Integral membrane protein structure: transmembrane α-helices as autonomous folding domains

Jean-Luc Popot

Institut de Biologie Physico-Chimique, Paris, France

The transmembrane region of many integral membrane proteins is made up of a bundle of hydrophobic α-helices. Such a structure could result from a two-stage folding process, during which preformed transmembrane helices with independent stability pack without topological rearrangement. This view was originally prompted by experiments in which fragments of transmembrane regions were separately refolded into lipid bilayers and subsequently brought together to yield a functional protein. Other lines of evidence, including the existence of 'one-helix' miniproteins, gene-fusion experiments, helix-driven oligomerization of bitopic proteins, and sequence rearrangements in the course of evolution support this view. Although it forms a useful basis for structural predictions, the limitations of the two-stage folding hypothesis are not clearly defined, and the proportion of integral membrane proteins to which it applies remains uncertain. The papers discussed in the present review illustrate recent progress along these lines.

Introduction

The way in which integral membrane proteins (IMPs) achieve their three-dimensional (3D) structure largely remains a matter of conjecture. There are a number of reasons for this, the most prominent of which are the extreme scarcity of high-resolution IMP structures and the great technical difficulties involved in obtaining direct structural information on the early steps of IMP folding. Because structurally well characterized IMPs are rare, it is both tempting and dangerous to try to extrapolate from these few data general rules which might help in understanding the structure and behaviour of other IMPs. One such attempt at generalization is the two-stage folding model [1,2]. This model postulates that the transmembrane region of many IMPs results from the packing of preformed, hydrophobic helices, each of which represents a small autonomous folding domain (see below).

The two-stage folding hypothesis has many implications (see Fig. 1) and has proven to be a useful tool for interpreting a number of experimental observations relating to, for example, IMP biosynthesis and oligomerization [3,4•,5,6•] and IMP evolution [6•]. It also raises numerous questions, in particular regarding its generality. Most IMPs whose sequences are known contain at least one stretch of amino acid residues with sufficient overall hydrophobicity to form a stable transmembrane helix. However, neither the number of these segments that do

form transmembrane helices nor the proportion of transmembrane regions that are pure α-helix bundles (rather than incorporating other structural elements) has been established. Similarly, many experiments have shown that some IMPs can be synthesized as fragments that insert independently into the membrane and subsequently assemble into a functional protein. It is not clear however to what extent insertion of the constitutive helices occurs independently *in vivo*.

This review discusses recent results that have a bearing on these questions. Some support the concept of two-stage folding, whereas others question its generality. The choice of articles has been guided mainly by the wish to illustrate the progress that has been made on various aspects of the problem. Space constraints mean this review cannot be exhaustive. For this reason, some relevant issues, such as model building and the formation of transmembrane channels by synthetic peptides or by toxins, have had to be largely, if not totally, omitted.

Transmembrane α-helices as autonomous folding domains

It is useful first to briefly recall the main arguments on which the concept of two-stage folding is founded

Abbreviations

AChR—acetylcholine receptor; BR—bacteriorhodopsin; 3D—three-dimensional; EM—electron microscopy; IMP—integral membrane protein; PLP—major proteolipid of central nervous system myelin; RC—reaction centre.
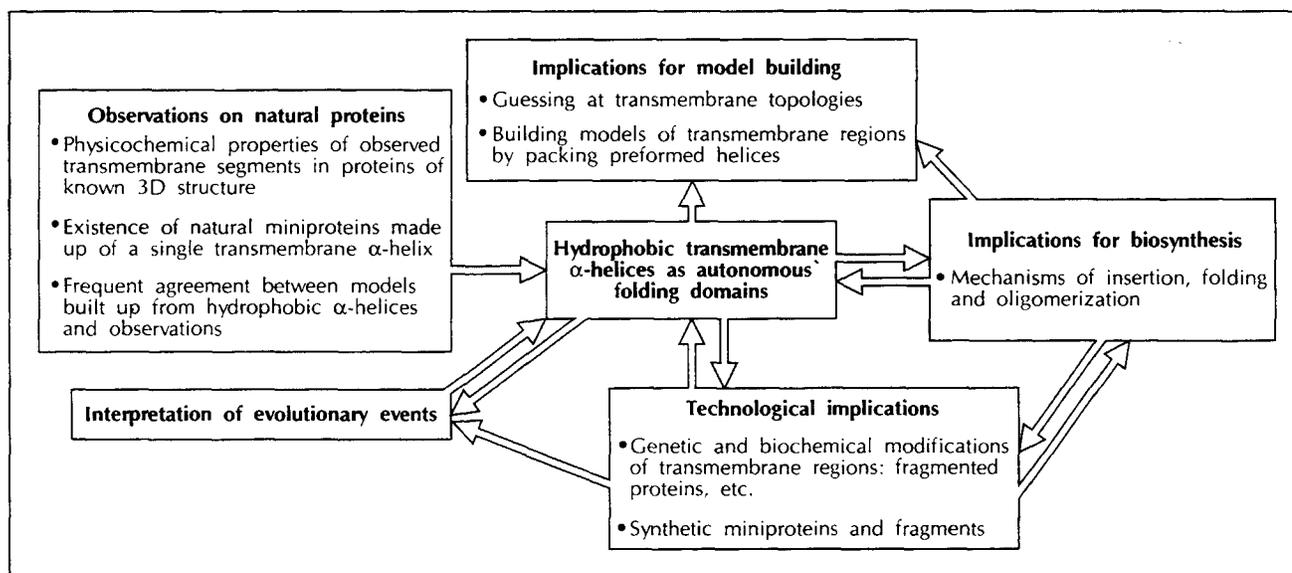
Fig. 1. A schematic view of the relationships between the concept of transmembrane α-helices behaving as autonomous folding domains and some related issues. Arrows symbolize the impact of one field upon the others. The main arguments in favour of the folding autonomy of transmembrane helices are summarized in Table 1.

and which underlie the whole discussion (Table 1). Hydrophobic environments induce polypeptides to take up periodic secondary structures, a process that is driven by the formation of main-chain hydrogen bonds [7]. For a short peptide (20–30 residues long) spanning the core of a lipid bilayer, this imposes the formation of an α- (or 3_{10}-) helix. The few plasma membrane proteins whose transmembrane secondary structure is known with certainty are made up from bundles of α-helices [6•]. Closed barrels, composed of β-sheets rolled upon themselves, can also satisfy main-chain hydrogen bonding in the membrane interior. All bacterial porins whose 3D structures are known are folded into β-barrels, the lumens of which form aqueous pores (reviewed in this issue, pp 501–507). Two-stage folding does not apply to β-barrels. Insertion of isolated β-strands (or pairs of β-strands) into lipids would be energetically very costly, because of the large number of unsatisfied hydrogen bonds. In keeping with this expectation, it seems that porins fold before, or during, membrane insertion [6•,8].

