9
Structural Biology of Membrane Proteins

David Salom and Krzysztof Palczewski

9.1
Introduction

Membrane proteins represent about a third of the proteins in living organisms [1], but knowledge of their various functions is hampered by the scarcity of structural information. Owing to their central role in basically all physiological processes, membrane proteins constitute around 60% of approved drug targets [2] and, therefore, their experimentally determined three-dimensional structures are eagerly sought to assist in structure-based drug design. Fortunately, the number of high-resolution membrane protein structures has grown exponentially since the first membrane protein crystal structure was solved [3], although most membrane protein crystal structures solved to date are from bacteria. To explain this bias one might assume that eukaryotic membrane proteins are just more difficult to crystallize. However, the fact that half of all eukaryotic membrane proteins crystallized were purified from native sources speaks of difficulties encountered with their heterologous expression. Of those eukaryotic membrane proteins expressed heterologously, less than one-quarter were expressed in *Escherichia coli* [4]. By contrast, the vast majority of soluble proteins and prokaryotic membrane proteins crystallized thus far were produced as recombinant proteins in bacterial systems [4, 5].

9.2
Folding and Structural Analysis of Membrane Proteins

9.2.1
Folding

The hydrophobic effect is a dominant contributor to the structural stabilization of soluble proteins and extramembranous regions of membrane proteins, but water is essentially absent in the hydrocarbon core of lipid bilayers. Consequently, the relative importance of other forces, such as van der Waals packing and hydrogen bonding, increase in the apolar environment of the membrane core. There are

Production of Membrane Proteins: Strategies for Expression and Isolation, First Edition.
Edited by Anne Skaja Robinson.

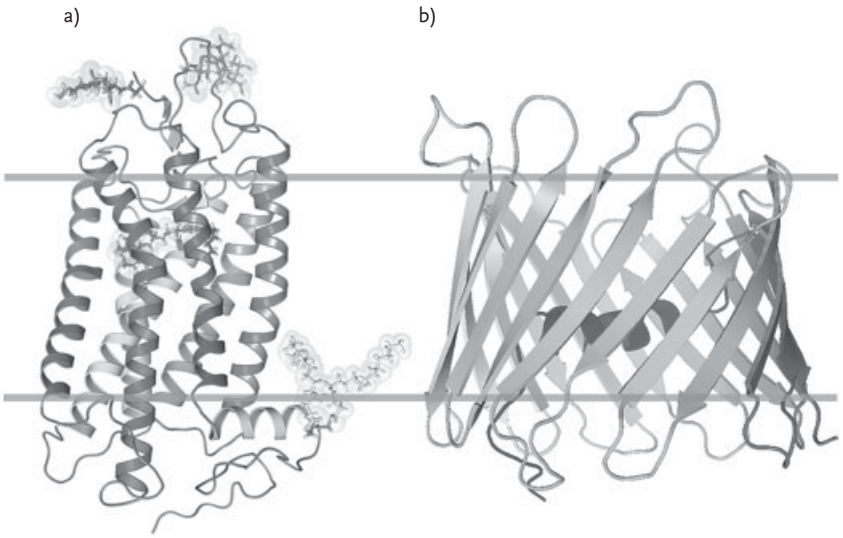

Figure 9.1 Examples of the two structural motives found in membrane proteins: the α-helical bundle (a, bovine rhodopsin) and a β-barrel protein (b, human mitochondrial voltage-dependent anion channel), in their approximate positions in the membrane bilayer. Retinal, carbohydrates, and palmitoyl chains are also shown in the rhodopsin illustration.

virtually no unfolded segments in transmembrane domains of membrane proteins because of the high thermodynamic cost of burying polar peptide bonds in a membrane, which also explains why the membrane interface is a potent catalyst of secondary structure formation. To date, two structural motifs – α-helices and β-barrels – have been found in membrane proteins (Figure 9.1). β-Barrels, made up of β-strands, are found in Gram-negative bacterial outer membranes as well as in mitochondrial and chloroplast membranes, and these structures function as channels or transporters for nutrients, proteins, hydrophobic toxic substances, and other molecules. α-Helical bundles are structurally and functionally more versatile, serving as receptors, channels, transporters, electron transporters, and redox facilitators.

Owing to their quasi-two-dimensional environment, the conformational space of membrane proteins is severely restricted. Over the past few years, the steady increase of published membrane protein structures has turned up numerous examples of structural similarities among apparently unrelated protein families. For example, structural similarity was found between the short acid transporter FocA and aquaporins [6], and between the transmembrane domain of the ionotropic glutamate receptor (iGluR) and bacterial K^+ channels [7]. However, it is still unclear whether these structural similarities reflect a divergence from common ancestral proteins so ancient that sequence similarities have been erased by genetic drift or whether they just represent a manifestation of the restricted folding space available to membrane proteins [8].

9.2.2
Prediction Methods

Early methods for predicting transmembrane helices were based on the premise that a protein segment would partition into a membrane if it was sufficiently hydrophobic and long enough to cross it. Starting with the method of Kyte and Doolittle [9], various algorithms for detecting transmembrane segments were proposed from experimental and computational data. These relied on an amino acid hydrophobicity scale based on the free energy needed to transfer them from hydrophilic to hydrophobic media. Typically, the algorithms searched for stretches of around 20 amino acid residues that could cross the roughly 30-Å hydrophobic core of a membrane. Sequence stretches scoring above a certain propensity threshold would be predicted to constitute transmembrane helices. However, these methods and others developed in the 1990s that relied on statistical inference provide only moderate success rates, and are unable to identify "irregular" structures, such as half helices and re-entrant loops found in the K^+ channel and aquaporins [10]. More sequences and structures then provided new insights, such as: aromatic Trp and Tyr residues tend to cluster near the ends of transmembrane segments; loops connecting transmembrane helices have different amino acid compositions depending on whether they face the inside or outside of the cell; and certain sequence motifs in transmembrane segments exhibit a higher-than-random occurrence [11]. Recently, von Heijne *et al.* proposed two methods based on first principles that employ an experimental scale of position-specific amino acid contributions to the free energy of membrane insertion; when coupled with the "positive-inside" rule [12], this combination predicts the topology of α-helical membrane proteins with performance levels rivaling the best machine learning methods [13]. Finally, a computational approach for optimizing the spatial arrangement of protein structures in lipid bilayers has been developed by minimizing their transfer energies from water to the lipid bilayer. The last method has been applied to all unique transmembrane domains and hundreds of peripheral proteins in the Protein Data Bank (PDB), and the results are available online [14].

Prediction of β-barrel motifs from primary sequences is still difficult because of the presence of short transmembrane stretches. Furthermore, the interior of the barrel is not always hydrophilic. Therefore, searching for alternating polar and nonpolar residues is not a promising approach. The C- and N-terminal β-barrel signature sequences are only moderately successful in identifying proteins with β-barrel motifs [15].

9.2.3
Membrane Insertion

With few exceptions, membrane protein synthesis on ribosomes and their membrane insertion, led by a signal sequence, are concerted steps. In many cases the signal sequence of eukaryotic membrane proteins is removed in the endoplasmic reticulum where their maturation is completed by different post-translational

modifications before export to their final cellular destinations. By contrast, the fold of soluble proteins is normally defined entirely within the sequence itself. It is now clear that the translocon complex itself plays an important role in determining the final membrane protein topology. von Heijne's group developed an experimental method for quantifying sequence-dependent translocon selection of transmembrane helices, which allows us to determine the membrane insertion efficiency of suspected transmembrane sequences [16]. The fact that important charged residues are not conserved among translocons of different species could be one of the reasons why it is so difficult to express eukaryotic membrane proteins in heterologous systems [17].

9.2.4 Estimating the Molecular Weight of Membrane Proteins

Most methods for protein structural characterization were first developed for soluble proteins and then adapted to membrane proteins. Thus, in many cases, cautious interpretation of experimental results is advisable when a membrane protein is investigated. For example, one of the most used biophysical methods for protein isolation is sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), which generally provides useful information about protein purity and molecular weight. However, this method often provides misleading information when used to analyze membrane proteins. In contrast to the relatively unstructured SDS-induced unfolded states of most water-soluble proteins, unfolded states of membrane proteins contain a significant percentage of secondary structure [18]. In addition, native (e.g., glycophorin A) or irreversible (e.g., G-protein-coupled receptors (GPCRs)) oligomerization may occur in SDS micelles. Also, if not boiled prior to electrophoresis, OmpA migrates to different positions on SDS–PAGE depending on the compactness of its structure. Native OmpA migrates with an apparent molecular weight of around 30 kDa, whereas completely unfolded OmpA migrates as an around 35-kDa protein [19]. In some cases, differential binding of SDS to membrane proteins during electrophoresis has been suggested to cause deviations up to ±50% in their apparent molecular weights [20]. Similar artifacts can arise when size-exclusion chromatography (SEC) is used to separate membrane proteins.

