Annu. Rev. Biochem. 2000. 69:881–922

HELICAL MEMBRANE PROTEIN FOLDING, STABILITY, AND EVOLUTION

Jean-Luc Popot[1] and Donald M. Engelman[2,3]

[1]Laboratoire de Physicochimie Moléculaire des Membranes Biologiques, Centre National de la Recherche Scientifique UPR 9052, Institut de Biologie Physico-Chimique, F-75005 Paris, France; e-mail: jean-luc.popot@ibpc.fr; [2]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven Connecticut 06520; [3]Chaire Internationale de Recherche Blaise Pascal de la Région Ile-de-France, Paris, France

Key Words membrane protein structure, two-stage model, helices, lipid bilayer, helix interactions

■ **Abstract** Helical membrane protein folding and oligomerization can be usefully conceptualized as involving two energetically distinct stages—the formation and subsequent side-to-side association of independently stable transbilayer helices. The interactions of helices with the bilayer, with prosthetic groups, and with each other are examined in the context of recent evidence. We conclude that the two-stage concept remains useful as an approach to simplifying discussions of stability, as a framework for folding concepts, and as a basis for understanding membrane protein evolution.

CONTENTS

INTRODUCTION

The protein folding problem has long been recognized as a major conceptual barrier to understanding the information content in genes. In particular, protein folding in aqueous environments has been the focus of considerable laboratory research and many books, articles, and meetings, but progress toward understanding the chemical processes involved has been slow. In this review, we examine whether the folding of transmembrane proteins may be more easily understood than that of soluble proteins, as a consequence of environmental constraints. Since hydrophobicity analyses suggest that membrane proteins represent more than a quarter of all proteins coded in genomes, an understanding of the principles that guide their stability and folding should prove useful (1, 2). Our point of departure is earlier reviews in which we proposed that helical membrane protein folding can be understood in terms of two energetically distinct stages of helix insertion (3) and helix interactions (4; Figure 1). Others have recently published several particularly relevant reviews (5–12).

Our contention is that the lipid bilayer environment strongly limits the range of possible structures for transmembrane proteins, simplifying the folding problem. The first part of the argument, in its simplest form, is that individual helical structures are stable because (*a*) hydrophobic side chains make contacts with the hydrophobic region of the lipid, thus stabilizing a transbilayer location, and (*b*) the hydrogen bonds that form in this process are strong in a low dielectric environment. Although other structures, such as β-barrels, can satisfy these constraints, helices do so locally and with shorter constrained sequences. Many helices found in polytopic proteins may, in fact, be independently stable folding units (13), that is, domains in the strict sense (14). A fascinating, recent observation consistent with this concept is the manner in which bacteriorhodopsin helices can be pulled out of the membrane mechanically, either as single helices or in pairs (F Oesterhelt, D Oesterhelt, M Pfeiffer, A Engel, HE Gaub & DJ Müller, *Science* 288:143–46).

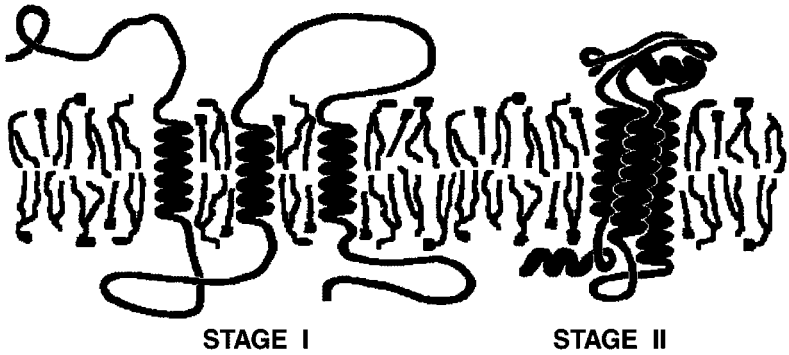

Figure 1 The two-stage model. This figure, from our review of 1990 (4), remains a good summary of our overview. Individual helices of a polytopic protein are postulated to be stable separately as domains in a lipid bilayer. Their stability as domains is a consequence of the hydrophobic effect and main-chain hydrogen bonding. Other interactions then drive side-to-side helix association, resulting in a functional protein. Specific folding energy is provided mainly by packing of the preformed helices with each other, by loop structures, and by interactions with prosthetic groups. Additionally, ion pairs and hydrogen bonds between helices are sometimes found, and general contributions are made by interactions with the lipid environment.

It is notable that helices found in membrane protein structures are mainly predictable by simple hydrophobicity searches of their amino acid sequences, even though polar amino acids may be present (15–17). As discussed below, the sequences in all of the structures known at high resolution (except bacterial outer membrane proteins) contain largely hydrophobic stretches of 18–20 amino acids (Table 1; see Figure 3), which is consistent with the idea that they are helices making favorable contacts with the lipid bilayer. Yet, in the folded proteins, these helix-forming sequences make more contacts with other helices and with prosthetic groups than with lipid. Thus, a concept based on the hydrophobic match of individual helices to the lipid bilayer predicts helices in the largely nonlipid environment within a folded polytopic protein, giving rise to the second part of our argument, which follows.

If individual helices, formed in response to main-chain hydrogen bonding and the hydrophobic effect, retain their identities in the final folded state of polytopic proteins, then other interactions must dominate in causing their side-to-side assembly. Such interactions might include packing, electrostatic effects, prosthetic-group interactions, loops between helices, and binding to components in the aqueous environments that surround the membrane. The formation of helices (the first stage) is then separated from their assembly into folded proteins (the second stage) in terms of the dominant energy contributions (Figure 1). This significantly simplifies our thinking about folding, because the helices and their orientations can often be predicted from the sequences, and side-to-side interactions are constrained in

TABLE 1 Integral membrane protein structures known at high resolution

Protein	Organism	Method[a]	Current best resolution (Å)[b]	Type[c]	First date[d]	References
Photosynthetic reaction centers	*Rhodopseudomonas viridis*	X-ray	2.3	α	1984	36, 56, 165, 168c
	Rhodobacter sphaeroides R-26	X-ray	2.2	α	1986	164, 166–167a, 168a
Light-driven ion pumps from halophilic bacteria						
Bacteriorhodopsin	*Halobacterium salinarium*	EM	3.0 (n.s.)	α	1990	35, 51, 52, 171–172
		X-ray	1.9	α	1997	169, 170, 173
Halorhodopsin[e]	*H. salinarium*	X-ray	1.8	α	2000[f]	
Bacterial outer membrane proteins						
Porin	*Rhodobacter capsulatus*	X-ray	1.8	β	1990	101, 174–176
Matrix porin (OmpF)	*Escherichia coli*	X-ray	2.4	β	1992	100
Phosphoporin (PhoE)	*E. coli*	X-ray	3.0	β	1992	100
Porin	*Rhodopseudomonas blastica*	X-ray	2.0	β	1994	177
Maltoporin (LamB)	*E. coli*	X-ray	3.1	β	1995	178
	Salmonella typhimurium	X-ray	2.4	β	1997	180
Porin	*Paracoccus denitrificans*	X-ray	3.1	β	1997	179
Sucrose-specific porin ScrY	*S. typhimurium*	X-ray	2.4	β	1998	180a
Outer membrane protein A (OmpA) transmembrane region	*E. coli*	X-ray	2.5	β	1998	62
Ferrichrome-iron receptor (FhuA)	*E. coli*	X-ray	2.5	β	1998	79, 102
Iron transporter (FepA)	*E. coli*	X-ray	2.4	β	1999	103
Outer membrane protein X (Omp X)	*E. coli*	X-ray	1.9	β	1999	180c
Osmoporin (OmpK36)	*Klebsiella pneumoniae*	X-ray	3.2	β	1999	180b
Outer membrane phospholipase A (OMPLA)	*E. coli*	X-ray	2.4	β	1999	180d

Prostaglandin H_2 synthase I	*Ovis aries*	X-ray	3.1	—	1994	181
Prostaglandin H_2 synthase II	*Musmusculus*	X-ray	2.5	—	1996	182
	Homo sapiens	X-ray	2.9	—	1996	183
Eukaryotic light-harvesting complex II (LHC II)	*Pisum sativum*	EM	3.4 (4.5)	α	1994	96
Bacterial light-harvesting complex 2 (LH2)	*Rhodopseudomonas acidophila*	X-ray	2.5	α	1995	161
	Rhodospirillum molischianum	X-ray	2.4	α	1996	162
Cytochrome *c* oxidase	*P. denitrificans*	X-ray	2.7	α	1995	54, 185b
	Bos taurus	X-ray	2.3	α	1995	44, 184–185a
α-Hemolysin	*Staphylococcus aureus*	X-ray	1.9	α	1995	186
Cytochrome *bc*1	*B. taurus*	X-ray	3.0	α	1997	97, 187, 188, 188a
	Gallus gallus	X-ray	3.2	α	1998	187
	Saccharomyces cerevisiae	X-ray	2.3	α	2000[e]	
Glycophorin A transmembrane region	*H. sapiens*	NMR	RMSD:0.4	α	1997	60
Squalene-hopene cyclase	*Alicyclobacillus acidocaldarius*	X-ray	2.0	—	1997	190, 191
K^+ channel	*Streptomyces lividans*	X-ray	3.2	α	1997	99
Mechanosensitive ion channel	*Mycobacterium tuberculosis*	X-ray	3.5	α	1998	189
Fumarate reductase	*E. coli*	X-ray	3.3	α	1999	192, 192a
ATP synthase F_1c_{10} subcomplex	*S. cerevisiae*	X-ray	3.9	α	1999	192b

[a]EM, electron microscopy; n.s., not specified; RMSD, root mean square derivation.

[b]For those structures established by EM, where resolution is anisotropic, the resolution normal to the plane of the membrane is indicated in parentheses.

[c]The type of structural organization of the transmembrane region (α-helix bundle or β-barrel) is indicated. Prostaglandin synthase and squalene-hopene reductase are integral but not transmembrane.

[d]The date indicated is that of the first atomic model. A compendium of known structures of membrane proteins with their PDB access codes is available at *http://www.biophys.mpg.de/michel/public/memprotstruct.html.*

[e]C Hunt, J Koejke, C Lange, T Rossmanith, H Michel, submitted for publication.

[f]M Kolbe, H Besir, L-O Essen, D Oesterhelt, submitted for publication.

relative tilts and translations. The skeptical reader will, at this point, raise a number of challenges to the simple concept just delineated, and we explore some of them in the following review of pertinent literature. We first discuss the solvent environment of a membrane protein, including lipid bilayers. We then examine the contention that helical structures are preferred, often stable, but somewhat variable entities in the transbilayer environment. Helix interactions and other folding factors are our next topics. We conclude that the two-stage folding concept has merit, even though the picture is not so simple as suggested above. The perspectives that we have developed regarding folding also lead to notions that may be pertinent in thinking about functional states and evolution, with which we conclude our discussion.

LIPID BILAYERS AND THE MEMBRANE PROTEIN ENVIRONMENT

The fluid lipid bilayer, in which membrane proteins fold and function, has regions where distinctively different environments interact with various parts of the protein's surface, creating a set of constraints that simplify folding concepts. At the simplest level, there are three regions: the hydrophobic bilayer center, the head group layers, and the aqueous regions (Figure 2). Each of these has a set of distinctive properties, which we outline briefly to establish a context for the subsequent discussion.