In contrast, provided its surface is sufficiently hydrophobic, an isolated α-helix spanning the membrane can be expected to be stable by itself. Estimates of transfer free energy indicate that each individual transmembrane α-helix in known structures [bacteriorhodopsin (BR) and bacterial reaction centres (RCs)], would be more stable when spanning the membrane than in the aqueous phase. Indeed, BR fragments can be inserted into lipid vesicles, where they form transmembrane α-helices, and can subsequently be reassembled into functional BR molecules (see below). This observation led to the proposal that the native structure of BR and similar proteins results from a two-stage folding process: in the first stage, hydrophobic segments insert into the membrane as α-helices; in the second stage, they pack without major rearrangement into an α-helix bundle [1,2]. In such a model, individual α-helices are considered as autonomous folding domains, able to achieve a defined structure in the

absence of the rest of the protein, even though they lack the structural complexity of the larger folding domains found in soluble proteins. The principal arguments in favour of the two-stage model are summarized in Table 1, together with some caveats.

New high-resolution and medium-resolution structures

New structures are the touchstone of ideas about IMP folding. The past year has seen some very interesting progress along familiar lines, as well as unexpected results which may require a revision of our ideas about IMP folding. The porin/β-barrel family has two new members whose structures have been established at atomic resolution by X-ray crystallography [9••]. Both have 16 transmembrane β-strands. The family of proteins folded into a transmembrane α-helix bundle now includes the photosystem I RC and rhodopsin. New, very unexpected medium-resolution data on the nicotinic acetylcholine receptor (AChR) are discussed in a later section.

A 3D electron-density map of photosystem I RC has been established by X-ray crystallography at 6Å resolution [10••]. Of the 28 expected transmembrane helices, 21 are resolved, as are most of the prosthetic groups involved in electron transfer and half of the 90-odd chlorophylls. Part of the structure shows approximate twofold symmetry and the arrangement of some of the chlorophylls is reminiscent of that in bacterial RCs, with which PSI is thought to share a distant ancestry (see this issue, pp 508–514).

The projection structure of rhodopsin has been established at 9Å resolution by high-resolution electron microscopy (EM) [11••]. It is compatible with the expected seven-helix topology. Four of the helices are nearly per-

Table 1. Principal arguments in favour of the folding autonomy of hydrophobic transmembrane α-helices.

Observation	Comments
Analysis of membrane-spanning segments in proteins of known 3D structure Lipid-facing segments in known structures strictly respect two constraints: a periodic secondary structure satisfying most main-chain hydrogen bonds; and the strongly hydrophobic character of lipid-facing surfaces.	Known structures represent only a very small subset of IMP structures and functions. Similar constraints apply to polypeptide segments buried within the transmembrane region of the protein, away from both lipids and water, but not to segments facing the aqueous phase inside hollow (pore-forming) proteins (e.g. the interior loop in porins). Sequence segments that are not by themselves hydrophobic, such as the strands making up a β-barrel, must be inserted in a concerted way.
Individual membrane-spanning helices are overall hydrophobic and predicted to partition more favourably into a lipid rather than an aqueous phase.	A transmembrane position may be just one among several possible arrangements for the isolated segment; e.g. an isolated helix may distribute between juxta- and trans-membrane positions. During biosynthesis, membrane insertion of two or more helices may be simultaneous (double-helical hairpin hypothesis).
One-helix proteins Organelle membranes contain many very short proteins (often 30–40 residues), predicted to each contain a single hydrophobic transmembrane α-helix.	In many cases, the transmembrane nature and even the existence of these proteins is still hypothetical.
Topologies by gene fusion Despite occasional ambiguities, topologies deduced from gene-fusion experiments generally are consistent with the assumption that all transmembrane segment are hydrophobic α-helices. The approach has been validated by the consistency of most of the observations with other sources of topological information, particularly in the case of the L subunit of *Rhodobacter sphaeroides* RC, where the actual topology is known from crystallographic studies.	By design, constructs in which the carboxy-terminal end of the protein under study is replaced with a reporter protein cannot detect some hypothetical structures, such as two successive transmembrane β-strands. Such a structure would not affect the distribution of other extramembrane loops on each side of the membrane. The constitutive strands of a β-barrel would not be expected to form stable transmembrane β-strands in the absence of their partners.
Fragmented proteins A host of experiments indicates that natural or synthetic fragments containing one or more putative transmembrane helices are able to insert into membranes and to recognize their partners within a protein or an oligomer.	One of the strongest arguments in favour of the idea of transmembrane helices behaving as autonomous folding domains.
Extramembrane loops can often be dispensed with.	The existence of loops nevertheless stabilizes the folded structure, even if it does not direct folding. Weaker helix–helix interactions are sufficient to ensure specific association, which may be a factor favouring structural plasticity.

pendicular to the plane of the membrane and the remaining helices are tilted to form a diffuse arc. One end of the arc intercalates between two of the well defined helices, yielding an overall structure that is only remotely reminiscent of BR. An analysis based on residue variability in > 200 sequences of G-protein coupled receptors and on constraints on the minimal length of the interhelix loops has permitted the putative transmembrane segments to be assigned to helices in the electron-density map [12•].

As in BR, the order of the helices would be the same as in the sequence, but their 3D arrangement would be somewhat different, with the third helix slanted into the middle of the bundle. These observations should lead to major improvements in current 3D models of G-protein coupled receptors [13,14].

As regards crystallographic studies of other IMPs whose structures are known to near atomic resolution, further evidence has been obtained in favour of a change of con-

formation of BR during proton pumping [15•]. The orientation of helices in resting-state BR has been further defined using neutron diffraction (FA Samatey *et al*, unpublished data) [16] and model building (P Tufféry *et al*, unpublished data) [17•]. In addition, extensive two-dimensional NMR studies have been made of the structure of various proteolytic fragments of BR in organic media or SDS micelles (see below).