9.2.5 Amino Acid Composition

From the first crystal structure of a membrane protein [21] it was clear that the distribution of protein residues along the lipid bilayer depended on their depth in the membrane. Leu, Ile, and Phe residues were the most abundant in the acyl chain areas of membrane lipids, Trp and Tyr residues in the interface, and polar residues in the aqueous zone. A decade later, analysis of available membrane protein crystal structures, as well as work with model peptides, confirmed these results [22, 23]. This observed hydrophobic distribution applies only to those resi-

dues facing the lipid acyl chains, because the interiors of α-helix membrane proteins are not significantly more hydrophobic than those of soluble proteins, although transmembrane domains bury smaller residues on average [24].

Trp and Tyr residues have such a marked tendency to locate in the interfacial area that most membrane proteins have what is known as an "aromatic belt." Whenever a hydrophobic mismatch occurs, the protein and/or the membrane can change their structures in order to match the length of their respective hydrophobic areas [25–28]. Lys and Arg residues, with their long and flexible side-chains, can "snorkel" towards the lipid/water interface region, where their positive charge can interact with the negatively charged phosphate groups of phospholipid [23]. Similarly, lipid molecules located close to a transmembrane helix can adapt to the presence of polar residues and water molecules can help solvate polar groups located well within the bilayer plane [29–32].

9.2.6
Transmembrane Helix Association Motifs and Membrane Protein Oligomerization

Consideration of individual transmembrane α-helices as independent folding units has facilitated a "divide-and-conquer" approach to the study of membrane protein assembly, where interactions between peptides recapitulate membrane-embedded portions of helical membrane proteins [33]. The utility of peptides as tools for investigating membrane protein folding originates from a series of early experiments that revealed that the native protein fold of bacteriorhodopsin could be regenerated from various fragments – an indication that transmembrane helix–helix interactions alone have sufficient specificity to generate tertiary and quaternary structures [34, 35]. The "divide-and-conquer" strategy has been successfully applied to the structural analysis of transmembrane segments of a number of membrane proteins [36, 37], such as the structural determination of the M2 proton channel from influenza A virus [38]. In addition, valuable structural information can be obtained in some instances from small transmembrane peptides or membrane protein fragments in organic solvents.

Work with model peptides has allowed investigating the thermodynamics and sequence dependence of transmembrane helix–helix association, for example, showing that polar residues can drive oligomerization of transmembrane helices [39–41]. Classic work by Engelman's group with glycophorin A showed that the GX_3G sequence also drives the association of transmembrane helices [42]; subsequently, this motif was found in both water-soluble [43] and membrane proteins [44]. In glycophorin the helices cross with a right-handed crossing angle of around 40°. Since then, a number of transmembrane helix self-association motifs have been determined experimentally [36] and an exhaustive analysis of helix-packing motifs in membrane proteins was published in 2006 [45]. The later study revealed that the most frequent association motifs showed a tendency to segregate small residues to the helix–helix interface every seven or four residues in the sequence, driving tight helix–helix associations. Thus, the universe of common transmembrane helix-pairing motifs is relatively simple [46].

Crystallographic and functional studies have demonstrated that a large number of membrane proteins can form homo- and hetero-oligomeric structures or even may require oligomerization for function. Some cases are obvious, such as tetrameric or pentameric ion channels where the pore runs through the symmetry axis of the helix bundle. In extreme cases, such as the photosystem II/light harvesting complex II in grana membranes, bacteriorhodopsin in purple membranes, and rhodopsin in rod outer segments (ROS), membrane proteins can be packed at high density in para-crystalline arrays, in order to increase the surface area capable of harvesting photons. There is mounting evidence for many GPCRs that homo- and heterodimers (and even larger oligomers) play an important role in the GPCR activity cycle. In some cases, GPCR hetero-oligomers give rise to pharmacological properties that differ from those of their individual GPCR components.

In the case of rhodopsin, the size of the minimum *in vivo* functional unit is still a matter of debate. Our laboratory, using transmission electron microscopy to study the effect of *n*-alkyl-β-D-maltoside (C_nM) chain length on the oligomerization state of rhodopsin, found that in micelles containing C_{12}M, rhodopsin exists as a mixture of monomers and dimers in equilibrium depending on the protein/detergent ratio. In C_{14}M and C_{16}M rhodopsin forms higher-ordered structures and, especially in C_{16}M, most of the particles are present in tightly packed rows of dimers similar to those seen in ROS membranes. The fact that the activity of rhodopsin increased with its oligomerization state in detergents suggests that oligomerization of GPCRs may be crucial for signal transduction [47]. However, in subsequent experiments we reconstituted monomeric bovine rhodopsin into an apolipoprotein A-I phospholipid particle derived from high-density lipoprotein. Photoactivation of rhodopsin in these particles resulted in rapid activation of transducin at a rate comparable with that found in native ROS and 20-fold faster than rhodopsin in detergent micelles. These data suggest that monomeric rhodopsin is the minimal functional unit for G-protein activation and that oligomerization is not absolutely required for this process [48]. Two points about these studies with nanoparticles, however: (i) they do not answer the real question (i.e., how does rhodopsin activate transducin in its native para-crystalline arrangement) and (ii) the structures of these nanoparticles are much more complex than initially thought [49].

9.2.7
Post-Translational Modifications

There are more than 150 known post-translational modifications in proteins [50]. Reversible and regulated post-translational modifications include phosphorylation, acetylation, *S*-palmitoylation, and ubiquitination. Prokaryotic systems lack the cellular machinery needed for post-translational modifications. Heterogeneous modification is a potential problem because the protein may run as several bands or a smear during electrophoresis, making it difficult to assess its identity, molecular weight and purity. In addition, heterogeneous modifications may inhibit the growth of protein crystals. For example, heterogeneous glycosylation would

prevent growth of crystal lattices involving specific carbohydrate–carbohydrate contacts, as was the case for two rhodopsin crystal forms grown in our laboratory (Figure 3 in [51]).

9.2.7.1 Glycosylation

The most prevalent post-translational modification of plasma membrane and secretory proteins is *N*-linked glycosylation [52]. This linkage participates in many important biological processes such as protein folding, intracellular targeting, immune responses, cell adhesion, and protease resistance. However, such glycosylation is also the most frequent source of heterogeneity in membrane proteins heterologously expressed in eukaryotic systems. In our laboratory, many GPCRs have been stably expressed in HEK-293 cells, but all, with the exception of the CB2 cannabinoid receptor, exhibited markedly heterogeneous (hyper)glycosylation (unpublished results).

Several strategies have been used to minimize or avoid heterogeneous glycosylation, the most common being enzymatic or mutational removal of the carbohydrates, or prior addition of the glycosylation inhibitor, tunicamycin. However, although many successful experiments have been performed with deglycosylated proteins, deglycosylation may negatively affect the expression, stability, folding, and/or and function of membrane proteins.

Several glycosylation-deficient expression systems also have been developed to address this problem [53]. Our laboratory developed a GPCR expression system that produces homogeneous glycosylated GPCRs by expressing them in mouse rod cells under the influence of the rhodopsin promoter [54, 55].

Rhodopsin's glycosylation pattern in rod cells is exquisitely homogeneous (Figure 9.2). However, when rhodopsin was heterologously expressed in liver [55], yeast [56], or mammalian cell cultures [57], heterogeneous glycosylation was observed. Rhodopsin in bovine retina is glycosylated at residues N2 and N15 [58]. Mutations of these residues or their surroundings (especially T4 and T17) can cause different degrees of retinal degeneration. Disruption of the normal N-terminal structure could result in loss of disk structure, potentially attributable to poor packing efficiency of rhodopsin within these membranes or decreased ability of rhodopsin to self-associate [52].

9.2.7.2 Palmitoylation

Palmitoylation consists of the reversible addition of a 16-carbon saturated fatty acyl chain, usually to cytoplasmic Cys residues via a thioester linkage (*S*-palmitoylation). Many other acyl lipids can be added to proteins as well [59].

Palmitoylation of soluble cytoplasmic proteins has been well documented to regulate their interaction with specific membranes or membrane domains. Less is known about the consequences of palmitoylation in transmembrane proteins, because of the difficulty of working with membrane proteins and the complexity of palmitoylation-induced behavior. Palmitoylation has been reported to be involved in regulation of membrane protein folding and targeting, trafficking, and protein–protein interactions [60]. It also has been suggested that palmitoylation

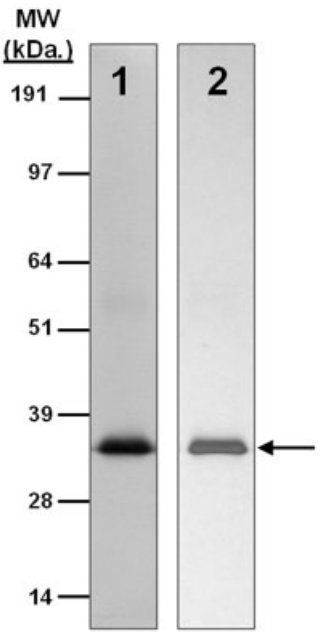

Figure 9.2 Bovine rhodopsin purified with immobilized 1D4 antibody. Lane 1: immunoblot with 1D4 used as the primary antibody followed by an alkaline phosphatase-conjugated secondary antibody. Signals were visualized by treatment with BCIP/NBT Color Development Substrate (Promega). Lane 2: silver-stained SDS–PAGE gel of purified rhodopsin. Arrow shows the position of rhodopsin. Mobility and molecular weight (MW) of markers (SeeBlue® Plus2; Invitrogen) are shown at the left.

determines the orientation of the transmembrane domain with respect to the plane of lipid bilayers [59].