For simplicity, we cover separately the interrelated aspects of bilayer planar organization and profile perpendicular to the plane. A bilayer in the L_α or liquid crystalline phase that is found in biological membranes is often considered a random two-dimensional liquid. This is unlikely, because the distribution of the many lipid components in a membrane can be random only if all pairwise interaction energies of the species are within thermal energies ($\sim$0.6 kcal/mol at room temperature) of each other, an unlikely condition. Thus, it should have been expected that regions of different composition would exist and that the environments of proteins should vary, because protein-lipid interaction energies also fail to match each other across all protein and lipid species in a membrane. Experimental observations show that differential interactions can lead to significant planar separations, as seen in the "rafts" resulting from sphingomyelin-cholesterol interactions (see 18–20).

A remarkable property of lipids is the extent to which the area per lipid molecule is conserved for a given head group. If the hydrocarbon chains in a bilayer with a given head group are varied in length, the thickness of the bilayer varies linearly with the chain length (21). Bilayer vesicles that are subjected to osmotic stress rupture and reseal to relieve the stress, rather than thinning to increase area by $>5\%$ (22). Small increases ($\sim$1%) in the area of one monolayer, which are achieved by the addition of amphiphiles, cause gross shape changes in a bilayer. The conserved area arises from the balance of three regions of interaction: forces between head groups (acting to increase area per lipid), exposure of hydrophobic areas of chains

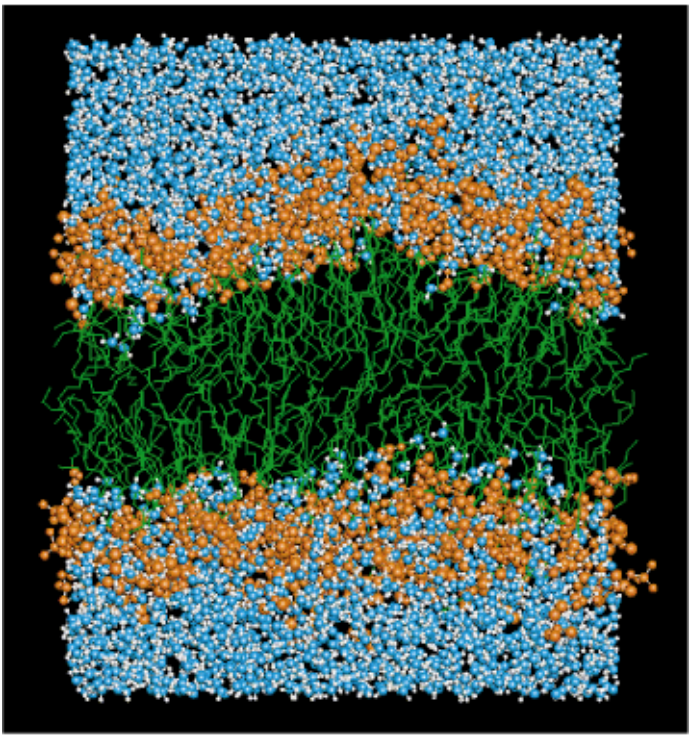

Figure 2 A simulation of a hydrated lipid bilayer. A molecular-dynamics simulation of a lipid bilayer in water (223) illustrates many features known from experiments. A partially disordered hydrocarbon layer ∼30 Å thick is formed by the lipid hydrocarbon chains, where virtually no water molecules are present. Two head group layers, each ∼15 Å thick, bound the hydrocarbon; the lowered concentration of water resulting from the high concentration of the head groups is evident.

to water (acting to decrease area per lipid), and the internal lateral pressure of the chains against each other (acting to increase area per lipid) (23). Consequences include the general orientation of the chains across a bilayer and the localization of terminal methyl groups near its center. There is evidence that lateral pressure, and its variations as a function of the depth in the bilayer, may affect both membrane protein folding and function (see e.g. 23a,b and references therein).

The profile of a lipid bilayer is often represented to emphasize its nonpolar region, but an approximately equal contribution to the thickness is made by the head group layers, each of which may be 10–15 Å or more, thus providing a total of >20–30 Å over which a membrane protein can interact (24; Figure 2). The head group layers are extremely concentrated if viewed as solutions (∼500–700 mg/ml for phosphatidyl choline). Most of the water in this region is used in hydrating the head groups, so the availability of water for the hydrophobic effect or for additional hydration is decreased compared with that of bulk aqueous solution. Not only

are the head group layers extensive and concentrated, but they are varied—the distribution of characteristic groups between the two layers is not equal, and the chemical variety within each layer can be considerable, including head groups with net charge, with dipoles, and with sugar groups. Given this diversity of chemistry and the diversity of proteins, it is almost certain that preferential associations of lipid species with particular proteins occur, adding to the nonrandom character in the bilayer plane and contributing to the stability of the proteins (25, 26).

The hydrophobic region of the lipid bilayer, a layer ~30 Å thick and largely impermeable to polar molecules and ions, defines the properties of membranes as permeability barriers. It also provides the most distinctive region of solvation for membrane proteins; ionic interactions are strong over long distances in the low dielectric, the hydrophobic effect is absent, and several other distinctive properties prevail (Table 2). Although often regarded as a liquid, the hydrocarbon chains are partially ordered, with pairs of ends covalently constrained to be next to each other and with their terminal methyl groups localized near the center of the bilayer. Moreover, sterols can further restrict chain motion in many membranes. However, a much larger perturbation of lipid structure is created in most membranes by the presence of transmembrane proteins, which have been estimated to occupy 20% of the lipid area of an erythrocyte membrane (27) and probably a larger fraction of mitochondrial and chloroplast inner membranes (see Table 2). The relatively rigid sides of the protein molecules restrict the motions of the lipid chains, and the hydrophobic thickness and polar interactions perturb the local lipid conformations and composition (26, 28).

Much has been made of the phase behavior of lipids. Starting with the seminal work of Luzzati (e.g. 29, 30) and stimulated by the finding that phase changes can occur in biological membranes (31, 32), a large body of literature has emerged (33). Mostly, the focus has been the "gel-liquid crystal" or L_β-L_α phase transition, in which the chains adopt a regular hexagonal packing in the low-temperature phase. It seems, however, that organisms seek to avoid the L_β state, suppressing it by adjusting chain length and saturation. Given the biology, it seems appropriate to set aside consideration of this aspect of lipid behavior in thinking about membrane protein folding. A more subtle and potentially important property of the hydrophobic region is the variation in fluidity that results from variations in chain length, sterol content, degree of unsaturation, and regional variation.

Finally, we come to the aqueous environment. The concentration of macromolecules in the vicinity of most membranes is high, and the aqueous spaces on the sides of the membrane differ in chemical composition, pH, and electrical potential (Table 2). It might be thought that these factors contribute to the stability and folding of membrane protein regions exposed to the aqueous environment as they do in the folding of soluble proteins, but there are differences that might be significant. First, the region adjacent to the membrane has gradients of composition and potential not seen in bulk solution. These arise from the transmembrane potentials as well as from the properties of the layer of charges in the characteristic groups and the adjacent influence of the low dielectric bilayer interior. Second,

TABLE 2 Properties of the cytosolic and membrane environments that are relevant to protein folding and association

Property	**Cytosol**	**Plasma membrane**[a]
Solvent chemical homogeneity	Yes	No
Chemical groups available	HOH, ions, −SH	$-CH_3$, $-CH_2-$, $=CH-$
Isotropy	∼Yes[b]	No
pH gradient	No	Yes[c]
Electric field ($V \cdot m^{-1}$)	∼0[b]	$\sim 2 \times 10^{6}$ [c]
Pressure gradient	No	Yes
Dielectric constant gradient	No	Yes
Redox potential gradient	No	Yes[c]
Volume or surface occupancy [protein/solvent (%)][d]	∼17	∼35[e]
Separation between two proteins:		
Distance (Å)	∼50	∼30–35
Intervening solvent molecules	∼15–20	∼4
Exchange time between solvent molecules(s)[f]	$\sim 10^{-11}$	$\sim 10^{-7}$
Viscosity at 20°C (η; $N.s.m^{-2}$)	0.001	0.1
Dimensions	3	∼2
Translational diffusion:[g]		
D_{lat} ($m^2 \cdot s^{-1}$)	$\sim 10^{-10}$	$\sim 10^{-11}$
Average range explored in 1 μs ($\bar{x}$; Å)	∼250	∼50
Dielectric constant (ε)	80	∼2[h]
ΔG° ($kcal \cdot mol^{-1}$) for:		
Breaking a main chain H-bond	∼0	+4−6[i]
Deprotonating a Glu side chain (pH7)	−4	>30[j]
Opening a salt bridge	<1	60[k]
Exposing one $Å^2$ of hydrophobic surface	+0.025[l]	∼0
Exposing a Leu side chain to the solvent	+2.8[j]	∼0
Associating two 50-kDa proteins ($T\Delta S$ at 20°C)	8[m]	5[m]

[a]For properties that vary as a function of the depth in the membrane, the data correspond to those at the membrane center. [b]Note, however, that the cytosol is heavily encumbered (cf. Reference 193). [c]In most but not all membranes. [d]Estimated from data compiled by Goodsell (193) for the cytosol and plasma membrane of an *E coli* cell; calculations for the cytosol are based on Goodsell's estimates for the average size of a soluble protein and assume a 1:2.5 w/w ratio (194) of RNA to protein (194); calculations for the plasma membrane assume the average integral protein (often an oligomer) to comprise ∼12 transmembrane helices and to have about half of its volume buried into the membrane (cf. Reference 39); estimates published in the literature vary from 17% to 50% (reviewed in 195; see also Reference 196). [e]Note that the percolation threshold for short-range diffusion of small molecules in the membrane plane is ∼50% (see References 195, 196). [f]In pure solvent. [g]For a middle-sized protein (∼50-kDa) in either pure water or pure lipids; in the cytosol and in real membranes, diffusion coefficients vary with the distance range considered; see References 195–197. [h]From Reference 198. [i]From Reference 199. [j]Based on Reference 3. [kj]From Reference 43. [l]From Reference 200. [m]Based on Reference 201.

loops in polytopic proteins are constrained to have their ends in the bilayer, perhaps positioned at a defined distance from one another. Defining the positions of its ends greatly restricts the possible conformations of a loop.

Thus, we see that the environment in which membrane proteins find themselves is nonuniform and complex along an axis perpendicular to the membrane, but more homogeneous within coplanar layers. We summarize a few properties in Table 2 and show some in Figure 2. Undoubtedly the details of this complexity influence the stability and folding of membrane proteins and must, in the end, be taken into account. It is therefore a pleasant surprise that useful insights appear to be possible when greatly simplified experimental and conceptual frameworks are adopted, as they are below.

Lipid Environment Summary

1. Bilayer profiles define a hydrophobic layer (∼30 Å thick) and two concentrated head group layers (each ∼15 Å thick).
2. Planar separation of regions of different lipid composition within a bilayer should be expected and is observed. Such regions vary in thickness and fluidity.
3. Lipid molecular areas tend to be conserved, and lateral forces vary across the bilayer profile.

CHARACTERISTICS OF SINGLE TRANSMEMBRANE HELICES

With knowledge of the secondary structures of bacteriorhodopsin and the photosynthetic reaction center came a general view that membrane proteins often comprise bundles of transmembrane helices (34–36) (Figure 3). Since then, conceptual and modeling exercises have often assumed a set of ideal α-helices traversing the bilayer. Here we examine the properties of actual helices, to develop a more realistic view based on new structural information in the last decade.