Topologies

Smooth sailing for MalG and central nervous system myelin proteolipid

Mapping topologies by gene-fusion techniques has continued to yield results (see [18] and this issue, pp 524–531). Particularly relevant to the present discussion are studies of *lac* permease [19], the MalG protein [20•] and AChR [21•]. In all three cases, the results obtained are consistent with interpreting the hydrophobicity profile in terms of one transmembrane helix per ~20-residue hydrophobic segment. The approach has been validated further by applying it to subunit L of the bacterial RC [22], whose transmembrane topology is known.

Mapping transmembrane topologies by gene-fusion experiments has its limitations however. For instance, the most common method is to replace the carboxy-terminal part of the protein under study with the reporter protein. This approach cannot be expected to reveal the presence of β-strands hairpins (such as have been proposed in models of voltage-gated channel proteins [23–25]) because such features would be unstable in the membrane in the absence of the strands completing the putative β-barrel, as discussed above [20•].

Another proposed deviation from the pure α-helix bundle besides β-strands is the interrupted α-helix, which spans a single leaflet, looping back on itself in the middle of the bilayer. For similar reasons to those discussed above, such an arrangement is expected to be energetically very costly [26]. This consideration was the starting point for revising existing topological models for the major proteolipid of central nervous system myelin (PLP) [27]. The revised topology, which features four α-helices spanning the full thickness of the membrane, has received good support from an examination of the state of the cysteine residues in the proteolipid: the four cysteines predicted to lie outside the cell form disulphide bridges; the four intramembrane cysteines are reduced; and the six cytosolic cysteines are acylated [28•]. There are a number of other proteins for which the proposed topologies deviate from the standard bundle of hydrophobic α-helices, for example, the mitochondrial ADP/ATP exchanger [29], whose structural study might lead to interesting insights.

Topological predictions often remain ambiguous, even if it is assumed that long stretches of hydrophobic residues must correspond to transmembrane α-helices, because there is no well defined threshold separating transmembrane from extramembrane hydrophobic seg-

ments [3,30]. Two recent studies on the composition of short [31•,32] and long [33•] extramembrane loops in extracellular versus intracellular compartments may help distinguish between alternative topologies.

Rough seas for the acetylcholine receptor

Whereas the above observations are consistent with the topologies of *lac* permease, MalG (and MalF) or PLP resulting from two-stage folding, some recent observations for AChR do not conform. The nicotinic AChR, the prototypic mediator-gated ion channel, is a complex of five ($\alpha_2\beta\gamma\delta$) homologous subunits, symmetrically arranged around a central pore. The topology of these subunits was a source of lively polemics in the mid-1980s [34]. The outcome was a consensus, questioned by the occasional *franc-tireur*, in which each subunit features four hydrophobic transmembrane helices (M1–M4). At least one of these, helix M2, was shown by numerous biochemical and molecular genetics experiments to be involved in channel formation, leading to the conclusion that the channel was lined by five M2 helices [35,36]. The three other hydrophobic segments have been shown to be in contact with lipids [37•]. Two experimental investigations have yielded results consistent with a four-helix topology for the AChR subunit. In one, the topology was mapped by fusing the carboxy-terminal domain of prolactin downstream of each of the postulated transmembrane segments [21•], and in the other it was observed that the five residues in M4 that were labelled by lipophilic reagents are distributed with α-helical periodicity [37•].

Recent results however have challenged this four-helix topology in two divergent and mutually incompatible ways. One series of experiments relied on electrophysiological measurements and the effect of sulphydryl-blocking agents on AChR molecules in which residues in and around the M2 region had been substituted with cysteines. The results were consistent with M2 lining the channel, but were considered more suggestive of a β-strand than an α-helix [38•].

This conclusion is hard to reconcile with a recent EM study [39••] that yielded a 9 Å 3D map of the *Torpedo* AChR. The map showed five helices, presumably M2, surrounding the pore. Surprisingly, the 15 other expected helices were not apparent in the map — a fact for which the limited resolution of the map does not seem to be responsible. Instead, the author proposed that the five M2 helices are surrounded by a vast (unresolved) β-barrel to which each subunit contributes five or seven strands, yielding an overall structure similar to those of two (water-soluble) bacterial toxins [40,41].

Space limitations preclude a detailed discussion of this proposal. However, it has wide-ranging implications for the assembly of the AChR. For instance, individual subunits (which are synthesized independently) would be expected to have a different topology in the isolated state than when assembled in the complex. As already discussed, it is quite unlikely that unassembled β-strands of a transmembrane β-barrel could be forced to face the

lipids. As segments M1, M3 and M4 are long and highly hydrophobic, it cannot be excluded that, in a truncated protein, they do insert as α-helices, which would explain the gene-fusion results [21•]. It is very difficult however to envisage how three of the four transmembrane α-helices (were they to be present in the unassembled subunits) could subsequently rearrange into five to seven β-strands upon oligomerization. Alternative possibilities, such as a transient co-existence (in isolated subunits) of a transmembrane helix (M2) with a seven-stranded β-barrel (that would open to associate with those of other subunits), are almost too embarrassing to mention given the mechanistic problems they would create at the time of insertion and oligomerization. Further investigations into the topology of the AChR and its unassembled subunits, as well as the identification of hypothetical factors that may stabilize the transmembrane structure of unassembled subunits and/or facilitate structural rearrangements, will clearly have important implications for IMPs extending far beyond the AChR field.

As regards the practice of sequence interpretation, rough days are ahead of us if we are to determine whether a stretch of ~ 25 hydrophobic residues represents a single transmembrane α-helix or three closely spaced β-strands with no intervening hydrophilic residues. Each putative transmembrane segment will have to be examined in more detail than before for tell-tale periodicities in amino acid composition and variability. Improved scales for discriminating protein–lipid from protein–protein interfaces may help in meeting this challenge (FA Samatey, J-L Popot, unpublished data) [42•].

Structural autonomy of integral membrane protein fragments

Further evidence that topology-determining information is dispersed throughout the sequence of IMPs has come from several experiments. For instance, alkaline phosphatase has been fused to fragments containing single *lac* permease transmembrane segments. Phosphatase activity indicated that each individual segment is able to insert in the plasma membrane of *Escherichia coli* with its correct orientation [43•]. Similar observations have been made *in vitro* and *in vivo* on the three transmembrane segments of hepatitis M protein [44•].