Many GPCRs have been experimentally shown to be mono-, bi-, or tripalmitoylated at conserved C-terminal Cys residues [59]. This trend indicates that these residues and post-translational modification are important for GPCR function and/or trafficking, but there is no consistency among GPCRs about the role that palmitoylation plays in receptor structure and function. We analyzed engineered mice lacking both C-terminal palmitoylated Cys residues of rhodopsin (Cys322 and Cys323) and found that absence of palmitoylation generally only had minor effects [61]. Not surprisingly, the palmitoyl chains adopted different conformations in the seven crystal structures of bovine rhodopsin solved so far.

9.2.8
Sequence Modifications

Most eukaryotic membrane proteins crystallized so far were purified from native sources [4]. However, heterologous expression of membrane proteins is sometimes needed or preferable because of their low expression levels in native sources, instability, difficulties with purification, and so on. As membrane proteins are translocated into membranes starting with their N-terminus during translation, modifications and additions to the C-terminus are in general less dysfunctional than those at the N-terminus. For example, placing a polyhistidine tag at the N-terminus adds several positive charges that may adversely affect membrane

insertion. This said, modifications of the N-terminus constitute a common strategy to improve expression levels. Also, C-terminal length variants such as splicing variants are common and these do not seem to affect function in many cases. This is why the C-terminus is generally preferred over the N-terminus to add tags and fusion proteins for different purposes (e.g., to facilitate purification, detection, crystallization, etc.).

We successfully expressed many different GPCRs in rod cells of *Xenopus* under the rhodopsin promoter with Green Fluorescent Protein (GFP) fused to their C-terminal for easy detection [62]. Recently we also expressed a 5-HT_{4R}–DsRed fusion protein in a *Sf*9/baculovirus system and found, despite a reasonable expression level, that most of the fusion protein was degraded by cleavage between its two protein components (unpublished results).

Fluorescence-activated cell sorting has been used to sort and enrich the best cell lines expressing membrane protein–GFP fusions [63]. The Gouaux group has used GFP fusions combined with fluorescence-detection SEC quite successfully to rapidly evaluate the expression level, stability, optimal detergent composition, and monodispersity of a number of membrane proteins expressed in HEK-293 cells [64]. Before crystallization trials, these constructs are cleaved with a specific protease and the GFP removed by SEC.

One strategy designed to enhance plasma membrane expression of GPCRs in heterologous cells consists of adding leader sequences to these receptors. The first example of this approach was the engineering of an artificial signal sequence onto the N-terminus of the β_2-adrenoceptor [65], that resulted in a several-fold increase in insertion of this receptor into the plasma membrane. Similarly, the addition of an N-terminal signal peptide or the partial truncation of the exceptionally long N-terminus of the cannabinoid CB1 receptor increased its expression level, with no effect on ligand binding. In the same line, truncation of 79 amino acids from the N-terminus of the α_{1D}-adrenergic receptor enhanced its expression. By contrast, grafting of the α_{1D}-adrenergic receptor N-terminus onto α_{1A}-adrenergic receptor or α_{1B}-adrenergic receptor reduced surface expression of these receptors in heterologous cells [66]. However, addition of a signal sequence provides no assurance that the protein will be expressed correctly and targeted to the plasma membrane [67]. In summary, N- and C-terminal modifications can be neutral in terms of membrane protein folding and function, but N-terminal modifications generally have the highest impact on membrane protein expression.

A more important and underappreciated problem occurs when researchers introduce sequence modifications in a membrane protein to stabilize it, reduce polydispersity, and eliminate heterogeneity as a strategy to increase the probability of obtaining crystals. Although important structural information can be gained from such engineered membrane proteins, caution should be exercised in extrapolating it to the wild-type protein. For example, a truncated, minimally functional acid-sensing ion channel ASIC1 was recently crystallized that provided more information about this channel's gating and ion selectivity properties than a previous structure of an inactive, more truncated version of ASIC1 [68]. In the case of GPCRs, thermostabilizing mutations, deletions, chemical modifications,

deglycosylation, and/or substitution of a loop with a soluble protein have allowed the crystallization of β-adrenergic receptors and adenosine A_{2A} receptors at the cost of losing their native-like ligand-binding properties [69–71].

9.2.9 Lipids and Water

It is generally known that addition of lipids is sometimes necessary for proper functioning of detergent-solubilized membrane proteins. For example, we have found that phospholipids are needed for the proper formation, stability, and function of the photoactivated rhodopsin–transducin complex [72]. Nevertheless, membrane protein crystallographers had initially assumed that lipids were a detrimental contaminant that should be removed as completely as possible before crystallization trials. However, it is now becoming clear that some proteins crystallize more readily in the presence of a number of specifically bound lipids, whereas others, such as the porins, only crystallize after delipidation [73]. For example, well-diffracting crystals of cytochrome b_6f can be obtained only by adding phospholipids that stabilize the intact complex [74]. Successful crystallization of a heterologously expressed rabbit sarcoplasmic–endoplasmic reticulum Ca^{2+}-ATPase isoform also required addition of native phospholipid [75]. Moreover, we obtained the first high-resolution crystals of rhodopsin in mixed phospholipid/detergent micelles [76].

However, excess lipid may be detrimental for crystal growth. Therefore, purification procedures leading to crystallization need to include optimization of the detergent : lipid : protein ratio. Examples of lipids found in crystal structures include those in photosystem I and II, bacteriorhodopsin, archaerhodopsin-2, spinach major light harvesting complex, ATP-binding cassette transporter, cytochrome b_6f and bc_1 complexes ([77] and references therein) and rhodopsin [78]. In addition, two crystal structures of the β_2-adrenergic receptor bound to cholesterol along with a sequence analysis have led to the hypothesis that most family A GPCRs have a cholesterol-binding motif [69, 79].

This is why it is hardly surprising that addition of lipids during or after chromatographic purification of membrane proteins is sometimes necessary or helpful in order to maintain their function and/or grow crystals. In the case of glycerol-3-phosphate transporter (GlpT) from *E. coli*, the extent of delipidation, that increases with successive purification steps, is critical for its crystallization. Two chromatography steps were found to be optimal wherein GlpT was found to copurify with about equal amounts of phosphatidylethanolamine, phosphatidylglycerol, and cardiolipin, at a total phospholipid/protein molar ratio of around 23 [80].

A systematic study of the effect of phospholipids on crystallization of lactose permease (LacY) resulted in three different crystal forms that diffracted to increasingly better resolution in a manner that correlated with the concentration of copurified phospholipid. Consistently, progressive addition of *E. coli* phospholipids to delipidated LacY led to different crystal forms. Tetragonal crystals were obtained with improved diffraction quality for a stable mutant by carefully adjusting phos-

pholipid content. Furthermore, crystals of good quality from wild-type LacY, a particularly difficult protein to crystallize, were also obtained by using the same approach [81, 82]. Thus, proper adjustment of the phospholipid environment is a good strategy for crystallizing membrane proteins.

Comparisons of high-resolution crystals of bacterial and mammalian cytochrome *c* oxidases have revealed that the positions occupied by native membrane lipids and detergent substitutes are highly conserved, along with amino acid residues in their vicinity. Well-defined detergent head-groups (maltose) were found associated with aromatic residues in a manner similar to phospholipid head groups, and this likely contributes to the success of alkyl-glycoside detergents in supporting membrane protein activity and crystallization [83].

Especially interesting are structures obtained by electron diffraction of two-dimensional crystals, which are grown in phospholipid planar bilayers faithfully mimicking cell membranes. Although technical difficulties usually prevent the high resolution expected for three-dimensional crystals, it has been possible to solve the structure of lens-specific aquaporin-0 (AQP0) at 1.9-Å resolution. This allowed atomic modeling of the lipid bilayer surrounding AQP0 tetramers and a description of the lipid–protein interactions. Indeed, close inspection of this model revealed that lipids bridge all the contacts between AQP0 tetramers within a layer and that these tetramers have virtually no direct lateral interaction [84].

Finally, an alternative to crystallization *in surfo* (in surfactant micelles) is to crystallize membrane proteins in bicelles (diskoidal phospholipid/detergent micelles) or to use *in cubo* methods wherein the protein is purified in detergent micelles and then embedded into a lipidic cubic phase [85]. However, the last method has the disadvantage of a high viscosity that makes crystal detection and collection challenging.

On the other side of the polarity spectrum, water is basically absent from the hydrophobic interior of phospholipid bilayers and also until recently was believed to be absent from the interior of membrane proteins, with the obvious exception of channel lumens. However, high-resolution crystal structures of a number of membrane proteins clearly show electron densities corresponding to water molecules that can act as cofactors or prosthetic groups.