Hydrophobicity

As mentioned in the introduction, the most conspicuous characteristic of transmembrane helices is the presence of a 15- to 20-amino-acid sequence that is, on average, hydrophobic. Hydrophobicity arises from the ordering of water by a nonpolar surface, and this property can favor association of a helix with the nonaqueous bilayer interior relative to the aqueous milieu. When hydrophobicity is assessed by any of a number of scales, the transmembrane helices of the known structures are usually identified with reasonable accuracy (see e.g. 16, 37, and references quoted in 38). However, errors of three kinds are made. First, some of the transmembrane segments in polytopic proteins have average hydrophobicities that overlap those of nontransmembrane segments, regardless of which scale or window length is used

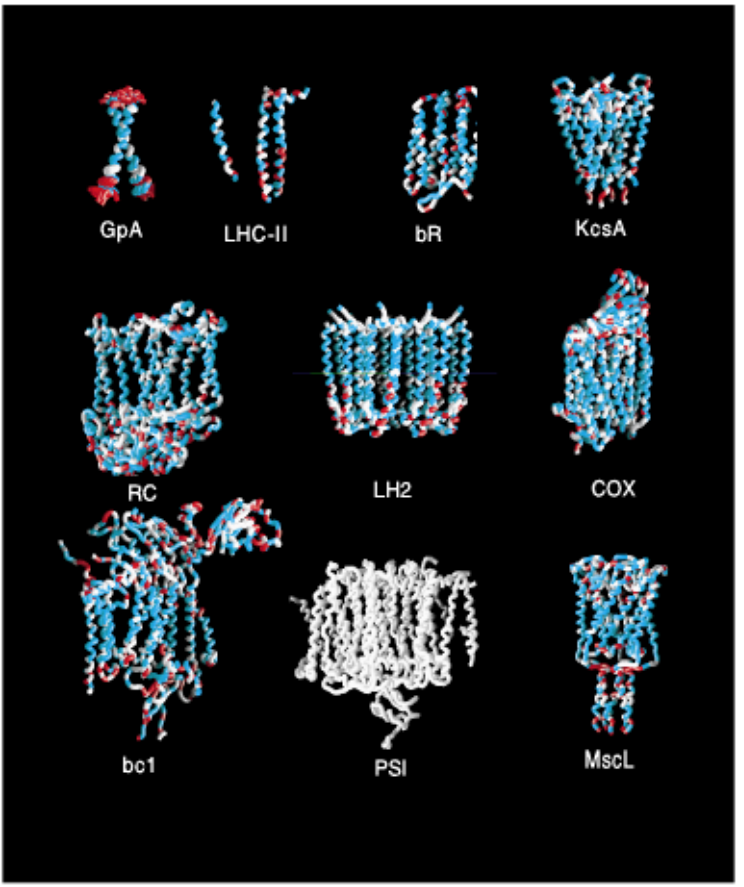

Figure 3 A gallery of known helical-membrane protein structures. The structures available as protein database entries in June 1999 are shown. Whenever several homologous structures are known, only one is shown. Hydrophobic side-chain positions (A, V, L, I, F, M) are in *cyan*, and strongly ionizable residues (D, E, R, K) are in *red*. The hydrophobic character of the transmembrane helices is clear. Prosthetic groups are omitted, resulting in the gaps seen. The proteins shown and their protein database files are as follows: GpA, glycophorin A structure from nuclear magnetic resonance [1AFO (60)]; LHC-II, eukaryotic light-harvesting complex II [electron microscopy at 3.4 Å; 1LHC (96)]; BR, bacteriorhodopsin [electron microscopy at 3.0 Å; 2AT9 (173)]; KcsA, K^+ channel [X-ray at 3.2 Å 1BL8 (99)]; RC, photosynthetic reaction center [X-ray at 2.2 Å; 1AIJ (224)]; LH2, bacterial light-harvesting complex [X-ray at 2.5 Å; 1KZU (225)]; COX, cytochrome oxidase [X-ray at 2.7 Å; 1AR1(54)]; bc1, cytochrome bc_1 complex (transmembrane subunits only) [X-ray at 3.7 Å; 3BCC (187)]; PSI, photosystem I (alpha carbons only) [X-ray at 4 Å; 2PPS (142)]; MscL, mechanosensitive ion channel [X-ray at 3.5 Å (189)]. This and following figures were created by using the GRASP program (226).

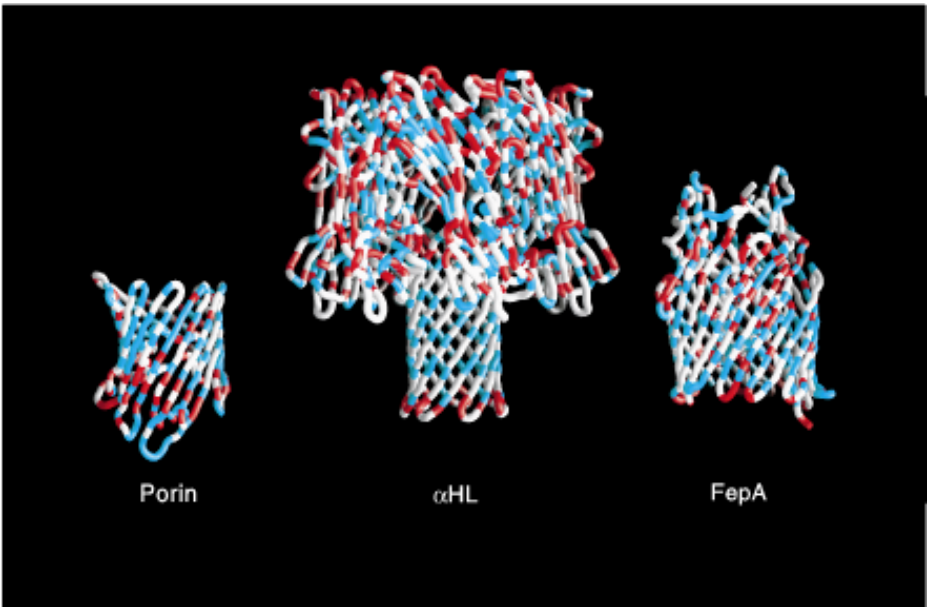

Figure 4 Examples of β-barrel structures. Three examples are shown for comparison with Figure 3: porin [2POR (101)]; αHL, α-hemolysin [7AH1 (186)]; and FepA, ferric enterobactin receptor [1FEP (103)]. Hydrophobic groups are shown in *cyan*, and strongly ionizable groups are in *red*.

for the analysis (37, 39). Ambiguities have to be removed either by consideration of the composition of nearby, extramembrane regions (40) or by experiment. Second, the beginning and end of a helix may not be well specified. Third, a helix tethered by a short extramembrane loop to a more hydrophobic helix (a "hairpin") may be assisted in partitioning into the membrane even though it is less hydrophobic than would be required for independent insertion (3, 41, 42).

The structures of helical membrane proteins that are known with enough resolution to identify the amino acids in their transmembrane regions are listed in Table 1 and shown in Figure 3, with hydrophobic amino acids highlighted. For comparison, some of the β-barrel proteins are shown in Figure 4. In helix searches, a moving window function is used to represent the bilayer profile, which is most often taken as a square wave of hydrophobicity, although trapezoidal and gaussian functions have been used as better representations of the water distribution (24; for a review, see 37). Some helices have longer hydrophobic regions than are required to traverse the bilayer; these may be tilted or may extend into the hydrophobic cores of domains outside the bilayer. Although most helices in the known structures are identified by hydrophobicity, it is clear from the structure of the eukaryotic light-harvesting complex (LHC-II) that other factors may influence the transmembrane structure in some cases. In the LHC-II, hydrophobic regions are identified, but

appear to be partially displaced from the bilayer in the final structure, in which polar groups are buried as ion pairs involved in prosthetic group binding. Ion pairing greatly facilitates the insertion of charges into the membrane (43).

Hydrophobicity Stabilizes Diversity

If a helix has completely hydrophobic amino acids (e.g. L, I, V), the favorable free energy of transfer from water to bilayer greatly exceeds the minimum required for stability. Thus, amino acids with unfavorable free energies of transfer or poor helix-forming propensities can be accommodated within a helix, which allows functional groups to be placed in the membrane. As mentioned above, excess stability of one helix may be used to stabilize another by forming a helical hairpin, in which two helices are connected by a short segment of polypeptide. Further, the formation of the main-chain hydrogen bonds and helical conformation can be seen as a consequence of the hydrophobic partitioning of the peptide.

π-Helix Turns

High-resolution crystal structures are rare for helical membrane proteins (Table 1), so the apparent deviations from ideal α-helicity that are observed in several current interpretations must be confirmed at higher resolution (<2 Å) before their status can be certain. Some particularly tantalizing suggestions, such as the existence of a largely disordered segment in the transmembrane core of mitochondrial cytochrome *c* oxidase (44) or the mixed α/β fold of the transmembrane region of the nicotinic acetylcholine receptor (45, 45b), are examples in which improved resolution would be welcome. There are cases, however, in which current structural information leaves little doubt about the existence of local 3_{10} or π conformation (helices with $n \leftrightarrow n + 3$ and $n \leftrightarrow n + 5$ main-chain hydrogen bonds, respectively, rather than $n \leftrightarrow n + 4$ as in α-helices, where n is the sequence number of the amino acid).

In the photosynthetic reaction centers from *Rhodopseudomonas viridis*, for instance, the single transmembrane helix of subunit H features, in its C-terminal region, a "π-bulge"—also called "looping out" (46, 47) or "alpha-aneurysm" (48; Figure 5). It is interesting that the homologous protein from *Rhodobacter sphaeroides* shows ideal α-helicity in the same region. In *R. viridis*, four residues are in position to form branched hydrogen bonds with an $n \leftrightarrow n + 4/5$ pattern, four hydrogen bond donors (the main-chain NH groups of residues 29–32) interacting with three hydrogen bond acceptors (the C═O groups of residues 25–27). This makes it possible for the helix to accommodate one extra residue in *R. viridis* as compared with *R. sphaeroides*, without creating a stagger between the two ends of the helix, so that its environment is not seriously perturbed either above or below the bulge (Figure 5). Such a response to insertion has been induced experimentally in engineered mutants of T4 lysozyme (46, 47). In the reaction center, the presence of the bulge frees two C═O groups pointing towards the periplasmic head group region, which are now unable to form hydrogen bonds with the surrounding residues

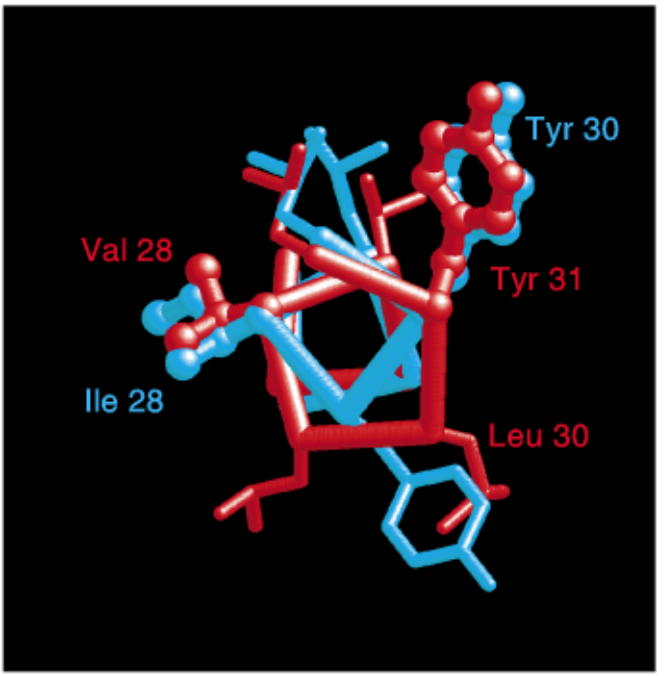

Figure 5 A π-bulge compared with an α-helix. In the single transmembrane helix of the H subunit of the photosynthetic reaction center of *Rhodopseudomonas viridis*, a π-bulge is found (*red*), but the homologous region in the *Rhodobacter sphaeroides* H subunit is 1 amino acid shorter and has no bulge (*cyan*). Bulges are a variation on helical structure that may be more widespread than is usually thought (see text).

and are presumably involved in other interactions either with water molecules or with lipids.