The large degree of structural autonomy of IMP fragments is supported by experiments in which IMPs have been reassembled from fragments that have been separately refolded (*in vitro*) or synthesized (*in vivo*). The most recent example concerns the muscarinic AChRs m2 and m3, two members of the seven-helix, G-protein-coupled receptor family [45]. These sorts of experiments have been reviewed elsewhere [3,6•]. In the most enlightening study to date, it was shown that the first two transmembrane helices in BR can be replaced with synthetic substitutes; these individually form transmembrane α-helices when inserted into artificial lipid vesicles (JF Hunt *et al.*, abstract, *Biophys J* 1991, 59:400) [46••].

These and earlier studies have shown that denatured proteolytic or synthetic fragments of BR reconstituted in lipid–detergent mixed micelles [47] and in lipid vesicles (JF Hunt *et al.*, abstract, *Biophys J* 1991, 59:400) [1] reform α-helices. Detailed two-dimensional NMR studies have been undertaken on the structure of such fragments in organic media and in SDS micelles [48••,49••,50•]. The most striking findings are that, even in isotropic organic media, these fragments form α-helices over a length that corresponds closely to that in the native molecule, and, in addition, the conformation of most of the residues in the α-helical section of each peptide is constrained. These observations have interesting implications both regarding the nature of the factors that determine helix length and for model building.

Transmembrane helices can drive oligomerization

Although the single membrane-spanning helix of bitopic proteins has long been considered a mere hydrophobic anchor, it is increasingly often found to play a role in oligomerization. These findings, and their implications for transmembrane signalling, have been discussed in a number of recent reviews [4•,5,51]. In the case of the erythrocyte glycoprotein glycophorin A, for instance, it has been shown that the transmembrane anchor, even taken out of its normal protein context, is sufficient to cause dimerization. SDS-resistant heterodimers have been obtained, in which an authentic molecule of glycophorin associates with a synthetic peptide mimicking its anchor [52]. In more recent studies, gene fusion was used to graft the glycophorin anchor onto a soluble, monomeric protein. Dimer formation was shown to depend on the exact sequence of the anchor [53••] and its sensitivity to amino acid substitutions was used to map the helix–helix interface [54••]. The observed periodicity suggests that the two helices form a right-handed supercoil, rather than the more usual left-handed one. This conclusion is supported by simulated annealing calculations [55••].

The maltose import system of *E. coli* is a 1:1:2 complex of two integral and one cytosolic protein, MalF, MalG and MalK, respectively [56]. MalF has eight transmembrane helices and MalG has six. Recent work shows that MalG′, a two-helix fragment of MalG, stabilizes MalF against the degradation by endogenous proteases that is observed when MalF is expressed alone (CJ Panagiotidis *et al.*, personal communication). It is not known whether the MalF–MalG′ interaction is mediated simply by helix–helix contacts, because MalG′ (128 residues) still harbours a rather extended periplasmic loop. Because earlier work had shown that MalF expressed along with MalG′ could subsequently associate with full length MalG [57••], it follows that the MalF–MalG′ interaction is reversible.

Another interesting case of association with an incomplete IMP may be offered by subunit D1 of the photosystem II RC, a five-helix protein that binds several cofactors. The kinetics of D1 synthesis exhibit several

pauses (of uncertain origin). These have been postulated to facilitate co-translational binding of prosthetic groups to freshly inserted helices, thereby protecting the incomplete protein against co-translational proteolytic degradation [58].

Plasticity of α-helical transmembrane regions

It has been shown that deletion of several transmembrane helices from BR or *lac* permease molecules by biochemical or genetic means results in non-functional species. However, functional complementation is observed when two such molecules bearing different deletions are either reconstituted or expressed simultaneously, even though the complex thus formed has an excess of transmembrane segments [59,60••]. The implication is that the two incomplete bundles are destabilized and hence can engage in intermolecular associations. They recruit missing helices from a complementary molecule, forming abnormal, but functional, heterodimers. Redundant helices in the complex presumably remain transmembrane, but are expelled to the side of a native-like, hybrid bundle [6•]. An example of redundant helices being inserted within an unfragmented IMP has also been described [61].

Elongation, shortening, fragmentation or fusion of transmembrane regions are likely to have been common events in the course of evolution [6•]. In the cytochrome *c* oxidase family, for instance, genes encoding proteins homologous to mitochondrial subunits I and III are fused into a single gene in one of the oxidases of *Sulfolobus acidocaldarius* (M Lübben *et al*, personal communication). In several other bacteria, subunits I and III are separate, but two of the transmembrane helices appear to have transferred from one subunit to the other, compared with the mitochondrial subunits [62]. It has been suggested recently that BR and G-protein coupled receptors have evolved from a common ancestor by shuffling blocks of transmembrane helices [63], but the statistical evidence presented is not entirely convincing.

Conclusion

Over the past few years, it has become increasingly obvious that IMPs whose transmembrane region is made up of a bundle of hydrophobic α-helices behave as two-tiered structures, with each helix exhibiting a large degree of structural autonomy. Application of this concept permits rationalization of some otherwise surprising observations in such areas as membrane integration during biosynthesis, oligomerization and evolution.

At the same time, it is still uncertain what proportion of IMPs belongs in the two-stage folding category. Determination of atomic structures has made it clear that membrane proteins strictly obey two expected constraints: firstly, polypeptide segments in contact with

lipids adopt periodic secondary structures which satisfy internal hydrogen-bonding requirements; and secondly, they expose predominantly hydrophobic side chains at the lipid–protein interface. The diversity of folding schemes allowed to IMPs within these constraints remains one of the most interesting open questions.

Progress is being made continuously in the development of high-resolution EM and in growing two-dimensional and 3D IMP crystals. Our database of atomic structures will inevitably increase, and, with a wider knowledge of structural diversity, the risks inherent in trying to identify general folding principles will diminish. These general rules will remain indispensable for interpreting data gathered on proteins whose high-resolution structure remains beyond our reach.

Acknowledgements

References and recommended reading

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

•   of special interest
••   of outstanding interest

1.   POPOT J-L, GERCHMAN S-E, ENGELMAN DM: Refolding of Bacteriorhodopsin in Lipid Bilayers: a Thermodynamically Controlled Two-Stage Process. *J Mol Biol* 1987, 198:655–676.