High-resolution crystals of bacteriorhodopsin in different signaling states along with spectroscopic studies have provided information about the cofactor role of several water molecules in this light-activated proton pump [86]. We analyzed the distribution of water molecules in the transmembrane region of all available GPCR structures and found conserved contacts with microdomains shown to be involved in receptor activation [87]. Armed with this knowledge, we used radiolytic labeling to identify structural water molecules in different GPCRs. Radiolytic labeling (or footprinting) involves generating hydroxyl radicals with synchrotron X-ray white light and then using liquid chromatography coupled to mass spectrometry to identify which residues react with those radicals. This method also can provide valuable information about the role of structural water in different conformational states of a membrane protein, either in bilayers or membrane mimetic systems [88].

9.2.10
Purity and Contaminants

Common experience dictates that proteins need to be purified to homogeneity for their crystallization in the presence of maximally purified chemicals. However, several exceptions to this precept exist wherein impurities not only do not seem to affect crystal growth, but may even be essential for it.

For the crystallization of urea transporter at 3.8 Å, the last purification step consists of gel filtration on a Superdex 200 10/300 (GE Health Sciences) while changing the detergent to 40 mM low-purity C_8M (Sol-Grade from Anatrace). Curiously, crystals grown from protein purified in high-purity C_8M (AnaGrade) failed to diffract to better than 4.5–5 Å [89].

As noted below with retinal pigmented epithelium-specific protein 65 kDa (RPE65), high protein purity is not always essential for growth of high-resolution crystals. Obtaining high-purity preparations can be especially difficult when the protein is obtained from native sources and purified by nonaffinity chromatographic methods. In a systematic study it was found that the photosynthetic reaction center was able to form crystals at 75% purity by using the vapor diffusion method (and just 50% purity for lipid cubic phase crystallization) [90].

Finally, gramicidin (gA′ or gD) is commercially available as a mixture of gA (Trp in position 11) and a small amount of gB (Phe11) and gC (Tyr11). Analysis of a high-resolution crystal structure of gA suggested that a small percentage of gC present in the sample acted as a nucleation agent for gA crystallization [91].

9.2.11
Current Trends in the Crystallization of α-Helical Membrane Proteins

In 2008, Iwata's group published an exhaustive study on the detergents and conditions used for crystallization of α-helical [92] and β-barrel [93] membrane proteins. The most successful detergents for α-helical membrane protein crystallization were, in this order, *n*-alkyl-β-D-maltosides ($C_{12}M$ and $C_{10}M$), *n*-alkyl-β-D-glucosides (C_8G and C_9G), lauryldimethylamine-*N*-oxide (LDAO), and polyoxyethylene glycol detergents ($C_{12}E_8$ and $C_{12}E_9$). This statistical analysis was facilitated by the Membrane Protein Data Bank [94] and Stephen White's data base of membrane proteins of known three-dimensional structure [3], both available online. High-resolution structures of native human membrane proteins in particular are intensely pursued because of their potential to serve as templates for structure-based drug design. Unfortunately, vertebrate proteins in general seem especially difficult to overexpress and crystallize [95]. Here, we describe a minisurvey performed to discern the latest trends in crystallizing vertebrate α-helical membrane proteins, including only those *unique* structures published in 2008 and 2009. In this order, the most successful detergents for crystallization (usually exchanged during purification) were $C_{12}M$ (three structures), C_9G (two structures), plus C_8G, $C_{11}M$, C_{11}-thio-M, and $C_{12}E_8$ with one structure. Several of these crystals were grown with variable amounts of phospholipids; $C_{12}M$ includes one GPCR structure

crystallized in lipidic cubic phase. The most used expression systems were *Sf*9 (four) cells, *Pichia pastoris* (three), plus *E. coli* and native tissue each with one structure. In conclusion, alkyl-glycosides are still the most successful detergents, and insect cells and yeast are being increasingly successful as expression system for the crystallization of vertebrate membrane proteins.

9.3 Test Cases

9.3.1 Rhodopsin

Protocol 1

Protein type	seven-transmembrane, family A GPCR
Crystal space group	tetragonal $P4_1$
Source	native (bovine retina)
Purification steps	organelle (ROS) fractionation, zinc acetate precipitation of apoprotein
Detergent	*n*-nonyl-β-D-glucoside (C_9G)

Rhodopsin is a photoreceptor membrane protein responsible for visual signal transduction [96]. As such, it is also the pharmacological target for drugs used to treat certain retinal diseases and age-related visual dysfunction. Notably, rhodopsin is a prototypical GPCR [97], by far the most important family of pharmacological receptors because about half of the therapeutic drugs in the market today target GPCRs [98].

Rhodopsin is glycosylated at Asn2 and Asn15 residues [99], palmitoylated at Cys322 and Cys323 residues [100], and acetylated at the N-terminal Met. It also contains eight Ser and Thr residues that are partially phosphorylated in a light-dependent manner [101]. These post-translational modifications are not completely homogeneous, so it was feared that they would negatively affect crystallization of this membrane protein. We addressed this challenge by purifying bovine rhodopsin from native sources without including a chromatographic step. The first and most important purification step was actually accomplished by the cow, because only properly folded rhodopsin is transported from the rod cell's endoplasmic reticulum to ROS where it forms para-crystalline arrays at millimolar concentrations [102]. Once in the laboratory, ROS were mechanically separated from bovine retinas in phosphate buffer (67 mM KP_i, pH 7.0, 1 mM $Mg(OAc)_2$, 0.1 mM EDTA, 1 mM dithiothreitol (DTT)). Then a stepwise sucrose gradient centrifugation allowed enrichment of the ROS at the interface between densities 1.11 and 1.13 g/ml (method adapted from [103]). The ROS suspension was washed several times with water to remove loosely attached proteins. Then, ROS were solubilized with C_9G to yield a final concentration of around 10 mg rhodopsin/ml, 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES), pH 6.3, and around 100 mM

$Zn(OAc)_2$. After 5–10 h incubation at room temperature, solubilized rhodopsin was separated from insoluble Zn^{2+}-opsin complexes by centrifugation. Solubilization of rhodopsin was optimal at a detergent/rhodopsin ratio of around 2.2 (w/w) and, after centrifugation, the sample of rhodopsin in its ground state was pure enough (around 98%) to grow diffracting crystals [104], which led to solving the three-dimensional structure of a GPCR for the first time [76]. All experimental procedures were done in the dark or under dim red light because these crystals were unstable under white light. Such instability most probably occurs because conformational changes upon photoactivation involve areas of rhodopsin engaged in protein–protein contacts that stabilize the crystalline array. Addition of heptanetriol, one of the most used additives for membrane protein crystallization, helped improve the resolution to 2.8 Å. Resolution was later enhanced to 2.2 Å by switching the detergent from C_9G to C_7-thio-G [105].

Protocol 2

Crystal space group	hexagonal $P6_4$
Source	native (bovine retina) and COS-1 cells
Purification steps	organelle (ROS) fractionation, and affinity (concanavalin A), gel-filtration, and anion-exchange chromatographies
Detergent	*n*-octyltetraoxyethylene (C_8E_4), LDAO

After our initial structure was published, the Schertler group independently solved the ground-state rhodopsin structure by using a different purification strategy and crystallization conditions [106, 107]. ROS membranes were solubilized in LDAO and the protein was purified by successive lectin-affinity, SEC and anion-exchange chromatographic steps, while the detergent was partially exchanged to C_8E_4. Furthermore, heterologous expression in COS-1 cells of a thermostable rhodopsin mutant and a similar purification scheme also yielded well-diffracting crystals [108]. In the last case, the cofactor, 11-*cis*-retinal, had to be added at the first step of purification. The space group was reinterpreted from trigonal $P3_1$ to hexagonal $P6_4$ [109].

Protocol 3

Crystal space group	trigonal $P3_112$ and rhombohedral R32
Source	native (bovine retina)
Purification steps	organelle (ROS) fractionation, zinc acetate precipitation of apoprotein, affinity (antibody) chromatography, NH_4SO_4-induced phase separation
Detergent	*n*-nonyl-β-D-glucoside (C_9G)

We designed a new purification scheme with the hope of solving the structure of photoactivated rhodopsin. $C_9G/Zn(OAc)_2$ extraction of rhodopsin from ROS membranes, as described above, was initially omitted, but was added later as an initial purification step in an attempt to improve crystal resolution.