Is the bulge in the *R. viridis* H helix spontaneous or induced upon interaction of the helix with the rest of the complex? Molecular-dynamics simulations of the transmembrane helix of the ErbB-2 receptor suggest that $\alpha \leftrightarrow \pi$ transitions need not be energetically very costly and that they are facilitated by the presence of β-branched side chains, which are frequent in transmembrane helices (49, 50). Their helix-breaking potential, may be important in facilitating deviations from ideal α-helix conformation, reintroducing the concept of sequence-induced conformational propensity.

π-Bulges can also have functional roles. In the recently published 1.9-Å-resolution structure of bacteriorhodopsin (51), a π-bulge is seen in the middle of helix G, in the vicinity of the Schiff base involved in proton pumping. Here, the altered H bonding and steric profile are likely to be components of the mechanism, contributing to the organization of water molecules (see Figure 8 below). This point is made even more clearly in a more recent higher-resolution structure

(52). Several other potential π-bulges have been identified in published membrane protein structures (53). The identification of some of them is tentative owing to the limited resolution of the structures, but others must be real. π-Bulges provide an opportunity for polar interactions between transmembrane segments; for instance, the existence of a bulge in the single transmembrane helix of subunit VIa of beef cytochrome *c* oxidase may free a C═O group for hydrogen bonding to the side chain of a tryptophan residue in helix VIII of subunit I, across the interface between the two 13-mers.

Three-Ten Helix

In the recent 2.7-Å-resolution structure of the core region of cytochrome *c* oxidase from *Paracoccus denitrificans* (subunits I and II), helix VII from subunit I has an extended region of transmembrane 3_{10} helix covering residues 320–331 (54). This section of the helix lies in a highly important region of the protein, where residues 325 and 326 provide two of the three histidine ligands to copper.

Deviations from ideal α-helix structure have two practical consequences: they reintroduce, in the apparently monotonous landscape of α-helix bundles, a measure of diversity and flexibility, which may be of great importance for function, and, unfortunately, they drop one more stone in the rocky garden of the model builder. Establishing a sequence-based prediction of alternative helical structures would be a very useful step, if it were possible.

Individual Amino Acid Roles

A bundle of completely hydrophobic α-helices is a very limited structure with which to accomplish transmembrane functions, so it is not surprising that more polar groups are found, including Q, N, H, D, E, R, and K. Many of these have been implicated in proton or electron transport or in binding prosthetic groups. When these groups are placed in a low dielectric environment, they are likely to be in their uncharged forms, but will nonetheless remain strongly polar. In general they are not exposed to direct contact with the lipid in the known polytopic structures, although this may be the case in the F_o region of the ATP-synthase (192b). Some may undergo functional protonation/deprotonation as for D85 and K216 in bacteriorhodopsin, or the acidic (*DorE*) residue in subunit c of ATP-synthases. In most cases, the net hydrophobicity of the helix still favors a transmembrane position or nearly so, based on free-energy scales. Although there are clear exceptions in various models for structures, such as the gating charges in some channel models, a generalization from the known structures is that polar groups can be included in a helix, provided either that the helix remains hydrophobic on average or that group polarity is diminished, as in the ion pairs of the LHC-II.

As previously noted, β-branched side chains I and V are accommodated (and abundant), as are G and P, even though they all tend to be helix breakers in soluble proteins. These occurrences emphasize the importance of the contribution of the hydrophobic environment to the stabilization of hydrogen bonding and helix

formation—the energy from the hydrophobic effect drives helix formation and main-chain H bonding, overwhelming other unfavorable terms.

A consequence of helix formation is that the rotamers available to some side chains are restricted. A rotamer is a set of allowed rotation angles of side-chain bonds, and restriction of rotamers is a major energy factor in protein folding, mainly for entropic reasons. For example, V has only one allowed rotamer in a helix, but has three in a coil. Thus, the restriction of rotamers is partly done by the formation of the helix, reducing the additional energy needed for helix-helix interactions and contributing to specificity (see 55 for discussion of glycophorin).

The hydroxyl groups of S and T participate in shared hydrogen bonds to carbonyl groups of the polypeptide chain (56). It is less frequently noted that the SH group in C is likewise hydrogen bonded in many cases (Figure 6), which is favored in

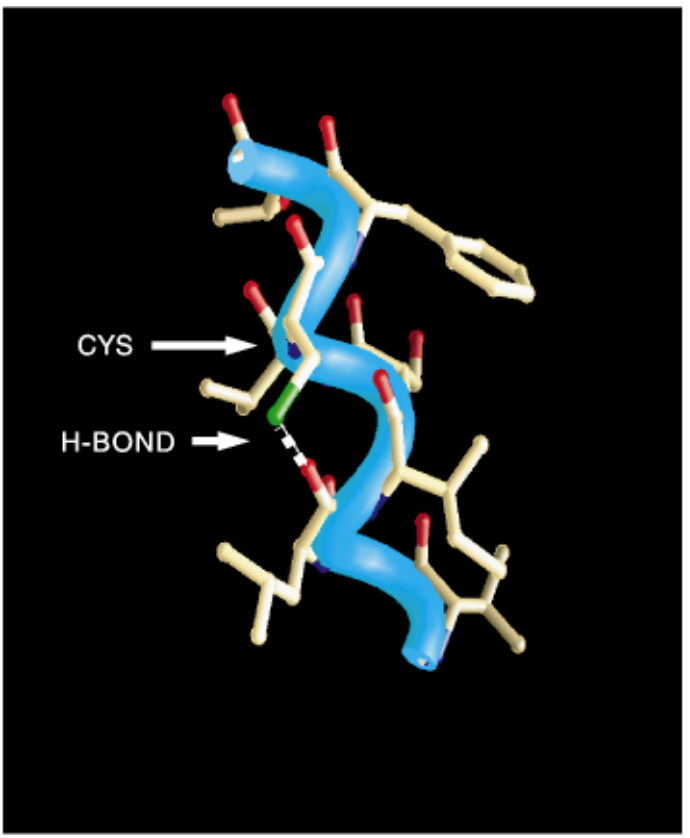

Figure 6 Cysteine hydrogen bond in a transmembrane helix. An example from the structure of cytochrome oxidase (54), showing the hydrogen bond made by the cysteine side-chain to a main-chain carbonyl, reducing the effective polarity of the SH group. The *dashed line* is from the sulfur atom to the carbonyl oxygen; the hydrogen would lie roughly along it.

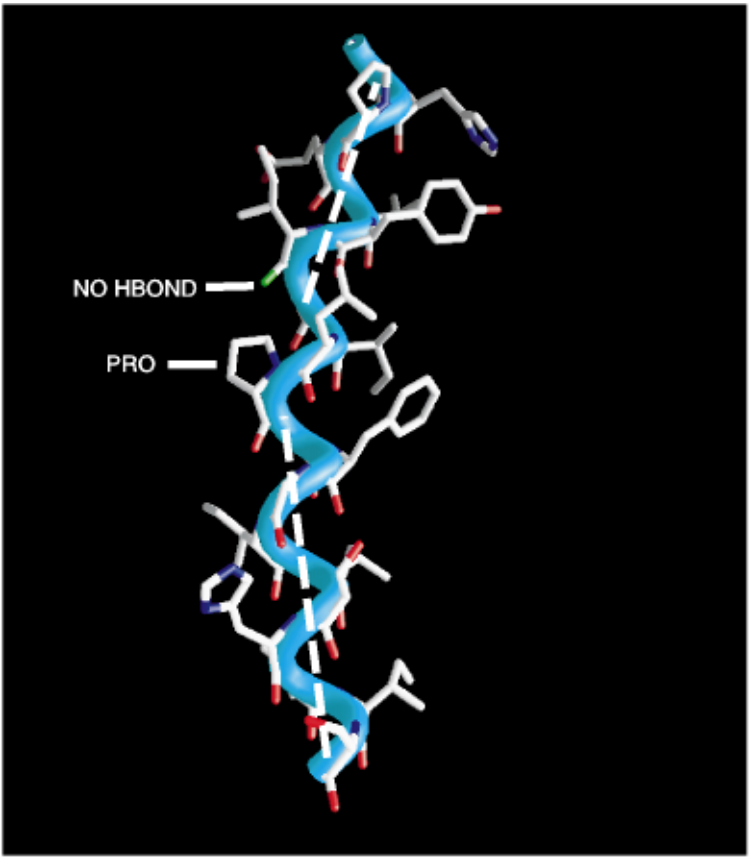

Figure 7 Proline bend in a membrane helix. An example from the cytochrome oxidase structure (54), showing a helix bend such as is often, but not always, induced by proline. Note the increased spacing of the main-chain turn above the proline and the carbonyl devoid of hydrogen bond.

a low dielectric environment, because the net polarity is reduced. Because of this effect, S, T, and C are more hydrophobic than might otherwise be supposed. They are frequently found in transmembrane structures.

P, because it has no NH group in the backbone and because one of the backbone rotation angles is locked by the ring, perturbs both the polarity and the geometry of a helix. The most common effect is to create a bend in the helix axis and leave a main-chain CO without a hydrogen-bonding partner (Figure 7). But other alternatives are seen, including a set of local conformations of the main chain that result in a straight helix even though a P is included. An example of such a structure is seen in helix 2 of subunit L of the photosynthetic reaction center, where P92 is in a straight helical segment (56).

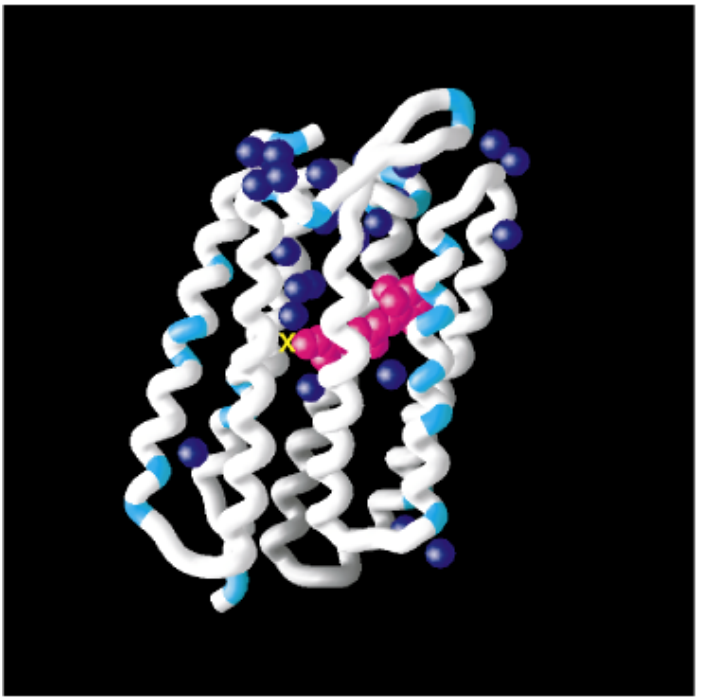

Figure 8 Water and glycines in bacteriorhodopsin. The recently published 1.9-Å structure of bacteriorhodopsin (51) shows the locations of water molecules (*blue*) among the transmembrane helices, particularly clustered near the retinal (*purple*) Schiff base (*yellow X*), where the water functions in the proton pathway. Glycine (*cyan*) positions are shown to illustrate their abundance in the transmembrane helices.