2.   POPOT J-L, ENGELMAN DM: Membrane Protein Folding and Oligomerization: the Two-Stage Model. *Biochemistry* 1990, 29:4031–4037.

3.   POPOT J-L, DE VITRY C: On the Microassembly of Integral Membrane Proteins. *Annu Rev Biophys Biophys Chem* 1990, 19:369–403.

4.   BORMANN BJ, ENGELMAN DM: Intramembrane Helix–Helix Association in Oligomerization and Transmembrane Signaling. *Annu Rev Biophys Biomol Str* 1992, 21:223–242.
An important review of the role of helix–helix interactions in driving oligomerization of bitopic (single-spanning) proteins and in transmembrane signalling.

5.   LEMMON MA, ENGELMAN DM: Helix–Helix Interactions Inside Lipid Bilayers. *Curr Opin Struct Biol* 1992, 2:511–518.

6.   POPOT J-L, DE VITRY C, ATTEIA A: Folding and Assembly of Integral Membrane Proteins: an Introduction. In *Membrane Protein Structure: Experimental Approaches*. Edited by White SH. Oxford: Oxford University Press; 1993: in press.
A discussion of data on the folding and assembly of IMPs in various cell compartments.

7.   SINGER SJ: The Properties of Proteins in Nonaqueous Solvents. *Adv Protein Chem* 1962, 17:1–68.

8.   NIKAIDO H: Porins and Specific Channels of Bacterial Outer Membranes. *Mol Microbiol* 1992, 6:435–442.

9.   COWAN SW, SCHIRMER T, RUMMEL G, STEIERT M, GHOSH R, PAUPTIT RA, JANSONIUS JN, ROSENBUSCH JP: Crystal Structures Explain Functional Properties of Two *E. coli* Porins. *Nature* 1992, 358:727–733.

The crystal structures of two *E. coli* porins, PhoE and OmpF, reveal interesting variations on the theme of the 16-strand transmembrane β-barrel.

10. KRAUSS N, HINRICHS W, WITT I, FROMME P, PRITZKOW W,
•• DAUTER Z, BETZEL C, WILSON KS, WITT HT, SAENGER W: Three-Dimensional Structure of System I of Photosynthesis at 6Å Resolution. *Nature* 1993, 361:326–331.

The newest member in the exclusive club of IMPs whose structure is known at a resolution sufficient to visualize elements of transmembrane secondary structure. It is a large (340 kDa) complex of 11 subunits, seven of which are integral. Only 21 of the expected 28 transmembrane helices have been identified. It is considered unlikely that further transmembrane helices will be revealed when the map improves (N Krauss, personal communication).

11. SCHERTLER GFX, VILLA C, HENDERSON R: Projection Structure
•• of Rhodopsin. *Nature* 1993, 362:770–772.

Purified bovine rhodopsin has been reconstituted into lipid bilayers at low lipid:protein ratios and samples screened for the presence of two-dimensional crystals. Images of frozen hydrated crystals, recorded at liquid nitrogen temperature using low-dose EM, are processed to calculate a 9Å projection electron-density map (see text).

12. BALDWIN JM: The Probable Arrangement of the Helices in
• G Protein-Coupled Receptors. *EMBO J* 1993, 12:1693–1703.

Sequence comparisons are used to predict the rotational orientation of each of the G-protein seven transmembrane α-helices with respect to the inside and outside of the helix bundle. Based on the probable degree of lipid-exposure of each helix (and each helix end), helices are postulated to be arranged in a flattened cylinder, with one end of helix III wedged into the bundle on the cytosolic side of the membrane.

13. MALONEYHUSS T, LYBRAND TP: Three-Dimensional Structure of the β2 Adrenergic Receptor Protein Based on Computer Modeling Studies. *J Mol Biol* 1992, 225:859–871.

14 TRUMPP-KALLMEYER S, HOFLACK J, BRUINVELS A, HIBERT M: Modelling of G Protein Coupled-Receptors. Application to Dopamine, Adrenaline, Serotonin, Acetylcholine and Mammalian Opsin Receptors. *J Med Chem* 1993, 35:3448–3462.

15. SUBRAMANIAM S, GERSTEIN M, OESTERHELT D, HENDERSON R:
• Electron Diffraction Analysis of Structural Changes in the Photocycle of Bacteriorhodopsin. *EMBO J* 1993, 12:1–8.

Structural changes during the photocycle of BR have been studied by EM. The M intermediate of either wild-type BR or the Asp96Gly mutant is trapped by rapid freezing 20 ms following actinic illumination. Difference Fourier projection maps at 3.5Å resolution are calculated from electron diffraction pattern intensities. The most prominent changes are observed in the vicinity of helices F and G, consistent with earlier, lower-resolution studies using neutron or X-ray powder patterns. Preliminary data from tilted grids indicate that changes are localized to the cytoplasmic half of the membrane.

16. SAMATEY FA, POPOT J-L, ETCHEBEST C, ZACCAI G: Rotational Orientation of Transmembrane α-Helices in Bacteriorhodopsin Studied by Neutron Diffraction. In *Proceedings of the 5th International Conference on Retinal Proteins.* Edited by Rigaud JL. Montronge: John Libbey Eurotext; 1992:9–12.

17. TUFFÉRY P, POPOT J-L, LAVERY R: Modelling Bundles of Trans-
• membrane Helices: a Test Study on Bacteriorhodopsin. In *Proceedings of the 5th International Conference on Retinal Proteins.* Edited by Rigaud JL. Montronge: John Libbey Eurotext; 1992:13–16.

The authors have examined whether the rotational arrangement of α-helices in BR can be predicted once their sequences and the position of their axes is known. Information about side-chain conformations and helix rotational positions (but not the shape of the backbone) has been erased from the EM model of BR. Models were generated by rotating helices (either in pairs or in triplets) around their axes and it was determined, for each model, which energy of interaction is associated with the most favourable combination of side-chain conformations. It is concluded that it is not possible, using this procedure, to identify uniquely how two given helices interact in BR, even though the correct arrangement always is among the best models; nevertheless, the rotational positions established by EM and neutron diffraction can be 'postdicted' to

18. BOYD D: The Use of Gene Fusions to Study Membrane Protein Topology. In *Membrane Protein Structure: Experimental Approaches.* Edited by White SH. Oxford: Oxford University Press; 1993; in press.