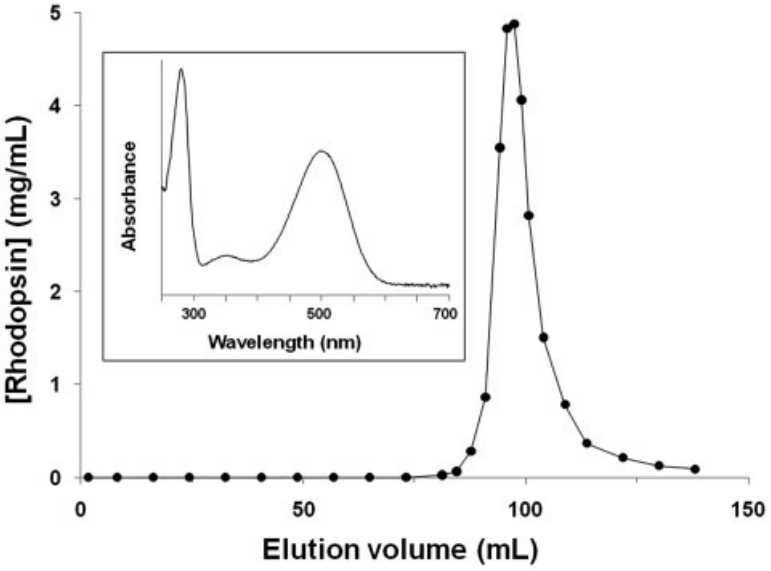

Figure 9.3 Chromatographic purification of bovine rhodopsin on a 2.5-cm × 20-cm column filled with 95 ml of Sepharose-immobilized 1D4 antibody. Inset: absorption spectrum of purified rhodopsin.

An affinity chromatographic support was prepared by immobilizing 1D4 antibody, which recognizes rhodopsin's C-terminus [110], in CNBr-activated Sepharose [111]. A detergent-solubilized ROS sample was loaded onto this affinity column at a ratio of about 0.6 mg rhodopsin/ml 1D4 gel. Then, the column was washed at a linear flow rate of around 18 cm/h with 10 column volumes of 50 mM Tris, pH 7.4, 280 mM NaCl, and 6 mM KCl containing 50 mM C_9G, and eluted with 0.75 mg/ml of a competing peptide (TETSQVAPA) at 7 cm/h (Figure 9.3). The rhodopsin concentration in each fraction was determined by diluting an aliquot into 10 mM Tris, pH 7.2, 1 mM $C_{12}M$, and 10 mM NH_4OH, and measuring the absorbance at 500 nm. Peak fractions reached around 5 mg/ml and the $A_{280\ nm}/A_{500\ nm}$ ratio of the fractions containing ground-state rhodopsin was 1.58 ± 0.2 (Figure 9.3, inset), indicating a purity of around 99% in agreement with electrophoresis results (Figure 9.2).

The most highly purified fractions were pooled together to achieve 1–2 mg/ml rhodopsin. Then, 0.25 volumes of 0.5 M MES, pH 6.3, were added and 0.69 g solid $(NH_4)_2SO_4$/ml of total solution was dissolved by stirring at room temperature. Addition of a high concentration of $(NH_4)_2SO_4$ induces phase separation in many detergent solutions, yielding a top phase rich in detergent where nearly all the rhodopsin partitioned. Treatment of purified rhodopsin with solid $(NH_4)_2SO_4$ resulted in a 12- to 15-fold increase in the concentration of this protein. A typical final yield was 0.2–0.3 mg of purified ground-state rhodopsin per bovine retina.

Crystals obtained from this purification scheme diffracted up to around 4 Å, allowing us to solve the structure of both the ground-state and photoactivated states of rhodopsin in a putatively physiological, parallel dimer orientation [51]. Crystals grown without the $(NH_4)_2SO_4$-induced phase separation step were larger, but diffracted poorly and lost integrity after photo-activation, much like our previous tetragonal crystals [104].

9.3.2 RPE65

Protein type	monotopic enzyme, seven-bladed β-propeller
Crystal space group	hexagonal $P6_5$
Source	native (bovine RPE); crystallization from heterologous expression systems failed
Purification steps	differential centrifugation and anion-exchange chromatography
Detergent	*n*-octyltetraoxyethylene (C_8E_4)

Vertebrate vision is maintained by a complex cyclic enzymatic pathway that continuously operates in the retina to regenerate the visual chromophore, 11-*cis*-retinal. A key enzyme in this pathway is the microsomal membrane protein RPE65, which catalyzes the conversion of all-*trans*-retinyl esters to 11-*cis*-retinol in the retinal pigmented epithelium (RPE). Mutations in RPE65 are known to be responsible of a subset of cases of the most common form of childhood blindness, Leber congenital amaurosis. Like other enzymes with hydrophobic substrates, RPE65 was suspected to be associated with the membrane, but the strength of this association was a matter of debate [112, 113].

9.3.2.1 Expression in *E. coli*

RPE65 was expressed as a fusion protein with maltose-binding protein and purified by metal-affinity chromatography. However, crystals obtained by this method corresponded to GroEL, which copurified with the RPE65 construct [114].

9.3.2.2 Expression in *Sf*9 Cells

RPE65 was purified by metal-affinity chromatography and, although the protein was more than 99% pure after gel-filtration chromatography, crystal trials failed to produce diffracting crystals.

9.3.2.3 Purification from Native Sources

Bovine RPE microsomes were solubilized for 1 h in 10 mM Tris acetate, pH 7.0, containing 1 mM DTT and 24 mM C_8E_4. The mixture was centrifuged at 100 000 *g* for 1 h to pellet insoluble material. RPE65 was purified from the supernatant by anion-exchange chromatography on a DEAE-Macroprep column (Bio-Rad) eluted with a 0–500 mM linear NaCl gradient. The pooled RPE65 sample was dialyzed against 10 mM Tris acetate, pH 7.0, containing 1 mM DTT and 19.2 mM C_8E_4. The

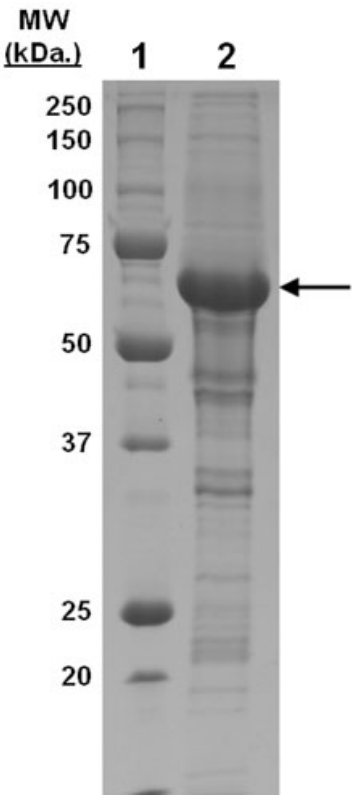

Figure 9.4 Coomassie-stained SDS–PAGE of purified RPE65 used for crystallization trials. Arrow shows the position of RPE65. Lane 1: molecular weight (MW) standards; lane 2: RPE65 sample.

resulting protein preparation was used directly for crystallization trials and biochemical experiments. On average, 150 µl of protein solution was obtained at a concentration of 10–15 mg/ml and a purity of around 90%, as judged by Coomassie-stained gels (Figure 9.4) from 300 bovine eyes. Crystals, obtained by a vapor diffusion method, diffracted to less than 2 Å. The structure of RPE65 revealed a β-propeller architecture, the membrane-binding domain [115], and bound product in the active site.

9.3.3
Transmembrane Domain of M2 Protein from Influenza A Virus

Protein type	proton channel, four-helix bundle
Crystal space group	primitive monoclinic and orthorhombic, $P2_1$ and $P2_12_12$
Source	chemical synthesis
Purification steps	C_4 reverse-phase chromatography
Detergent	*n*-octyl-β-D-glucoside (C_8G)

Flu virus kills hundreds of thousands of people in nonpandemic years worldwide and millions in pandemic years. M2, a proton channel essential for

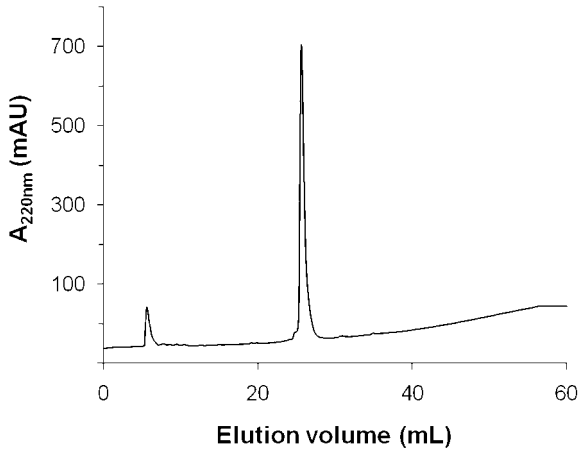

Figure 9.5 Analytical HPLC of purified M2TM on a reverse-phase C_4 column. Elution occurred in a nonlinear gradient of 60–100% buffer B, at 1 ml/min (see text).

viral replication, is the target for the related drugs, amantadine and rimantadine, although resistance to these drugs is now generalized. Both the 97-residue M2 protein (prepared by native chemical ligation of two synthetic peptide segments) and a synthetic 50-residue variant (representing a stable tryptic digest of M2) [116] failed to grow crystals.