It is remarkable that G is so abundant in membrane helices, given its tendency to be in the nonhelical regions of soluble proteins. For examples, the fifth helix of bacteriorhodopsin has five (57, 58), as shown in Figure 8, and the glycophorin helix has four (59). In glycophorin, two of the glycyl residues permit the close packing that stabilizes the helix dimer (60), and GXXXG motifs are abundant in other cases (60a,b).

Aromatic side chains, particularly W and Y, are often located near the boundary between the hydrophilic and hydrophobic regions of the lipid bilayer in the structure of both helical and β-barrel membrane proteins (9, 36, 61, 62, 99, 175, 189). Statistical studies of sequence profiles for putative helices in the databases also show that these amino acids tend to partition in this region. The physical basis for an association with the transitional region of polarity is thought to be the relative polarity of the π electrons in the aromatic ring structures and, possibly, the shapes of these side chains (63). The finding of an association of this kind and of a longitudinal amino acid variation in the hydrophobic region gives further opportunities for refining prediction methods for transmembrane helices.

The occurrence of longitudinal compositional variation raises interesting issues concerning the fit of transmembrane helices to the bilayer profile. To date, this kind of fit has been examined only approximately (64–66). The regions of high head group concentration, the transition between head groups and hydrophobic chains, and the hydrophobic bilayer interior may each select for subpopulations of amino acids. Most intensively examined have been the bilayer interior and the interface; it remains to examine which rules may apply in the head group layers (64). An excellent review has appeared, with careful consideration of these issues (11).

Water

As high-resolution (<2-Å) structures are determined, a number of water molecules in the transbilayer region can be identified with confidence. Two recent bacteriorhodopsin structures, one from untwinned crystals at 1.9 Å (51) and the other from partly twinned crystals at 1.55 Å (52), reveal networks of waters that support function and also bridge helices (Figure 8). Few such waters are seen in other cases, such as the reaction center, but these structures are at a lower resolution (Table 1). The presence of water molecules inside transmembrane helix bundles further increases the diversity of chemical groups available to carry out function.

Helix Dipoles

The alignment of hydrogen bond dipoles in a helix gives rise to a net separation of charge between the helix ends, the helix dipole (67). Each end of a helix has about half a unit of charge, distributed over the end in a delocalized volume. For helix ends exposed at the membrane surface, solvation and counterion effects strongly damp the helix charge effects (68). Placing the end of a helix in the nonpolar bilayer interior would remove these influences, creating a comparatively high free energy of transfer. Thus, it may not be surprising that lipid buried helix ends are not (to date) found.

Single-Helix Summary

1. Single transmembrane helices are characterized by largely (but not entirely) hydrophobic sequences of 15–20 amino acids.
2. These may include a limited number of polar or potentially charged groups, and C, T, and S with H bonds to the main chain.
3. Helix formation is driven by the hydrophobic effect and includes the restriction of side-chain rotamers.
4. Longitudinal composition varies, placing more Y and W side-chains near the transition between the hydrophobic and hydrophilic regions of the bilayer.
5. Variations in helix parameters to include regions of π and 3_{10} helix need to be kept in mind, because they are observed to occur in known structures, but no reliable means of predicting them exists at this time.

6. P and G are found more commonly in membrane helices than in soluble-protein helices. As in soluble proteins, P can induce a bend in a helix axis.
7. Water can be an essential component of transmembrane bundles.

INTERACTIONS BETWEEN TRANSMEMBRANE HELICES

Given the stability of transmembrane helices in response to main-chain hydrogen bonding and the hydrophobic effect, what are the interactions that might influence their association to form a higher-order structure? Which interactions are general and which specific for folding? The major possibilities are as follows:

1. Differential lipid effects, such as the relative packing and entropy of lipids interacting with helices rather than with other lipids, may favor helix association.
2. The loops between helices constrain their location and may organize specific orientations if the loops are short or form ordered domains.
3. Specific interactions with prosthetic groups located between helices may provide a driving force.
4. Single, strongly polar interactions, such as interhelical hydrogen bonds, bridging water molecules, or ion pairs, could result in association.
5. Favorable Van der Waals interactions promoted by a detailed fit between surfaces may be involved.
6. Steric clashes at helix-helix interfaces and restrictions of side-chain rotamers can be expected to act against helix interactions.

Below, we argue that, although the general energy for helix interactions may involve any or perhaps all of the above, specific associations giving rise to unique folded structures are more likely to be a property of prosthetic-group interactions and detailed steric fits of surfaces.

General Factors in Helix Interactions

A clear feature of the known structures, both of porins and of helical proteins, is a cylindrical band of hydrophobic surface that surrounds each protein (Figures 3 and 4). The bands have dimensions that approximately match the hydrophobic dimension of a lipid bilayer, and, where lipids or detergents have been structurally visualized, they interact with these regions (Figure 9). That the same features are found in the subunits of complexes such as the reaction centers and the cytochrome oxidase is as expected if the subunits are individually folded in the membrane and subsequently assemble into complexes, and all known cases follow this pattern. In the two-stage model (4), the concept of matching bilayer properties is taken one step further, to the level of individual helices. In the following, we examine the

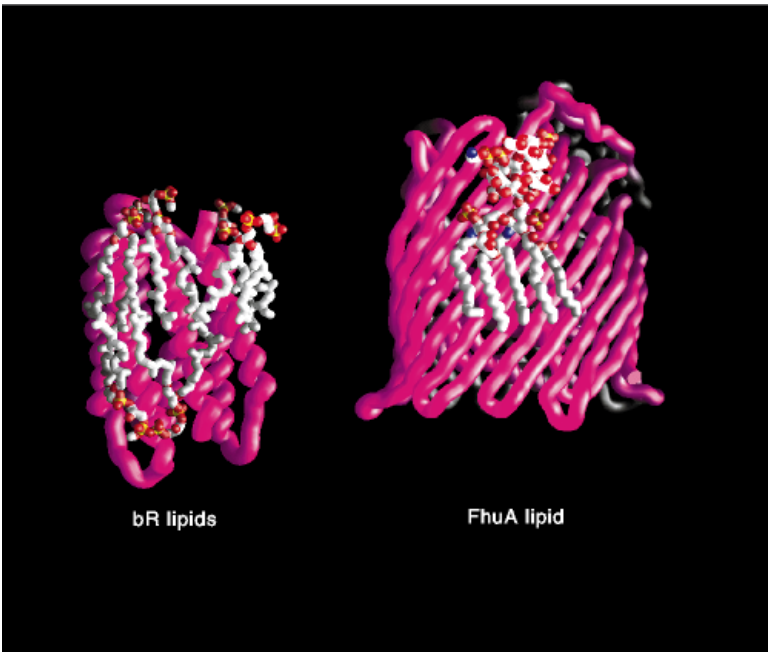

Figure 9 Lipids strongly bound by membrane proteins. Examples of lipids that are seen in defined positions in X-ray structures are illustrated, one for a helical protein [bacteriorhodopsin (173)], and one for a β-barrel protein [FhuA (79)].

idea that membrane protein folding and oligomerization can be understood in part as the side-to-side association of domains that are hydrophobically matched to the bilayer.

Helices can form right- or left-handed pairs, dictated by the spacing of side chains at their surfaces (69–71). Left-handed pairing may be a simple crossing or the helices may wrap around each other to form a coiled coil. Left-handed supercoiling creates a contact between the helices at every 3.5 amino acids, resulting in a repeat every 7 amino acids (the heptad repeat″). In a right-handed crossing of helices, the effect is in the opposite direction; coiled coils are rare, and the contacts occur every 3.9–4.0 amino acids. In both cases the basic 3.6-amino-acid-per-turn motif and the right-handedness of each individual α-helix are maintained. In membrane proteins, both kinds of helix pairing are found (e.g. bacteriorhodopsin contains predominantly left-handed contacts, whereas glycophorin dimers are right-handed; cf Figure 3). Right-handed bundles appear to be frequent in channels (see below) and in ion pumps and transporters (71a–c). The supercoiling of left-handed helix pairs facilitates close helix contacts across the membrane, on the other hand, right-handed contacts facilitate interaction of straight helices at a larger crossing angle, allowing the formation of the funnel-shaped folds seen

in the potassium-specific (99) and mechanosensitive (189) ion channels, in aquaporins (105–107), and in the nicotinic acetylcholine receptor channel (45, 45a).

Folding Pathways In Vivo

A discussion of the insertion and folding pathways is beyond the scope of this article, but it should be recognized that the ideas we present suggest pathways in which helices enter the bilayer singly or in small groups. This is compatible with recent views that helices exit the translocon one or a few at a time (71d and references therein), and with translocon-independent insertion of helical hairpins (3).

The Roles of the Lipids

When helices interact, there are changes in lipid-lipid and lipid-helix interactions as well as those between the helices themselves, and each set of interactions has a number of possible energy terms. The lipid can be thought of as having two kinds of roles—general and specific.

As a general solvent, the lipids promote folding, but may not be needed to maintain the folded structure. Many membrane proteins can maintain their structure when precipitated from detergent in the absence of added lipid. The situation is reminiscent of lyophilized soluble proteins: Water is essential for folding, but folded proteins can be stable in its near-absence. If the lipids were generally responsible for stability, it should be difficult to stabilize membrane proteins in detergents or to reconstitute them in non-native lipid environments, but often it is not, and a bewildering variety of solvent environments can maintain structure in some cases (72–77). A general lipid effect favoring helix association, if it were dominant, might be expected to produce aggregated rafts of randomly associated helices. Further, many membrane proteins change lipid environments during cytoplasmic transport between compartments. Thus, the specific properties of lipids, such as internal pressure or surface charge density, may not be generally essential for stability.

On the other hand, examples of specific associations of individual lipids or of classes of lipids with certain proteins are rapidly accumulating, and some proteins are very particular about the lipidic company they keep. Often proteins will stand one detergent but not another and may require specific lipids to be present, and it is a common observation that complete delipidation is detrimental. It may be that specific lipids can stabilize a protein through specific interactions or shift equilibria between conformational states. In this respect, lipids may act as prosthetic groups or regulatory ligands, and examples of detailed complex formation have been found (44, 78, 79; C Hunte, C Lange J Keojke & H Michel, In press; Figure 9). Hydrophobic prosthetic groups, such as chlorophyll, carotenoids, or retinal, are known to influence folding, oligomerization, and stability of such proteins as the LHC-II or bacteriorhodopsin, and some lipids do the same. The mere fact that bound lipids may favor one protein structure over another should serve as a warning when attempting to predict the most stable structure of, say, a bundle of transmembrane helices, by using calculations in a simplified lipid environment or

in vacuo. The problem is serious given that current model-building approaches, including molecular-dynamics calculations, are unable to exhaustively explore the variety of possible molecular arrangements between helices and lipids (for a review, see 12).

Loops Connecting Helices

Between the helices of polytopic proteins, loops are defined that range in size from a few amino acids to very large domains. What contributions do these parts of the structure make to stability and folding? Many proteins will assemble from fragments to form functional entities, so the entropic constraint is not always essential (see below). Calorimetric measurements show that the loops can add to the stability of bacteriorhodopsin, contributing stabilization energies comparable with the binding of retinal (80). It is also possible that loops can destabilize the protein, as when the folding of a loop is potentiated by holding its ends in defined locations or domains are forced together during oligomerization. The loops could act either to promote or to resist folding events that bring transmembrane helices together.