19. CALAMIA J, MANOIL C: *lac* Permease of *Escherichia coli:* Topology and Sequence Elements Promoting Membrane Insertion. *Proc Natl Acad Sci USA* 1990, 87:4937–4941.

20. BOYD D, TRAXLER B, BECKWITH J: Analysis of the Topology
• of a Membrane Protein by Using a Minimal Number of Alkaline Phosphatase Fusions. *J Bacteriol* 1993, 175:553–556.

Mapping of the topology of MalG by alkaline phosphatase fusions is consistent with the existence of six membrane-spanning segments.

21. CHAVEZ RA, HALL ZW: Expression of Fusion Proteins of the
• Nicotinic Acetylcholine Receptor from Mammalian Muscle Identifies the Membrane-Spanning Regions in the α and δ Subunits. *J Cell Biol* 1992, 116:385–393.

An investigation of the transmembrane topology of the α and δ subunits of the AChR using a gene-fusion technique in which a domain of prolactin is attached downstream of the various putative transmembrane segments. The results are consistent with the four-helix topology (see text).

22. YUN C-H, VAN DOREN SR, CROFTS AR, GENNIS RB: The Use of Gene Fusions to Examine the Membrane Topology of the L-Subunit of the Photosynthetic Reaction Center and of the Cytochrome *b* subunit of the *bc₁* Complex from *Rhodobacter sphaeroides. J Biol Chem* 1991, 266:10967–10973.

23. YOOL AJ, SCHWARZ TL: Alteration of Ionic Selectivity of a K⁺ Channel by Mutation of the H5 Region. *Nature* 1991, 349:700–704.

24. DURELL SR, GUY HR: Atomic Scale Structure and Functional Models of Voltage-Gated Potassium Channels. *Biophys J* 1992, 62:238–250.

25. BOGUSZ S, BOXER A, BUSATH DD: An SS1–SS2 β-Barrel Structure for the Voltage-Activated Potassium Channel. *Protein Eng* 1992, 4:285–293.

26. JACOBS RE, WHITE SH: The Nature of the Hydrophobic Binding of Small Peptides at the Bilayer Interface: Implications for the Insertion of Transbilayer Helices. *Biochemistry* 1989, 28:3421–3437.

27. POPOT J-L, PHAM-DINH D, DAUTIGNY A: Major Myelin Proteolipid: the 4-α-Helix Topology. *J Membr Biol* 1991, 120:233–246.

28. WEIMBS T, STOFFEL W: Proteolipid Protein (PLP) of CNS
• Myelin: Positions of Free, Disulfide-Bonded, and Fatty-Acid Thioester-Linked Cysteine Residues and Implications for the Membrane Topology of PLP. *Biochemistry* 1992, 31:12289–12296.

Four different topological models for PLP have been advocated over the years. This paper presents a thorough study of the state of the cysteinyl residues in PLP from bovine brain. The only model that correctly predicts that all cystine residues lie outside the cell and all acylated residues face the cytosol is that compatible with two-stage folding [27].

29. MARTY I, BRANDOLIN G, GAGNON J, BRASSEUR R, VIGNAIS PV: Topography of the Membrane-Bound ADP/ATP Carrier Assessed by Enzymatic Proteolysis. *Biochemistry* 1992, 31:4058–4065.

30. WHITE SH: Hydropathy Plots and the Prediction of Membrane Protein Topology. In *Membrane Protein Structure: Experimental Approaches.* Edited by White SH. Oxford: Oxford University Press; 1993: in press.

31. VON HEIJNE G: Membrane Protein Structure Prediction.
• Hydrophobicity Analysis and the Positive-Inside Rule. *J Mol Biol* 1992, 225:487–494.

A formalization and test of the use of the 'positive inside' rule to straighten ambiguous topologies. Out of 24 *E. coli* IMPs tested, the only one on which the analysis fails is... the pet protein of the author, signal peptidase.

32.   VON HEIJNE G: Decoding the Signals of Membrane Protein Sequences. In *Membrane Protein Structure: Experimental Approaches* Edited by White SH. Oxford: Oxford University Press; 1993: in press.

33. • NAKASHIMA H, NISHIKAWA K: The Amino Acid Composition Is Different between the Cytoplasmic and Extracellular Sides in Membrane Proteins. *FEBS Lett* 1992, 303:141–146.
The composition of large (> 50 residues) extramembrane segments in bitopic and polytopic proteins is examined. Strongest deviations from randomness are found for alanine and arginine residues (inside) and for threonine and cysteine residues (outside). The origin of most the differences is not clear, but they are similar to differences observed between cytosolic and extracellular proteins. These data can probably help when topologies are difficult to sort out.

34.   MCCREA PD, POPOT J-L, ENGELMAN DM: Transmembrane Topography of the Nicotinic Acetylcholine Receptor δ Subunit. *EMBO J* 1987, 6:3619–3626.

35.   CHANGEUX J-P, GALZI J-L, DEVILLERS-THIÉRY A, BERTRAND D: The Functional Architecture of the Acetylcholine Nicotinic Receptor Explored by Affinity Labelling and Site-Directed Mutagenesis. *Q Rev Biophys* 1992, 25:395–432.

36.   CAFISO D: Experimental Determination of the Topology of Membrane Proteins: Lessons from the Nicotinic Acetylcholine Receptor, a Multisubunit Polytopic Protein. In *Membrane Protein Structure: Experimental Approaches* Edited by White SH. Oxford: Oxford University Press; 1993: in press.

37. • BLANTON MP, COHEN JB: Mapping the Lipid-Exposed Regions in the *Torpedo californica* Nicotinic Acetylcholine Receptor. *Biochemistry* 1992, 31:3738–3750.
Two lipophilic labels are used to identify residues in the AChR that face the lipid phase. The most detailed results are obtained for segment M4, a strongly hydrophobic sequence segment located close to the carboxy terminus of each AChR subunit. The data confirm earlier conclusions that M4 is transmembrane and lies at the protein–lipid interface, and suggest that it is indeed an α-helix.

38. • AKABAS MH, STAUFFER DA, XU M, KARLIN A: Acetylcholine Receptor Channel Structure Probed in Cysteine-Substitution Mutants. *Science* 1992, 258:307–310.
The authors introduce cysteines at various positions in a nine-residue stretch in the M2 (putative channel-lining) segment of mouse muscle AChR α-subunit. They express the mutant receptors in *Xenopus* oocytes and determine their sensitivity to small, charged, sulphydryl-blocking agents purported to enter the channel. They conclude that the pattern of blockade observed is more compatible with the nine-residue segment in M2 forming a β-strand rather than a helix. Alternative interpretations are possible (see [39••], for example).