Then, we attempted the crystallization of a 25-residue peptide containing the only transmembrane helix of M2 (M2TM) able to recapitulate the tetramerization, drug binding/resistance, low-pH activation, gating, and proton transport properties of the full-length protein [116–118]. Udorn strain wild-type plus multiple natural and artificial single-point mutants of M2TM were prepared by solid-phase peptide synthesis with standard Fmoc chemistry. Purification proceeded by reverse-phase high-performance liquid chromatography (HPLC) with a preparative C_4 column (Vydac) and gradients of buffer B (6 : 3 : 1 2-propanol/acetonitrile/H_2O containing 0.1% trifluoroacetic acid (TFA)) versus 0.1% aqueous TFA. Most peptide variants eluted at 80–84% buffer B. M2TM purity was around 99%, as assessed by analytical HPLC on a C_4 column (Figure 9.5) and its molecular mass was confirmed by matrix-assisted laser desorption/ionization-time of flight mass spectrometry. Purified peptide fractions were lyophilized and the peptide powder was stored at –20 °C or in organic solvent at –80 °C. To prepare samples for biophysical studies, an aliquot of M2TM stock in methanol, ethanol, or trifluoroethanol was transferred to a glass vial, and the organic solvent was evaporated under a N_2 stream and overnight under high vacuum. Finally, a detergent solution in the appropriate buffer (typically 50 mM C_8G in water for crystal trials, to achieve a final peptide concentration of around 1 mM) was added and the sample was vortexed vigorously [38, 116]. Crystals grew robustly from C_8G-solublized samples under different conditions. Crystals of M2TM, free and bound to amantadine, diffracted

at 2.0 and 3.5 Å, respectively, revealing the 4-helix bundle structure of the proton channel and the amantadine-binding site [38].

Acknowledgments

This research was supported in part by grants EY008061, EY019478, GM079191 and P30 EY11373 from the National Institutes of Health. We want to thank to Drs. Leslie T. Webster, Jr, David T. Lodowski, and Ismael Mingarro for valuable comments, Philip Kiser for help in preparing Figure 9.4, and Dr. Amanda L. Stouffer for help in preparing Figure 9.5.

Abbreviations

DTT	dithiothreitol
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled receptor
HPLC	high-performance liquid chromatography
LDAO	lauryldimethylamine-*N*-oxide
M2TM	transmembrane helix of M2
MES	2-(*N*-morpholino)ethanesulfonic acid
PAGE	polyacrylamide gel electrophoresis
PDB	Protein Data Bank
ROS	rod outer segments
RPE	retinal pigmented epithelium
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
TFA	trifluoroacetic acid

References

1 Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.*, **305**, 567–580.

2 Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabasi, A.L., and Vidal, M. (2007) Drug-target network. *Nat. Biotechnol.*, **25**, 1119–1126.

3 White, S.H. (2009) Biophysical dissection of membrane proteins. *Nature*, **459**, 344–346.

4 Carpenter, E.P., Beis, K., Cameron, A.D., and Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.*, **18**, 581–586.

5 Loll, P.J. (2003) Membrane protein structural biology: the high throughput challenge. *J. Struct. Biol.*, **142**, 144–153.

6 Wang, Y., Huang, Y., Wang, J., Cheng, C., Huang, W., Lu, P., Xu, Y.N., Wang, P., Yan, N., and Shi, Y. (2009) Structure of the formate transporter FocA reveals

a pentameric aquaporin-like channel. *Nature*, **462**, 467–472.
7 Sobolevsky, A.I., Rosconi, M.P., and Gouaux, E. (2009) X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature*, **462**, 745–756.
8 Theobald, D.L. and Miller, C. (2010) Membrane transport proteins: surprises in structural sameness. *Nat. Struct. Mol. Biol.*, **17**, 2–3.
9 Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–132.
10 Fleishman, S.J. and Ben-Tal, N. (2006) Progress in structure prediction of alpha-helical membrane proteins. *Curr. Opin. Struct. Biol.*, **16**, 496–504.
11 Elofsson, A. and von Heijne, G. (2007) Membrane protein structure: prediction versus reality. *Annu. Rev. Biochem.*, **76**, 125–140.
12 Nilsson, J., Persson, B., and von Heijne, G. (2005) Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. *Proteins*, **60**, 606–616.
13 Bernsel, A., Viklund, H., Falk, J., Lindahl, E., von Heijne, G., and Elofsson, A. (2008) Prediction of membrane-protein topology from first principles. *Proc. Natl. Acad. Sci. USA*, **105**, 7177–7181.
14 Lomize, M.A., Lomize, A.L., Pogozheva, I.D., and Mosberg, H.I. (2006) OPM: Orientations of Proteins in Membranes database. *Bioinformatics*, **22**, 623–625.
15 Brosig, A., Nesper, J., Boos, W., Welte, W., and Diederichs, K. (2009) Crystal structure of a major outer membrane protein from *Thermus thermophilus* HB27. *J. Mol. Biol.*, **385**, 1445–1455.
16 Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S.H., and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, **433**, 377–381.
17 Bowie, J.U. (2005) Solving the membrane protein folding problem. *Nature*, **438**, 581–589.
18 Booth, P.J. and Curnow, P. (2009) Folding scene investigation: membrane proteins. *Curr. Opin. Struct. Biol.*, **19**, 8–13.
19 Kleinschmidt, J.H., Wiener, M.C., and Tamm, L.K. (1999) Outer membrane protein A of *E. coli* folds into detergent micelles, but not in the presence of monomeric detergent. *Protein Sci.*, **8**, 2065–2071.
20 Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. (2009) Detergent binding explains anomalous SDS–PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. USA*, **106**, 1760–1765.
21 Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985) Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospeudomonas viridis* at 3 Å resolution. *Nature*, **318**, 618–624.
22 von Heijne, G. (1996) Principles of membrane protein assembly and structure. *Prog. Biophys. Mol. Biol.*, **66**, 113–139.
23 Killian, J.A. and von Heijne, G. (2000) How proteins adapt to a membrane–water interface. *Trends Biochem. Sci.*, **25**, 429–434.
24 Oberai, A., Joh, N.H., Pettit, F.K., and Bowie, J.U. (2009) Structural imperatives impose diverse evolutionary constraints on helical membrane proteins. *Proc. Natl. Acad. Sci. USA*, **106**, 17747–17750.
25 Killian, J.A. (1998) Hydrophobic mismatch between proteins and lipids in membranes. *Biochim. Biophys. Acta*, **1376**, 401–415.
26 Salom, D., Perez-Paya, E., Pascal, J., and Abad, C. (1998) Environment- and sequence-dependent modulation of the double-stranded to single-stranded conformational transition of gramicidin A in membranes. *Biochemistry*, **37**, 14279–14291.
27 Lee, A.G. (2003) Lipid–protein interactions in biological membranes: a structural perspective. *Biochim. Biophy. Acta*, **1612**, 1–40.
28 Orzaez, M., Lukovic, D., Abad, C., Perez-Paya, E., and Mingarro, I. (2005) Influence of hydrophobic matching on association of model transmembrane

fragments containing a minimised glycophorin A dimerisation motif. *FEBS Lett.*, **579**, 1633–1638.
29 Johansson, A.C. and Lindahl, E. (2006) Amino-acid solvation structure in transmembrane helices from molecular dynamics simulations. *Biophys. J.*, **91**, 4450–4463.
30 Dorairaj, S. and Allen, T.W. (2007) On the thermodynamic stability of a charged arginine side chain in a transmembrane helix. *Proc. Natl. Acad. Sci. USA*, **104**, 4943–4948.
31 Freites, J.A., Tobias, D.J., von Heijne, G., and White, S.H. (2005) Interface connections of a transmembrane voltage sensor. *Proc. Natl. Acad. Sci. USA*, **102**, 15059–15064.
32 Krepkiy, D., Mihailescu, M., Freites, J.A., Schow, E.V., Worcester, D.L., Gawrisch, K., Tobias, D.J., White, S.H., and Swartz, K.J. (2009) Structure and hydration of membranes embedded with voltage-sensing domains. *Nature*, **462**, 473–479.
33 White, S.H., Ladokhin, A.S., Jayasinghe, S., and Hristova, K. (2001) How membranes shape protein structure. *J. Biol. Chem.*, **276**, 32395–32398.
34 Liao, M.J., London, E., and Khorana, H.G. (1983) Regeneration of the native bacteriorhodopsin structure from two chymotryptic fragments. *J. Biol. Chem.*, **258**, 9949–9955.
35 Kahn, T.W., and Engelman, D.M. (1992) Bacteriorhodopsin can be refolded from two independently stable transmembrane helices and the complementary five-helix fragment. *Biochemistry*, **31**, 6144–6151.
36 Rath, A., Tulumello, D.V., and Deber, C.M. (2009) Peptide models of membrane protein folding. *Biochemistry*, **48**, 3036–3045.
37 Bordag, N. and Keller, S. (2010) Alpha-helical transmembrane peptides: a "divide and conquer" approach to membrane proteins. *Chem. Phys. Lipids*, **163**, 1–26.
38 Stouffer, A.L., Acharya, R., Salom, D., Levine, A.S., Di Costanzo, L., Soto, C.S., Tereshko, V., Nanda, V., Stayrook, S., and DeGrado, W.F. (2008) Structural basis for the function and inhibition of an influenza virus proton channel. *Nature*, **451**, 596–599.
39 Gratkowski, H., Lear, J.D., and DeGrado, W.F. (2001) Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. USA*, **98**, 880–885.
40 Zhou, F.X., Merianos, H.J., Brunger, A.T., and Engelman, D.M. (2001) Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci. USA*, **98**, 2250–2255.
41 Lear, J.D., Gratkowski, H., Adamian, L., Liang, J., and DeGrado, W.F. (2003) Position-dependence of stabilizing polar interactions of asparagine in transmembrane helical bundles. *Biochemistry*, **42**, 6400–6407.
42 Lemmon, M.A., Treutlein, H.R., Adams, P.D., Brunger, A.T., and Engelman, D.M. (1994) A dimerization motif for transmembrane alpha-helices. *Nat. Struct. Biol.*, **1**, 157–163.
43 Kleiger, G., Grothe, R., Mallick, P., and Eisenberg, D. (2002) GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry*, **41**, 5990–5997.
44 Senes, A., Gerstein, M., and Engelman, D.M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.*, **296**, 921–936.
45 Walters, R.F. and DeGrado, W.F. (2006) Helix-packing motifs in membrane proteins. *Proc. Natl. Acad. Sci. USA*, **103**, 13658–13663.
46 Oberai, A., Ihm, Y., Kim, S., and Bowie, J.U. (2006) A limited universe of membrane protein families and folds. *Protein Sci.*, **15**, 1723–1734.
47 Jastrzebska, B., Maeda, T., Zhu, L., Fotiadis, D., Filipek, S., Engel, A., Stenkamp, R.E., and Palczewski, K. (2004) Functional characterization of rhodopsin monomers and dimers in detergents. *J. Biol. Chem.*, **279**, 54663–54675.
48 Whorton, M.R., Jastrzebska, B., Park, P.S., Fotiadis, D., Engel, A., Palczewski, K., and Sunahara, R.K. (2008) Efficient