Some of the key observations that led to the two-stage model concerned the sequential folding of fragments of bacteriorhodopsin, first to form helices as separate fragments in detergent or lipid and then to assemble to give a functional protein (72, 73), which suggested that the fragments were acting as domains do in soluble proteins. This kind of observation was subsequently extended to a large variety of polytopic proteins under many conditions (Table 3). Except for bacteriorhodopsin, whose split, reassembled form has been studied crystallographically (81–83), characterization of the fragment complexes has been limited to observing the extent of functional recovery, which is often partial, and/or physical association. In many cases, helix-helix recognition suffices to form a folded structure, but one in which stability or functionality may be partially compromised.

Specific Factors in Helix Interactions

The dimer of the helices of glycophorin A is a known structure to which mutagenesis, modeling, energy measurements, and spectroscopic tools have been applied, and this dimer provides an opportunity to examine the interplay of structure and energy in helix interactions. The two helices form a closely packed interface that creates a right-handed dimer that is apparently stabilized by Van der Waals interactions, with little displacement of side-chain rotamers from preferred positions (60; Figure 10). Key interfacial residues include G, forming the motif LIxxGVxxGVxxT. Most point mutations in the interface destabilize the dimer, as seen in altered behavior in sodium dodecyl sulfate gel assays of dimerization (84). Based on the structure, these effects can be interpreted as arising from altered Van der Waals interactions and rotamer restrictions. A quantitative, first-order calculation has been developed from the interpretation, and it appears to have predictive power (55). Analytical ultracentrifugation and small-angle X-ray scattering give

TABLE 3 Integral membrane proteins assembled from α-helical fragments that had been either refolded or synthesized independently

Origin of fragments	Medium where assembly takes place	Protein	N[a]	Transmembrane helices per fragment[b]	References
Proteolysis	Mixed micelles 72, 206	Bacteriorhodopsin	7	2 + 5	72, 206
				5 + 2	118, 207
				5 + 5*	118
Proteolysis	Lipid vesicles	Bacteriorhodopsin	7	2 + 5, 5 + 2	73, 80, 81
Proteolysis; chemical synthesis				1 + 1 + 5	80
Plasmids	Yeast	Bacteriorhodopsin	7	2 + 5	208
Engineered & purified	Mixed micelles	Bacteriorhodopsin	7	2 + 5, 3 + 4, 4 + 3, 5 + 2	119
				3 + 5*, 4 + 4*, 5 + 3*	119
				4 + 5*, 5 + 4*, 5 + 5*	119
cRNA	*Xenopus* oocyte	β_2-Adrenergic receptor	7	5 + 2	209
Plasmids	COS-7 cells	m2- & m3-muscarinic ACh receptors	7	5 + 2	210
Plasmids	COS-1 cells	Rhodopsin	7	3 + 4, 4 + 3, 5 + 2, 3 + 2 + 2	211, 212
Plasmids	COS-7 cells	NK1 receptor	7	5 + 2	204
Plasmids	Yeast	α-Factor STE2 receptor	7	1 + 6, 2 + 5, 3 + 4	120
				4 + 3, 2 + 5, 1 + 6	120
Plasmids	*E. coli*	*lac* permease	12	2 + 10	213–215
				3 + 9, 6 + 6, 7 + 5	214–216
				3 + 10*, 2 + 11*	214
				4 + (2Δ8)*	214

cRNA	*Xenopus* oocyte/ in vitro microsomes	Anion exchanger (band 3)	14	8 + 6, 12 + 2 5 + 9, 7 + 7 1 + 13, 3 + 11 5 + 7 + 2 etc., 5 + Δ + 7	217 218 219 121
cRNA	*Xenopus* oocyte	Voltage-gated chloride channel	10/12	8 + 2/4	203
Plasmids	Yeast ER	Sec61 protein	10	1 + 9, 2 + 8 etc.	220
Natural	Thylakoid	Cytochromes *b* vs b_6	8	4 + 3 + Δ	221
Natural	*E. coli*	Nicotinamide nucleotide transhydrogenase	14	4 + Δ + 9	222
Natural	Mitochondria	Cytochrome *c* oxidase I + III	19	14 + 5, 12 + 7	132
Natural	Thylakoid	PSI vs PSII reaction centers	11	5 + 6	141, 145, 147

[a]Number of transmembrane helices in each fragment.

[b]Number of transmembrane helices in each fragment, in the order of the fragments in the sequence. Δ, One or more transmembrane helices are missing; *, the complex of fragments contains one or more redundant helices. All of the proteins listed are thought to have a transmembrane region entirely composed of largely hydrophobic α-helices. In addition to those listed, a number of other complementation experiments have been reported in which some of the putative transmembrane segments are redundant (see in particular Reference 202), as well as experiments on integral proteins containing internal repeats (reviewed in References 116, 122, 203, 204), and an experiment involving the reassociation of fragments of a β-barrel protein (205). Cases listed as natural correspond to integral proteins that are made up of one polypeptide chain in one type of membrane and two complementary, nonhomologous transmembrane subunits in another (see text).

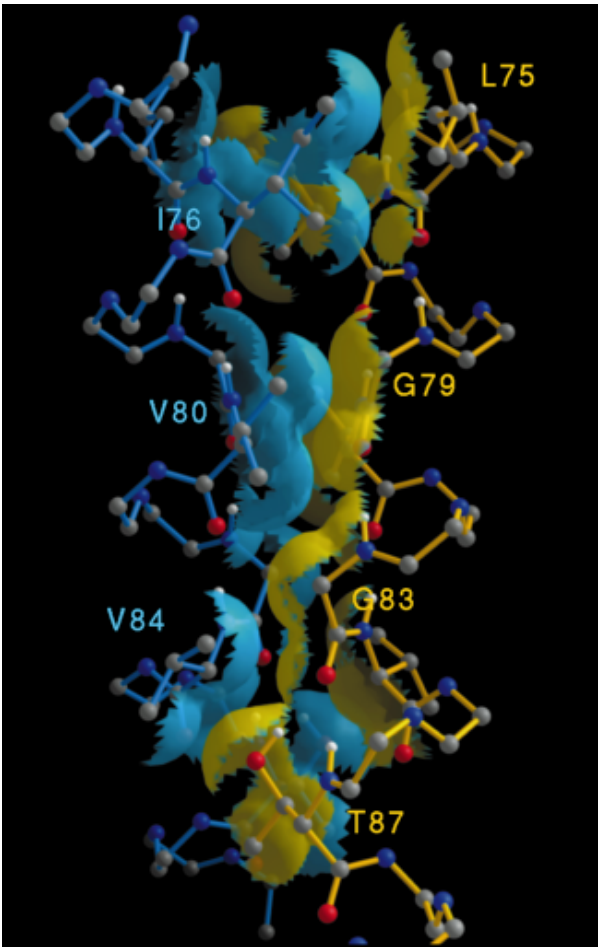

Figure 10 Glycophorin helix dimer interface. The human glycophorin A transmembrane helices dimerize in lipidic environments by coordinating close Van der Waals contacts (60). Note the excellent fit between the surfaces. In the two-stage model, the individual helices are proposed to be stable, forming the surfaces before contact is made. Thus, interactions are coordinated over the surfaces when the helices interact, whereas such coordination of contacts will not take place between the more fluid lipid and the helix surfaces.

measures of thermodynamic variables, which are in fair agreement with the calculated affinities (85, 86). Further tests with multiple substitutions should lead to refinements in the representation and prediction of helix interactions.

For this discussion, the main point is that close packing of helix surfaces inside the bilayer stabilizes the glycophorin dimer (Figure 10). This result might seem surprising, except that the surfaces are largely formed by the separate helices as individually stable entities, resulting in preformed surfaces that allow Van der Waals

energies to be coordinated among all points of contact. By contrast, coordination of contacts would be absent in the interfaces with the more fluid bilayer or micelle. The consequence is that good Van der Waals contacts involve a low entropy penalty of association at the helix-helix interface, whereas they do involve such a penalty at the helix-lipid interface. Thus, in this case the concept of independently stable helices explains a key feature of their assembly in a higher-order structure.

If we take an overview of the interactions that are known to be involved in the helix-helix contacts, setting aside prosthetic group and lipid interactions, we arrive at the following views based on current evidence. A dominant factor is the way that the helices fit together, guided by Van der Waals interactions and side-chain rotamers. Although interhelical hydrogen bonds had been anticipated as a source of stability, not very many have been seen so far (see e.g. 56). They are known to occur, however, in the outer membrane of *E. coli*, in oligomer formation from beta barrels (86a). Likewise, ion pairs are not a dominant theme to date, the only ones found so far being in the LHC-II (96). No disulfides have been found in the bilayer region, even though C is abundant. What is the message behind these observations? An intriguing possibility is that strong, local interactions are not favored for biological reasons, perhaps including turnover and/or function.

The danger presented by the strength of hydrogen bonds or ion pairs between helices is that they can overwhelm all other interactions, potentially leading to nonspecific aggregation or misfolding (86b–d and references therein). A well known example of such a perturbation is that of the neu oncogene (86e,f). Yet, strongly polar groups are required in helices for functional reasons (e.g. 35, 96, 192b). Thus, it may be that folding pathways and/or the presence of membrane chaperones prevent exposure of polar groups to lipids, as aqueous chaperones shield the exposure of hydrophobic groups to water. Once formed, internal H bonds or ion pairs would add significantly to the stability of helix association. Outside of the membrane, a number of proteins are known to play a role in assisting membrane protein folding (86g and references therein).

Helix-Lipid vs Helix-Helix Interactions

Early experimental data on bacteriorhodopsin and photosynthetic reaction centers (87, 88) led to a suggestion that transmembrane helix bundles might be "inside-out" as compared with soluble proteins, in the sense that their outside (membrane-exposed) surface is more hydrophobic than their core. Recent sequence analyses have considered the propensity of the various classes of residues for contacts with lipids vs other helices (50). By a mathematical procedure first developed by Cornette et al (89), contact propensity was analyzed by purely statistical methods, involving no a priori assumptions about the physicochemical properties of the residues studied. This method resulted in a scale that characterizes the propensity of the α-carbon of various types of residues to lie on either one or the other face in a large database of putative helices. An interesting outcome of these analyses is that residue distribution is not dominated by a single physicochemical parameter,

such as hydrophobicity (50). Furthermore, there is no universal scale that correctly describes the distribution of residues in any set of transmembrane helices; whereas large aliphatic residues (L, I, V) tend to be located on the same helix face, presumed to face the lipids, other residues, e.g. F or C, may be preferentially located on the same face as L, I, and V in one bank of segments and on the opposite face in another (N Ugolin, J-L Popot, C Etchebest, submitted for publication). Helix-specific constraints related to function, local environment, and perhaps biosynthetic constraints appear more important in determining residue distribution than the physicochemical properties of the side chains. Thus, a simple, unique rule is unlikely to correctly describe the variety of helix-helix packing in transmembrane bundles. A practical consequence is that attempts at predicting helix orientations based on lateral amphipathy alone are not very reliable (N Ugolin, J-L Popot, C Etchebest, submitted for publication).

Functional Perspectives

The problem of creating functionality with chemically limited helices has been mentioned above, and it is overcome in part by the inclusion of polar and prosthetic groups, and perhaps by π, 3_{10}, or P-bend structures. Additionally, the helices may change their relationships to one another through translations, rotations, and changes in relative tilt. In the chemotaxis receptor (Tar), cross-linking data are best explained by a translation of one helix relative to the other (92). The functional cycles of bacteriorhodopsin (see 93 and references therein) and lactose permease (see 5), involve relative helix motions. From the earliest consideration of helix packing (69), it has been seen that there are many local minima in the relationships of a pair of helices produced by rotations about their long axes. For membrane helices, both right- and left-handed crossing angles are seen. Global computational-chemistry searches for favorable helix interactions in glycophorin and phospholamban reveal a number of local minima that are near one another in energy, although it is not known what the energy barriers are for moving from one to another (94, 95). Taken as a whole, it seems that the data are beginning to indicate a role for alternative states of helix interactions in function. If this is the case, it may help to explain the relative scarcity of strong local hydrogen or ion pair bonds between helices, which would restrict such alternatives.