39. •• UNWIN N: Nicotinic Acetylcholine Receptor at 9 Å Resolution. *J Mol Biol* 1993, 229:1101–1124.
A 3D map of the AChR from *Torpedo* is calculated from images of tubular two-dimensional crystals embedded in amorphous ice. *Torpedo* AChR is a heteromeric pentamer homologous to the skeletal muscle AChR analyzed in [38•]. At 9 Å resolution, the only secondary structure elements that can be resolved are α-helices. In the bilayer-spanning section, only one transmembrane helix is resolved per subunit, close to the fivefold axis of symmetry. It is presumed to correspond to the putative transmembrane segment M2. It is proposed that the remaining transmembrane segments form a wide β-barrel surrounding the five central helices.

40.   SIXMA TK, PRONK SE, KALK KH, WARTNA ES, VAN ZANTEN BAM, WITHOLT B, HOL WGJ: Crystal Structure of a Cholera Toxin-Related Heat-Labile Enterotoxin from *E. coli. Nature* 1991, 351:371–377.

41.   STEIN PAE, BODHOO A, TYRRELL GJ, BRUNTON JL, READ RJ: Crystal Structure of the Cell-Binding B Oligomer of Verotoxin-1 from *E. coli. Nature* 1992, 355:748–750.

42. • DONNELLY D, OVERINGTON JP, RUFFLE SV, NUGENT JHA, BLUNDELL TL: Modeling α-Helical Transmembrane Domains: the Calculation and Use of Substitution Tables for Lipid-Facing Residues. *Protein Sci* 1993, 2:55–70.
This paper describes an improved method for predicting the helical nature and orientations of putative transmembrane segments, based upon estimates of the probability of one type of residue being substituted by another on lipid-exposed helix faces. The approach has been tested by using the sequences of BR, halorhodopsin and sensory rhodopsin I and gives a fair prediction of helix orientations in BR.

43. • CALAMIA J, MANOIL C: Membrane Protein Segments as Export Signals. *J Mol Biol* 1992, 224:539–543.
The alkaline phosphatase (*PhoA*) gene is fused to the 3' end of gene fragments encoding selected individual transmembrane segments of *lac* permease and the extramembrane segment that precedes them in the sequence. All except one segment act as export signals, translocating alkaline phosphatase across the plasma membrane into the periplasm. The presence of a transmembrane arginine residue in segment IX reduces its export efficiency. It is suggested that, in the natural situation, insertion of segment IX may be coupled with that of segment X, with which it is thought to interact in mature *lac* permease.

44. • KRIJNSE LOCKER J, ROSE JK, HORZINEK MC, ROTTIER PJM: Membrane Assembly of the Triple-Spanning Coronavirus M Protein. Individual Transmembrane Domains Show Preferred Orientation. *J Biol Chem* 1992, 267:21911–21918.
The M (previously E1) protein of hepatitis virus is synthesized without a signal sequence. Its topology probably features three closely spaced transmembrane α-helices (a, b and c), with a short amino terminus facing the lumen of the endoplasmic reticulum and a large carboxy terminus facing the cytosol. All combinations of one- or two-helix deletions are generated. The mutant proteins are expressed in a cell-free system, and their topology is examined using glycosylation, protease sensitivity and the binding of antibodies as probes. All proteins insert as expected if the hydrophobic segment(s) adopt the same transmembrane orientation as in the intact protein. The mutant in which helix b has been deleted, which contains conflicting vectoriality signals, ends up with a scrambled orientation. Together with previous experiments of the sort (see also [43•]), these experiments further demonstrate that topological information is distributed throughout polytopic proteins rather than being determined only by the orientation of the first transmembrane segment. The topologies of the mutant proteins, however, are difficult to explain simply on the basis of the 'positive inside' rule and it is suggested that additional signals must be involved.

45.   MAGGIO R, VOGEL Z, WESS J: Reconstitution of Functional Muscarinic Receptors by Co-Expression of Amino- and Carboxy-Terminal Receptor Fragments. *FEBS Lett* 1993, 319:195–200.

46. •• KAHN WT, ENGELMAN DM: Bacteriorhodopsin Can Be Refolded from Two Independently Stable Transmembrane Helices and the Complementary Five-Helix Fragment. *Biochemistry* 1992, 31:6144–6151.
Three bacterioopsin fragments are refolded into lipid vesicles and subsequently reassociate in the presence of retinal to regenerate BR. Two of the fragments are synthetic peptides, each containing one of the first two transmembrane α-helices of BR; the third is a five-helix segment derived from BR by protease digestion. Each of the two synthetic peptides is shown to form transmembrane α-helices before reassociating with the rest of the complex. This is the most telling demonstration to date that an IMP can be built from preformed transmembrane helices.

47.   LIAO M-J, LONDON E, KHORANA HG: Regeneration of the Native Bacteriorhodopsin Structure from Two Chymotryptic Fragments. *J Biol Chem* 1983, 258:9949–9955.

48. •• PERVUSHIN KV, ARSENIEV AS: Three-Dimensional Structure of (1–36) Bacterioopsin in Methanol–Chloroform Mixture and SDS Micelles Determined by 2D ¹H-NMR Spectroscopy. *FEBS Lett* 1992, 308:190–196.
This paper and the two that follow [49••,50•] describe two-dimensional NMR studies of synthetic or proteolytic fragments of bacterioopsin in organic media or detergent micelles. A proteolytic fragment encompassing the first transmembrane segment (helix A, residues 1–36) adopts an α-helical conformation from Pro8 to Met32, which corresponds closely to EM observations on native BR. Interestingly, the conformation of most side chains within the α-helical region is restricted.

49. BARSUKOV IL, NOLDE DE, LOMIZE AL, ARSENIEV AS: Three-
•• Dimensional Structure of Proteolytic Fragment 163–231
of Bacterioopsin Determined from Nuclear Magnetic Res-
onance Data in Solution. *Eur J Biochem* 1992, 206:665–672.
In this NMR study, a larger proteolytic fragment (residues 163–231,
which encompass BR helices F and G) is found to comprise two α-
helical regions connected by a disordered segment. The helices corre-
spond, within a few residues, to the transmembrane helices observed
in native BR. As in a similar study on helices A and B [50•], there is no
evidence that the helices interact in solution, as they do in the native
molecule.