coupling of transducin to monomeric rhodopsin in a phospholipid bilayer. *J. Biol. Chem.*, **283**, 4387–4394.

49 Wu, Z., Gogonea, V., Lee, X., Wagner, M.A., Li, X.M., Huang, Y., Undurti, A., May, R.P., Haertlein, M., Moulin, M., Gutsche, I., Zaccai, G., Didonato, J.A., and Hazen, S.L. (2009) Double superhelix model of high density lipoprotein. *J. Biol. Chem.*, **284**, 36605–36619.

50 Aitken, A. (1996) Protein chemistry methods, post-translational modification, consensus sequences, in *Proteins LabFax* (ed. N.C. Price), BIOS/Academic Press, Oxford, pp. 253–285.

51 Salom, D., Lodowski, D.T., Stenkamp, R.E., Le Trong, I., Golczak, M., Jastrzebska, B., Harris, T., Ballesteros, J.A., and Palczewski, K. (2006) Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl. Acad. Sci. USA*, **103**, 16123–16128.

52 Tam, B.M. and Moritz, O.L. (2009) The role of rhodopsin glycosylation in protein folding, trafficking, and light-sensitive retinal degeneration. *J. Neurosci.*, **29**, 15145–15154.

53 Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002) Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous *N*-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl. Acad. Sci. USA*, **99**, 13419–13424.

54 Li, N., Salom, D., Zhang, L., Harris, T., Ballesteros, J.A., Golczak, M., Jastrzebska, B., Palczewski, K., Kurahara, C., Juan, T., Jordan, S., and Salon, J.A. (2007) Heterologous expression of the adenosine A1 receptor in transgenic mouse retina. *Biochemistry*, **46**, 8350–8359.

55 Salom, D., Wu, N., Sun, W., Dong, Z., Palczewski, K., Jordan, S., and Salon, J.A. (2008) Heterologous expression and purification of the serotonin type 4 receptor from transgenic mouse retina. *Biochemistry*, **47**, 13296–13307.

56 Mollaaghababa, R., Davidson, F.F., Kaiser, C., and Khorana, H.G. (1996) Structure and function in rhodopsin: expression of functional mammalian opsin in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **93**, 11482–11486.

57 Reeves, P.J., Thurmond, R.L., and Khorana, H.G. (1996) Structure and function in rhodopsin: high level expression of a synthetic bovine opsin gene and its mutants in stable mammalian cell lines. *Proc. Natl. Acad. Sci. USA*, **93**, 11487–11492.

58 Hargrave, P.A. (1977) The amino-terminal tryptic peptide of bovine rhodopsin. A glycopeptide containing two sites of oligosaccharide attachment. *Biochim. Biophys. Acta*, **492**, 83–94.

59 Chini, B. and Parenti, M. (2009) G-protein-coupled receptors, cholesterol and palmitoylation: facts about fats. *J. Mol. Endocrinol.*, **42**, 371–379.

60 Charollais, J. and Van Der Goot, F.G. (2009) Palmitoylation of membrane proteins (Review). *Mol. Membr. Biol.*, **26**, 55–66.

61 Park, P.S., Sapra, K.T., Jastrzebska, B., Maeda, T., Maeda, A., Pulawski, W., Kono, M., Lem, J., Crouch, R.K., Filipek, S., Muller, D.J., and Palczewski, K. (2009) Modulation of molecular interactions and function by rhodopsin palmitylation. *Biochemistry*, **48**, 4294–4304.

62 Zhang, L., Salom, D., He, J., Okun, A., Ballesteros, J., Palczewski, K., and Li, N. (2005) Expression of functional G protein-coupled receptors in photoreceptors of transgenic *Xenopus laevis*. *Biochemistry*, **44**, 14509–14518.

63 Mancia, F., Patel, S.D., Rajala, M.W., Scherer, P.E., Nemes, A., Schieren, I., Hendrickson, W.A., and Shapiro, L. (2004) Optimization of protein production in mammalian cells with a coexpressed fluorescent marker. *Structure*, **12**, 1355–1360.

64 Kawate, T. and Gouaux, E. (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure*, **14**, 673–681.

65 Guan, X.M., Kobilka, T.S., and Kobilka, B.K. (1992) Enhancement of membrane insertion and function in a type IIIb membrane protein following

introduction of a cleavable signal peptide. *J. Biol. Chem.*, **267**, 21995–21998.
66 Hague, C., Chen, Z., Pupo, A.S., Schulte, N.A., Toews, M.L., and Minneman, K.P. (2004) The N terminus of the human alpha1D-adrenergic receptor prevents cell surface expression. *J. Pharmacol. Exp. Ther.*, **309**, 388–397.
67 O'Malley, M.A., Mancini, J.D., Young, C.L., McCusker, E.C., Raden, D., and Robinson, A.S. (2009) Progress toward heterologous expression of active G-protein-coupled receptors in *Saccharomyces cerevisiae*: linking cellular stress response with translocation and trafficking. *Protein Sci.*, **18**, 2356–2370.
68 Gonzales, E.B., Kawate, T., and Gouaux, E. (2009) Pore architecture and ion sites in acid-sensing ion channels and P2X receptors. *Nature*, **460**, 599–604.
69 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., and Stevens, R.C. (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*, **318**, 1258–1265.
70 Jaakola, V.P., Griffith, M.T., Hanson, M.A., Cherezov, V., Chien, E.Y., Lane, J.R., Ijzerman, A.P., and Stevens, R.C. (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science*, **322**, 1211–1217.
71 Warne, T., Serrano-Vega, M.J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G., Tate, C.G., and Schertler, G.F. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*, **454**, 486–491.
72 Jastrzebska, B., Goc, A., Golczak, M., and Palczewski, K. (2009) Phospholipids are needed for the proper formation, stability, and function of the photoactivated rhodopsin–transducin complex. *Biochemistry*, **48**, 5159–5170.
73 Pebay-Peyroula, E., Garavito, R.M., Rosenbusch, J.P., Zulauf, M., and Timmins, P.A. (1995) Detergent structure in tetragonal crystals of OmpF porin. *Structure*, **3**, 1051–1059.
74 Zhang, H., Kurisu, G., Smith, J.L., and Cramer, W.A. (2003) A defined protein–detergent–lipid complex for crystallization of integral membrane proteins: the cytochrome b_6f complex of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. USA*, **100**, 5160–5163.
75 Jidenko, M., Nielsen, R.C., Sorensen, T.L., Moller, J.V., Maire, M., Nissen, P., and Jaxel, C. (2005) Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **102**, 11687–11691.
76 Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. *Science*, **289**, 739–745.
77 Yoshimura, K. and Kouyama, T. (2008) Structural role of bacterioruberin in the trimeric structure of archaerhodopsin-2. *J. Mol. Biol.*, **375**, 1267–1281.
78 Ruprecht, J.J., Mielke, T., Vogel, R., Villa, C., and Schertler, G.F. (2004) Electron crystallography reveals the structure of metarhodopsin I. *EMBO J.*, **23**, 3609–3620.
79 Hanson, M.A., Cherezov, V., Griffith, M.T., Roth, C.B., Jaakola, V.P., Chien, E.Y., Velasquez, J., Kuhn, P., and Stevens, R.C. (2008) A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. *Structure*, **16**, 897–905.
80 Lemieux, M.J., Song, J., Kim, M.J., Huang, Y., Villa, A., Auer, M., Li, X.D., and Wang, D.N. (2003) Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily. *Protein Sci.*, **12**, 2748–2756.
81 Guan, L., Smirnova, I.N., Verner, G., Nagamori, S., and Kaback, H.R. (2006) Manipulating phospholipids for crystallization of a membrane transport protein. *Proc. Natl. Acad. Sci. USA*, **103**, 1723–1726.
82 Guan, L., Mirza, O., Verner, G., Iwata, S., and Kaback, H.R. (2007) Structural

determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. USA*, **104**, 15294–15298.
83 Qin, L., Hiser, C., Mulichak, A., Garavito, R.M., and Ferguson-Miller, S. (2006) Identification of conserved lipid/detergent-binding sites in a high-resolution structure of the membrane protein cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. USA*, **103**, 16117–16122.
84 Gonen, T., Cheng, Y., Sliz, P., Hiroaki, Y., Fujiyoshi, Y., Harrison, S.C., and Walz, T. (2005) Lipid–protein interactions in double-layered two-dimensional AQP0 crystals. *Nature*, **438**, 633–638.
85 Seddon, A.M., Curnow, P., and Booth, P.J. (2004) Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta*, **1666**, 105–117.
86 Hirai, T., Subramaniam, S., and Lanyi, J.K. (2009) Structural snapshots of conformational changes in a seven-helix membrane protein: lessons from bacteriorhodopsin. *Curr. Opin. Struct. Biol.*, **19**, 433–439.
87 Angel, T.E., Chance, M.R., and Palczewski, K. (2009) Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proc. Natl. Acad. Sci. USA*, **106**, 8555–8560.
88 Angel, T.E., Gupta, S., Jastrzebska, B., Palczewski, K., and Chance, M.R. (2009) Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. *Proc. Natl. Acad. Sci. USA*, **106**, 14367–14372.
89 Levin, E.J., Quick, M., and Zhou, M. (2009) Crystal structure of a bacterial homologue of the kidney urea transporter. *Nature*, **462**, 757–761.
90 Kors, C.A., Wallace, E., Davies, D.R., Li, L., Laible, P.D., and Nollert, P. (2009) Effects of impurities on membrane-protein crystallization in different systems. *Acta Crystallogr. D*, **65**, 1062–1073.
91 Burkhart, B.M., Gassman, R.M., Langs, D.A., Pangborn, W.A., Duax, W.L., and Pletnev, V. (1999) Gramicidin D conformation, dynamics and membrane ion transport. *Biopolymers*, **51**, 129–144.
92 Newstead, S., Ferrandon, S., and Iwata, S. (2008) Rationalizing alpha-helical membrane protein crystallization. *Protein Sci.*, **17**, 466–472.
93 Newstead, S., Hobbs, J., Jordan, D., Carpenter, E.P., and Iwata, S. (2008) Insights into outer membrane protein crystallization. *Mol. Membr. Biol.*, **25**, 631–638.
94 Raman, P., Cherezov, V., and Caffrey, M. (2006) The Membrane Protein Data Bank. *Cell. Mol. Life Sci.*, **63**, 36–51.
95 Muller, D.J., Wu, N., and Palczewski, K. (2008) Vertebrate membrane proteins: structure, function, and insights from biophysical approaches. *Pharmacol. Rev.*, **60**, 43–78.
96 Palczewski, K. (2006) G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.*, **75**, 743–767.
97 Mirzadegan, T., Benko, G., Filipek, S., and Palczewski, K. (2003) Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry*, **42**, 2759–2767.
98 Lundstrom, K. (2006) Latest development in drug discovery on G protein-coupled receptors. *Curr. Protein Pept. Sci.*, **7**, 465–470.
99 Fukuda, M.N., Papermaster, D.S., and Hargrave, P.A. (1979) Rhodopsin carbohydrate. Structure of small oligosaccharides attached at two sites near the NH_2 terminus. *J. Biol. Chem.*, **254**, 8201–8207.
100 Ovchinnikov Yu, A., Abdulaev, N.G., and Bogachuk, A.S. (1988) Two adjacent cysteine residues in the C-terminal cytoplasmic fragment of bovine rhodopsin are palmitylated. *FEBS Lett.*, **230**, 1–5.
101 Ohguro, H., Palczewski, K., Ericsson, L.H., Walsh, K.A., and Johnson, R.S. (1993) Sequential phosphorylation of rhodopsin at multiple sites. *Biochemistry*, **32**, 5718–5724.
102 Fotiadis, D., Liang, Y., Filipek, S., Saperstein, D.A., Engel, A., and Palczewski, K. (2003) Atomic-force microscopy: rhodopsin dimers in native disc membranes. *Nature*, **421**, 127–128.
103 Papermaster, D.S. (1982) Preparation of retinal rod outer segments. *Methods Enzymol.*, **81**, 48–52.

104 Okada, T., Le Trong, I., Fox, B.A., Behnke, C.A., Stenkamp, R.E., and Palczewski, K. (2000) X-ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. *J. Struct. Biol.*, **130**, 73–80.

105 Okada, T., Sugihara, M., Bondar, A.N., Elstner, M., Entel, P., and Buss, V. (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *J. Mol. Biol.*, **342**, 571–583.

106 Li, J., Edwards, P.C., Burghammer, M., Villa, C., and Schertler, G.F. (2004) Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.*, **343**, 1409–1438.

107 Edwards, P.C., Li, J., Burghammer, M., McDowell, J.H., Villa, C., Hargrave, P.A., and Schertler, G.F. (2004) Crystals of native and modified bovine rhodopsins and their heavy atom derivatives. *J. Mol. Biol.*, **343**, 1439–1450.

108 Standfuss, J., Xie, G., Edwards, P.C., Burghammer, M., Oprian, D.D., and Schertler, G.F. (2007) Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.*, **372**, 1179–1188.

109 Stenkamp, R.E. (2008) Alternative models for two crystal structures of bovine rhodopsin. *Acta Crystallogr. D*, **D64**, 902–904.

110 MacKenzie, D., Arendt, A., Hargrave, P., McDowell, J.H., and Molday, R.S. (1984) Localization of binding sites for carboxyl terminal specific anti-rhodopsin monoclonal antibodies using synthetic peptides. *Biochemistry*, **23**, 6544–6549.

111 Salom, D., Le Trong, I., Pohl, E., Ballesteros, J.A., Stenkamp, R.E., Palczewski, K., and Lodowski, D.T. (2006) Improvements in G protein-coupled receptor purification yield light stable rhodopsin crystals. *J. Struct. Biol.*, **156**, 497–504.

112 McBee, J.K., Kuksa, V., Alvarez, R., de Lera, A.R., Prezhdo, O., Haeseleer, F., Sokal, I., and Palczewski, K. (2000) Isomerization of all-*trans*-retinol to *cis*-retinols in bovine retinal pigment epithelial cells: dependence on the specificity of retinoid-binding proteins. *Biochemistry*, **39**, 11370–11380.

113 Travis, G.H., Golczak, M., Moise, A.R., and Palczewski, K. (2007) Diseases caused by defects in the visual cycle: retinoids as potential therapeutic agents. *Annu. Rev. Pharmacol. Toxicol.*, **47**, 469–512.

114 Kiser, P.D., Lodowski, D.T., and Palczewski, K. (2007) Purification, crystallization and structure determination of native GroEL from *Escherichia coli* lacking bound potassium ions. *Acta Crystallogr. F*, **63**, 457–461.

115 Kiser, P.D., Golczak, M., Lodowski, D.T., Chance, M.R., and Palczewski, K. (2009) Crystal structure of native RPE65, the retinoid isomerase of the visual cycle. *Proc. Natl. Acad. Sci. USA*, **106**, 17325–17330.

116 Kochendoerfer, G.G., Salom, D., Lear, J.D., Wilk-Orescan, R., Kent, S.B., and DeGrado, W.F. (1999) Total chemical synthesis of the integral membrane protein influenza A virus M2: role of its C-terminal domain in tetramer assembly. *Biochemistry*, **38**, 11905–11913.

117 Ma, C., Polishchuk, A.L., Ohigashi, Y., Stouffer, A.L., Schon, A., Magavern, E., Jing, X., Lear, J.D., Freire, E., Lamb, R.A., DeGrado, W.F., and Pinto, L.H. (2009) Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. *Proc. Natl. Acad. Sci. USA*, **106**, 12283–12288.

118 Salom, D., Hill, B.R., Lear, J.D., and DeGrado, W.F. (2000) pH-dependent tetramerization and amantadine binding of the transmembrane helix of M2 from the influenza A virus. *Biochemistry*, **39**, 14160–14170.