In some membrane proteins, such as the photosynthetic reaction centers and bacteriorhodopsin, the close packing among helices gives an interior with the density of soluble proteins, but in other cases it does not. In the LHC-II and bc_1 complexes (96, 97), the packing is much looser, with lipid and prosthetic groups mixed among helices. The constraint of being in a bilayer means that the solvent volume containing interacting helices is small and the helices are stabilized as separate entities partially oriented relative to each other. Under these conditions, small energies can produce association, and it may be possible to have protein regions with less defined structures that are loosely held together and have many alternative helix interactions. Such structures might be useful, for example, in the escape of

transmembrane helices from the secretion apparatus of the endoplasmic reticulum into the surrounding lipids (98). The state would be analogous to the molten globule observed in the folding of some soluble proteins, in which secondary structure has formed and the structure is compact, but detailed folding has not occurred.

Escape from Bilayer Constraints

One of the challenges faced by a cell is to create polar regions that traverse the permeability barrier of the bilayer, for example in ion channels, gap junctions, and porins. The recent structure of a K^+ channel (99) and the earlier structures of porins (e.g. 100, 101) and siderophore receptors (79, 102, 103) reveal a recurrent feature of membrane protein design—that structures can be devised by using β-barrels or rings of helices to build barriers that sequester regions from the bilayer to permit a wider scope of structure. In the porins, the wall of β-strands allows a loop of polypeptide to form inside it, avoiding the constraints of bilayer contact. In siderophore receptors, a whole folded domain occupies the lumen of the barrel. In the K^+ channel, a central region enables short helices to place their ends near the center of the bilayer, where their polar character helps to lower the energy barrier for transmembrane K^+ movement. Gap junctions (104), aquaporins (105–107), and the nicotinic acetylcholine receptor (45) use helical arrays to create aqueous channels.

Summary of Helix Interactions

1. Helices make closely packed right- and left-handed contacts in helical membrane proteins.
2. Packing is likely to provide the key to specific helix-helix recognition and contributes to stabilization as well.
3. Extramembrane links between helices may enhance stability, but many are not required for folding and function.
4. Weak hydrogen bonding between helices does occur, including via water molecules.
5. Strong hydrogen bonds or ion pairs may occur, but create a danger of nonspecific aggregation or misfolding.
6. The stability of individual helices facilitates their interaction in a membrane.

EXPERIMENTAL REARRANGEMENT OF STRUCTURE

Deletion or Flipping of Helices

Gene fusion experiments have been invaluable in sorting out the factors that determine the transmembrane topology of integral proteins (see e.g. 108). Five major conclusions can be drawn:

1. As a rule, deleting a C-terminal fragment from an integral protein and replacing it with a reporter protein does not affect the topology adopted by the remaining N-terminal region when the fusion protein is expressed.
2. Vectoriality signals are distributed along the sequence; transmembrane orientation and topology are not determined solely by the vectoriality of insertion of the first transmembrane segment, followed blindly by insertion in alternating orientations of the hydrophobic segments downstream (see e.g. 109, 110).
3. Transmembrane helices per se apparently contain little or no information about their orientation in the membrane, vectoriality being primarily determined by the distribution of positively charged residues in the vicinity of helix ends (111). Genetic manipulation of the distribution of lysine and arginine residues can cause a transmembrane segment to insert with inverted vectoriality (see e.g. 109, 112–114, and references therein).
4. Depending on the extramembrane loops considered, vectoriality signals may be strong or weak; when deletions are introduced such that the orientation of the N-terminal region of MalF will be inverted, some of the downstream segments may follow suit and insert with an inside-out orientation, but the C-terminal region contains strong vectoriality signals and will not invert (109, 114). Similar observations have been made on *lac* permease (115).
5. Whether a given segment will adopt a stable transmembrane position does not depend only on its hydrophobicity, but also on its environment (see 40, 108). This is in keeping with the observation that the hydrophobicity distributions of transmembrane and extramembrane segments overlap (see above).

Together with the study of split proteins, these experiments indicate that insertion of integral protein fragments with their proper transmembrane topology can occur in the absence of the rest of the protein. They show, furthermore, that both insertion and topology can be manipulated by appropriate mutations, primarily by altering the distribution of basic residues in extramembrane segments.

Creating Redundant Helices; Deleting Internal Helices

There are now many reports of experiments in which extranumerary segments, presumed to form transmembrane helices, have been introduced into a protein without abolishing its function (reviewed in 116). This was generally achieved by reassembling fragments comprising overlapping sequence regions (bacteriorhodopsin, *lac* permease, α-factor receptor; cf Table 3), but can also be observed after internal duplication of a segment comprising several putative helices [FhuB protein (117)]. In some cases, it is certain that the redundant helices were part of the final structure (e.g. 118, 119), whereas in others they may be trimmed after fragment assembly (120).

Conversely, there is at least one case in which function can be recovered by reassociating fragments that do not fully complement: stilbene disuphonate-sensitive anion transport is exhibited by complexes of N-terminal and C-terminal fragments of band 3 that, taken together, lack two of the internal helices (121; Table 3).

EVOLUTION

One can easily imagine that events such as those just described, were they to occur naturally, might result in a protein with extra transmembrane α-helices. Insofar as they do not interfere with folding of the functional part of the protein, these adulterations may well be tolerated. This could permit, for instance, the formation of redundant binding sites for lipophilic or hydrophilic ligands, which would serve as a starting point for subsequent evolution. Less extensive modifications, such as fusing several polytopic subunits into a single one or splitting one into several, may be neutral or may provide a basis for subsequent evolution of biosynthetic or regulatory processes. Below we describe a few cases in which evolution appears to have taken advantage of the inherent stability of transmembrane helices to reorganize transmembrane regions (further examples have been reviewed in 13, 116, 122).

Naturally Fused (or Split) Integral Membrane Proteins

Internal repeats within the sequence of some integral membrane protein families are a testimony to gene duplication and fusion. In several cases, the basic repeat or module can be found as an independent subunit in another membrane or organism (see e.g. 116, 123, and references therein). Such is the case in the bacterial plasma membrane ATP-driven carrier family. Most of these carriers comprise two polytopic subunits (MalF and MalG in the maltose system) associated with a cytosolic ATP-binding protein (reviewed in 124). Evidence has been presented for a distant homology between the two integral proteins (125, 126). Similar situations are encountered e.g. in the FhuB protein, in the tetracycline resistance (Tet) protein, and in voltage-gated channels, all of which can be manipulated by separating modules or splicing them one to another (see e.g. 117, 127–130).

It is more interesting that polypeptides that lack internal repeats may exist in one organism or cellular compartment as a single polypeptide and in another as several separately encoded subunits (Table 3, lower part). A particularly spectacular case is that of the quinol oxidase from *Sulfolobus acidocaldarius*, whose single apoprotein (131) contains three regions, respectively, that are homologous to each of the three separate polytopic subunits of mitochondrial cytochrome *c* oxidases (132, 133).

Additional helices at either end of the sequence of an integral membrane protein are not an exceptional feature: witness the case of cytochromes b/b_6 (Table 3) or that of cytochrome oxidase, whose subunit I contains 12 transmembrane helices in *Paracoccus denitrificans* and in mitochondria (44, 134) and is thought to comprise 14 transmembrane helices in the thermophilic bacterium *Bacillus* PS3 (135) and

in *B. subtilis* (136; with two extra C-terminal helices) and 15 in the cytochrome *o* terminal oxidase complex of *Escherichia coli* [one extra N-terminal and two extra C-terminal helices (137)]. There are many other cases, for instance in the family of ATP-driven bacterial carriers, in which related proteins are proposed to present partially different topologies.

Another interesting case of splitting or fusion of polytopic proteins is emerging in the field of photosynthesis. Oxygen-evolving organisms harbor two types of reaction centers (RCs), respectively associated with photosystem I (PSI) and photosystem II (PSII). The core of each type of RC is made up of a dimer of closely related proteins, D1/D2 in PSII (five transmembrane helices each, homologous to the L and M subunits of purple bacteria RCs) and PsaA/PsaB in PSI (11 transmembrane helices each). The two types of dimers differ in many respects: the nature of the prosthetic groups involved in electron transfer, the sizes of the proteins, and the fact that numerous molecules of chlorophyll (called the "core antenna") are carried by PsaA/B in PSI, and by distinct subunits (called CP43 and CP47) in PSII. There are no obvious sequence similarities between PsaA/B and D1/D2. Functional considerations nevertheless led to the proposal that both types of centers originated from a common ancestor (138, 139; reviewed in 140). Weak similarities in the sequences of putative transmembrane helices were proposed as evidence that the N-terminal region of PsaA/B is distantly related to CP43/CP47 (141).

Recent structural data on the two RCs appear to support these views: (*a*) In the 4-Å map of the PSI RC from the cyanobacterium *Synechococcus elongatus*, the arrangement of the last five transmembrane helices of PsaA and PsaB and of the pigments bound to them is reminiscent of that of the L and M helices and the associated pigments in purple bacteria RCs (142–144); (*b*) according to recent medium-resolution three-dimensional electron microscopic data on a subcomplex of the PSII RC (145), the D1/D2 dimer appears very similar to the L/M dimer in purple bacteria RCs, as anticipated (146), whereas the arrangement of the six transmembrane helices in CP47 resembles that of the six N-terminal antenna chlorophyll-binding helices in PsaA/B. PsaA/B therefore appear to combine into single-polypeptide helices that bind the electron transfer prosthetic groups—distant homologs of D1/D2—and helices that carry the inner antenna chlorophylls, homologous to CP43/CP47. This conclusion has been further reinforced by recent electron microscopy projection data obtained on a PSII complex containing both CP43 and CP47 (147). Comparison of the PSI and PSII three-dimensional structures suggests that the two blocks of helices have moved compared with each other, in the course of evolution (145). It is not clear, from evolutionary data, whether a fusion or a splitting is more likely to have generated the two different primary structures (140).

Removing a single helix from the middle of the sequence may become more likely once the protein is split, as seen for nicotinamide nucleotide transhydrogenase (Table 3). We are not aware of cases in which it is clear that helices have been added to (or removed from) the middle of a full-length protein in the course of evolution. Experiments resulting in an internal helix duplication in the FhuB

protein (117) or in a functional anion exchanger lacking two internal helices (121), however, indicate that such events should not be ruled out a priori.

Helix Shuffling

Successive fusion and scission events may lead to the transfer of helices from one subunit to another. [It is worth noting, in this context, that an early analysis, which deserves to be expanded now that much more structural and sequence data are available, concluded that introns present in the structural genes encoding polytopic proteins tend to be located in regions that code for extramembrane loops (148)]. Given the limited set of residues that transmembrane segments contain, reconstructing helix shuffling between integral membrane proteins by sequence comparison may be difficult. One example does exist in the cytochrome oxidase family (see 122 and references therein).