50. SOBOL AG, ARSENIEV AS, ABDULAEVA GV, MUSINA LYU, BYSTROV
• VF: Sequence-Specific Resonance Assignment and Sec-
ondary Structure of (1–71) Bacterioopsin. *J Biomol NMR*
1992, 2:151–171.
See [49••].

51. MURRAY-RUST J: Do Specific Interactions Between Trans-
membrane Helices Play a Part in Signalling Transduc-
tion? Exploration with the Insulin Receptor. *Bioessays* 1993,
15:61–62.

52. BORMANN B-J, KNOWLES WJ, MARCHESI VT: Synthetic Polypep-
tides Mimic the Assembly of Transmembrane Glycopro-
teins. *J Biol Chem* 1989, 264:4033–4037.

53. LEMMON MA, FLANAGAN JM, HUNT JF, ADAIR BD, BORMANN B-
•• J, ENGELMAN DM: Glycophorin A Dimerization Is Driven by
Specific Interactions between Transmembrane α-Helices. *J
Biol Chem* 1992, 267:7683–7689.
A genetic grafting experiment in which the anchoring transmembrane
helix of glycophorin A is attached to the soluble nuclease from *Staphylo-
coccus aureus*, causing the chimeric protein to dimerize. This opened
the way for more detailed studies of the sequence requirements for
dimerization (see [54••]).

54. LEMMON MA, FLANAGAN JM, TREUTLEIN HR, ZHANG J, ENGELMAN
•• DM: Sequence Specificity in the Dimerization of Transmem-
brane α-Helices. *Biochemistry* 1992, 31:12719–12725.
The effects on dimerization of 282 mutations in the anchoring segment
of glycophorin A were examined. Amino acid substitutions which re-
placed one hydrophobic residue by another revealed the great sensitiv-
ity of dimer formation to minor sequence changes. Mutation-sensitive
positions occurred with a periodicity of 3.9 residues, suggesting that the
two α-helices in the dimer form a slightly right-handed supercoil. The
region of contact between the two helices is mainly hydrophobic and
contains no strongly polar residues, at variance with the postulated con-
tact areas in several tyrosine-kinase receptors. Close packing probably
plays an important role in the association of glycophorin A anchors.

55. TREUTLEIN HR, LEMMON MA, ENGELMAN DM, BRÜNGER AT: The
•• Glycophorin A Transmembrane Domain Dimer: Sequence-
Specific Propensity for a Right-Handed Supercoil of Helices.
Biochemistry 1992, 31:12726–12733.
Modelling by simulated annealing is used to examine the way trans-
membrane anchors may pack in the glycophorin A dimer. Models
with the lowest (most favourable) interaction energy tend to cor-
respond to the left- and right-handed supercoil arrangements. The
right-handed configuration gives a much better fit to the mutagene-
sis data of [54••]. It predicts that the axes of the two helices approach

quite closely (~6.9 Å) near Gly79, a region of great sensitivity in the mu-
tational analysis. The paper is of particular interest because it illustrates
for a 'simple' case both the difficulties and promises of predicting, *ab
initio*, the structure of transmembrane regions by packing preformed
α-helices.

56. DAVIDSON AL, NIKAIDO H: Purification and Characterization
of the Membrane-Associated Components of the Maltose
Transport System from *Escherichia coli. J Biol Chem* 1991,
266:8946–8951.

57. TRAXLER B, BECKWITH J: Assembly of a Hetero-Oligomeric
•• Membrane Protein Complex. *Proc Natl Acad Sci USA* 1992,
89:10852–10856.
This reports a study of the synthesis of MalF and its stabilization against
exogenous proteases in mutants affected in the synthesis of other sub-
units of the maltose transport complex. The observations strongly sug-
gest that each of the two IMPs in the complex is inserted independently
in the membrane, where it diffuses and associates stochastically with
the complementary subunit. This situation is reminiscent of the mode
of trimerization of influenza virus haemagglutinin in the endoplasmic
reticulum.

58. KIM J, GAMBLE KLEIN P, MULLET JE: Ribosomes Pause at Spe-
cific Sites During Synthesis of Membrane-Bound Chloro-
plast Reaction Center Protein D1. *J Biol Chem* 1991,
266:14931–14938.

59. LIAO M-J, HUANG K-S, KHORANA HG: Regeneration of Native
Bacteriorhodopsin Structure from Fragments. *J Biol Chem*
1984, 259:4200–4204.

60. BIBI E, KABACK HR: Functional Complementation of Internal
•• Deletion Mutants in the Lactose Permease of *Escherichia
coli. Proc Natl Acad Sci USA* 1992, 89:1524–1528.
This paper describes the construction of a series of *lac* permease mu-
tants with large in-frame deletions. None of the molecules catalyze lac-
tose accumulation when expressed alone in *E. coli*. However, some but
not all pairwise combinations of constructs showed functional comple-
mentation, most probably resulting from physical association of two
deleted molecules.

61. KÖSTER W, BRAUN V: Iron (III) Hydroxamate Transport of
Escherichia coli: Restoration of Iron Supply by Coexpres-
sion of the N- and C-Terminal Halves of the Cytoplasmic
Membrane Protein FhuB Cloned on Separate Plasmids. *Mol
Gen Genet* 1990, 223:379–384.

62. SANTANA M, KUNST F, HULLO M-F, RAPOPORT G, DANCHIN A,
GLASER P: Molecular Cloning, Sequencing, and Physiological
Characterization of the *qox* Operon from *Bacillus subtilis*
Encoding the aa₃-600 Quinol Oxidase. *J Biol Chem* 1992,
267:10225–10231.

63. PARDO L, BALLESTEROS JA, OSMAN R, WEINSTEIN H: On the
Use of the Transmembrane Domain of Bacteriorhodopsin as
a Template for Modeling the Three-Dimensional Structure
of Guanine Nucleotide-Binding Regulatory Protein-Coupled
Receptors. *Proc Natl Acad Sci USA* 1992, 89:4009–4012.

J-L Popot, Institut de Biologie Physico-Chimique and Collège de France,
CNRS URA 1187, 11 rue Pierre et Marie Curie, F-75005, Paris, France.