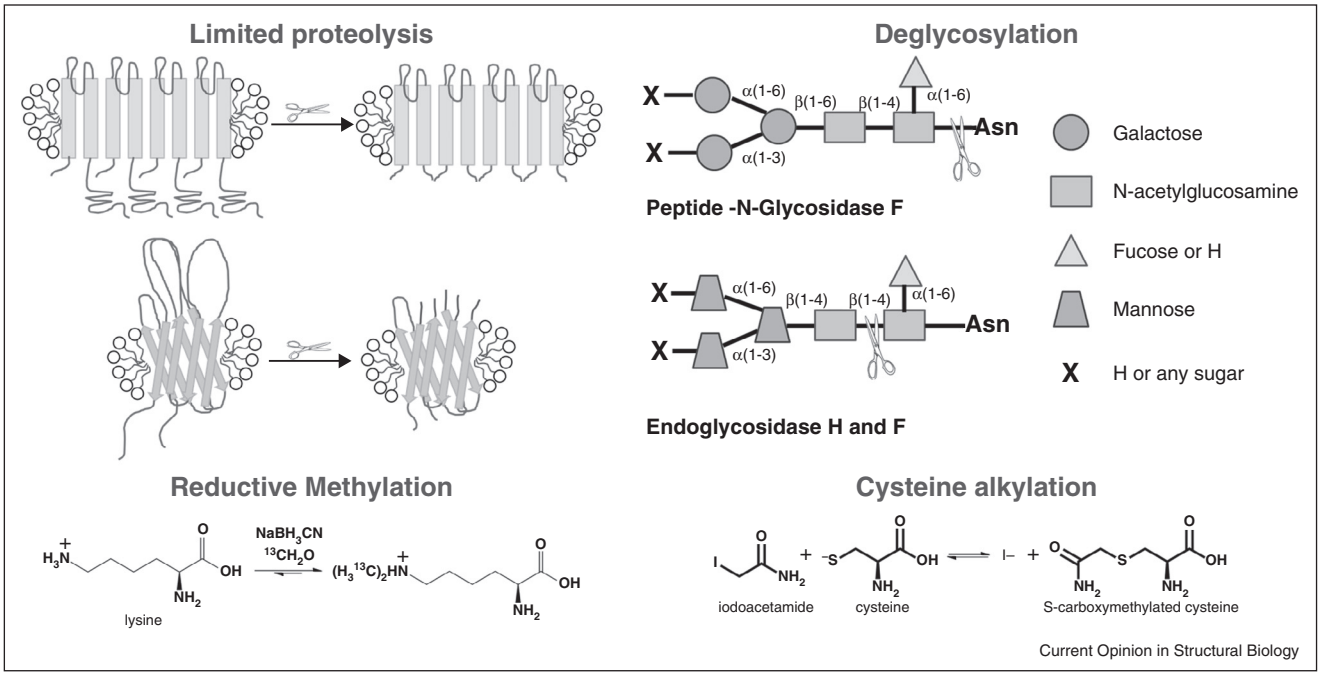
Membrane protein structural biology has required a multi-tier, highly empirical approach to obtain samples that are amenable to the structural techniques commonly used for soluble proteins. Much attention is paid to the construct (e.g. organism, chimeras, and mutations), expression, and membrane mimic selection [1,2]. However, several structures have required post-expression protein modifications. The post-expression modifications improve structure 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. For membrane proteins, the transmembrane regions are protected from protease cleavage by the membrane or membrane mimic (Figure 1). The proteolytic product is typically evaluated using sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) 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 kDa fragment corresponding to the outer membrane translocation domain [3]. The 55 kDa 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-opioid and δ-opioid receptors [4,5]). Beyond tag removal and domain mapping, limited proteolysis has facilitated

**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 crytallogaphy 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 modulate conformational equilibria or facilitate crystallization, respectively.

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; Figure 2a) 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 $Opa_{60}$, 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 may not result in denaturation. For $Opa_{60}$ [10••], $Opa_{50}$, and OprH [11], the resulting five fragments after trypsin proteolysis maintained the β-barrel fold (Figure 2b) and the corresponding resonances were nearly superimposable with that of the untreated proteins [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 $Opa_{60}$ 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