Helix Flipping

Intuition suggests that biosynthetic constraints and functional requirements will make it both difficult and potentially counterproductive for a protein or fragment thereof to insert into a membrane with inverted vectoriality. As noted above, experimental manipulation of the distribution of basic residues has shown that the vectoriality of the insertion of a given protein can be inverted and suggests that mismatches between the orientations of the N-terminal and C-terminal ends of a protein can be created. It is therefore not impossible a priori that, for instance, duplication of a protein region that does not contain strong vectoriality signals may result in the upside-down insertion of the duplicated part.

Such a mechanism has been postulated to account for the internal homology of the lens fiber major intrinsic protein (MIP) family (149, 150), whose six transmembrane helices (105, 106) appear to have arisen by duplication of a 3-helix module (123). This proposal is consistent with the three-dimensional structure of aquaporins (105–107). An interesting twist to this problem is that, as a rule, misfolded eukaryotic membrane proteins never make it from the endoplasmic reticulum to the Golgi complex (reviewed in 86f, 116, 151, 152). One would suppose—although we are not aware of any such demonstration—that an upside-down protein stands little chance of escaping this quality control process. Presumably, it would be recognized as misfolded and degraded before ever reaching the plasma membrane, leaving nothing for evolution to act on. Note that, in the phylogenetic tree proposed for the MIP family, duplication actually predates the divergence between eubacteria and eukaryotes (149).

Intrahelical Stagger

Throughout this discussion, we have considered helices as blocks within which insertions or deletions are forbidden. We have seen above, however, how a transmembrane helix can accommodate the insertion of an extra residue by forming a π-bulge (Figure 5). Another interesting case of insertion is to be found in the

homologous *b* and b_6 cytochromes. Two hemes are bound by cytochrome *b* or b_6, each heme iron receiving as its fifth and sixth ligands nitrogen atoms from two histidine residues, one from transmembrane helix 2 and the other from transmembrane helix 4 (see EA Berry & AR Crofts, this volume). In helices 2 and 4 of cytochrome *b*, as well as in helix 2 of cytochrome b_6, the two histidine residues are separated by 13 residues (about four helix turns). In helix 4 of cytochrome b_6, however, the two histidines are separated by 14 residues (154). Recent spectroscopic data suggest that this insertion of one residue is accommodated by a local distortion without stagger and does not lead to a change in the geometry of heme liganding (155).

Evolution Summary

Taken together, therefore, sequence data and structural data strongly indicate that, in the course of evolution, helices or bundles of helices may be split, fused, and recombined in various manners, an observation that fits with the idea that adopting a correct transmembrane topology does not (or does not systematically) depend on association with other transmembrane segments.

Domain shuffling represents a major feature in the evolution of soluble proteins or extramembrane protein regions. It may also have played an important role in the making of α-helical transmembrane regions. These regions, however, can be conceived of as having a two-tiered structural organization, which results from the independent stability exhibited by individual transmembrane helices. Domains that are shuffled around can correspond either to helix bundles with a defined geometry or to fragments thereof, down to individual helices. It may seem difficult to remove helices from an existing bundle without abolishing functionality; nevertheless, we have seen several examples of such events, resulting in the loss of (presumably outlying) transmembrane helices. Newly gained helices, on the other hand, may be retained even though initially they may not pack well with the rest of the transmembrane region.

To end on a more speculative note, it is provocative to think that, very early in evolution, miniproteins made up of a single stretch of hydrophobic residues may have played a role as prototypes of present-day integral proteins. Many reports describe the formation of ion channels by synthetic peptides that are thought to organize into hollow bundles of transmembrane α-helices (e.g. 156, 157), and such seems to be the structure of present-day phospholamban (158–160). Hydrophobic peptides carrying a metal-binding residue like histidine into a membrane may have represented early forms of heme- or chlorophyll-binding proteins. Today's light-harvesting antennae of purple and green bacteria consist of pigments bound to small one-helix polypeptides (161, 162). Cytochrome b_{559}, whose proposed structure comprises two such peptides liganding a heme (163), may perhaps be considered a contemporary illustration of what early membrane-buried cytochromes may have looked like. The remarkable ability of short hydrophobic peptides to adopt a stable α-helical conformation on insertion into a lipid bilayer may lie at the root of

the chance appearance and subsequent evolution of functional helix bundles and polytopic proteins.

SUMMARY OF MAIN POINTS

1. Transmembrane helices with largely hydrophobic side chains are a major theme in membrane proteins.
2. The idea that single helices are often stable transmembrane domains is consistent with a considerable body of observations.
3. The energy of the hydrophobic effect for some of the side chains in a helix can be used to drive the membrane insertion of unfavored β-branched or strongly polar side chains.
4. Side-to-side helix association is a major theme in the stability of many membrane proteins and complexes.
5. Interactions favoring specific helix associations include the coordination of favorable Van der Waals contacts over complementary surfaces and the binding of prosthetic groups.
6. Single, strong interactions through hydrogen bonds, covalent links, or ion pairs are seldom seen between helices.
7. Interactions opposing helix associations include side-chain rotamer restrictions and steric clashes.
8. Variations in helix structure can include proline-induced bends, π-helix or 3_{10} helix.
9. Variation of helix-helix interactions may be a theme in functional states.
10. Arrays of helices may create interior spaces that allow inclusion of more polar entities, such as channels, gating charges, or ions.
11. The two-stage model leads to concepts that may be helpful in rationalizing membrane protein evolution.

ACKNOWLEDGMENTS

We thank Jean-Pierre Duneau, Kevin MacKenzie, Christian Ostermeier, and Daniel Picot for help with the figures; C Breyton, Richard Cogdell, Philippe Devaux, Mark Dumont, Lars-Oliver Essen, Jonathan Groves, Werner Kühlbrandt, Kevin MacKenzie, Hartmut Michel, Wolfgang Nitschke, Dieter Oesterhelt, and Wolfram Welte for discussions and/or communication of unpublished information; and Edith Gardet and Janice Murphy for help with the preparation of the manuscript. Our collaboration was made possible by the award of a Chaire de Recherche Blaise Pascal of the Région Ile-de-France to D. M. E. and by support from the Centre National de la Recherche Scientifique, Paris, France; National Institutes of Health,

Bethesda, Maryland; and National Science Foundation, Washington, DC. Thanks are due to the Fondation de l'Ecole Normale Supérieure for administrating the Chaire Blaise Pascal contract.

LITERATURE CITED

1. Boyd D, Schierle C, Beckwith J. 1998. *Protein Sci.* 7:201–5
2. Wallin E, von Heijne G. 1998. *Protein Sci.* 7:1029–38
3. Engelman DM, Steitz TA. 1981. *Cell* 23: 411–22
4. Popot J-L, Engelman DM. 1990. *Biochemistry* 29:4031–37
5. Kaback HR, Wu J. 1997. *Q. Rev. Biophys.* 30:333–64
6. Kaback HR, Voss J, Wu J. 1997. *Curr. Opin. Struct. Biol.* 7:537–42
7. Booth PJ. 1997. *Fold. Des.* 2:R85–R92
8. Fillingame RH, Jones PC, Jiang W, Valiyaveetil FI, Dmitriev OY. 1998. *Biochim. Biophys. Acta* 1365:135–42
9. Sakai H, Tsukihara T. 1998. *J. Biochem.* 124:1051–59
10. Booth PJ, Curran AR. 1999. *Curr. Opin. Struct. Biol.* 9:115–21
11. White SH, Wimley WC. 1999. *Annu. Rev. Biophys. Biomol. Struct.* 28:319–65
12. Biggin PC, Sansom MS. 1999. *Biophys. Chem.* 76:161–83
13. Popot J-L. 1993. *Curr. Opin. Struct. Biol.* 3:532–40
14. Wetlaufer DB. 1973. *Proc. Natl. Acad. Sci. USA* 70:697–701
15. Kyte J, Doolittle RF. 1982. *J. Mol. Biol.* 157:105–32
16. Engelman DM, Steitz TA, Goldman A. 1986. *Annu. Rev. Biophys. Biophys. Chem.* 15:321–53
17. Engelman DM, Goldman A, Steitz TA. 1982. *Methods Enzyzmol.* 88:81–88
18. Ahmed S, Brown D, London E. 1997. *Biochemistry* 36:10944–53
19. Harder T, Scheiffele P, Verkade P, Simons K. 1998. *J. Cell Biol.* 141:929–42
20. Harder T, Simons K. 1998. *Curr. Opin. Cell Biol.* 9:534–42
21. Lewis BA, Engelman DM. 1983. *J. Mol. Biol.* 166:211–17
22. Hallett FR, Marsh J, Nickel BG, Wood JM. 1993. *Biophys. J.* 64:435–42
23. Marsh D. 1996. *Biochim. Biophys. Acta* 1286:183–223

23a. Curran AR, Tenyler RH, Booth PJ. 1999. *Biochemistry* 38:9328–36

23b. Cantor RS. 1997. *Biochemistry* 36:2339–44

24. Wiener MC, White SH. 1992. *Biophys. J.* 61:437–47
25. Sabra MC, Uitdehaag CM, Watts A. 1998. *Biophys. J.* 75:1180–88
26. Essen L, Siegert R, Lehmann WD, Oesterhelt D. 1998. *Proc. Natl. Acad. Sci. USA* 95:11673–78
27. Engelman DM. 1969. *Nature* 233:1279–80
28. Mouritsen OG, Bloom M. 1984. *Biophys. J.* 46:141–53
29. Reiss-Husson F, Luzzati V. 1967. *Adv. Biol. Med. Phys.* 11:87–107
30. Tardieu A, Luzzati V, Reman FC. 1973. *J. Mol. Biol.* 75:711–33
31. Steim JM, Reinert JC, Tourtellotte ME, McElhaney RN, Rader RL. 1969. *Proc. Natl. Acad. Sci. USA* 63:109–13
32. Engelman DM. 1970. *J. Mol. Biol.* 47:115–17
33. de Kruijff B. 1997. *Curr. Opin. Chem. Biol.* 1:564–69
34. Henderson R, Unwin PNT. 1975. *Nature* 257:28–32
35. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH. 1990. *J. Mol. Biol.* 213:899–929

36. Deisenhofer J, Epp O, Miki K, Huber R, Michel H. 1985. *Nature* 318:618–24
37. White SH. 1994. *Membrane Protein Structure: Experimental Approaches*, ed. SH White, pp. 97–123. New York: Oxford Univ. Press
38. Etchebest C, Popot J-L. 1997. *Membrane Protein Assembly*, ed. G von Heijne, pp. 221–49. Austin, TX: Landes
39. Popot J-L, de Vitry C. 1990. *Annu. Rev. Biophys. Biophys. Chem.* 19:369–403
40. von Heijne G. 1996. *Membrane Protein Assembly*, ed. G von Heijne, pp. 55–62. Austin, TX: Landes
41. Hunt JF, Earnest TN, Bousche O, Kalghatgi K, Reilly K, et al. 1997. *Biochemistry* 36:15156–76
42. Luneberg J, Widmann M, Dathe M, Marti T. 1998. *J. Biol. Chem.* 273:28822–30
43. Honig BH, Hubbell WL. 1984. *Proc. Natl. Acad. Sci. USA* 81:5412–16
44. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, et al. 1996. *Science* 272:1136–44
45. Unwin N. 1993. *J. Mol. Biol.* 229:1101–24

45a. Unwin N. 1995. *Nature* 373:37–43

45b. Unwin N. 1996. *J. Mol. Biol.* 257:586–96

46. Heinz DW, Baase WA, Dahlquist FW, Matthews BW. 1993. *Nature* 361:561–64
47. Heinz DW, Baase WA, Zhang XJ, Blaber M, Dahlquist FW, Matthews BW. 1994. *J. Mol. Biol.* 236:869–86
48. Keefe LJ, Sondek J, Shortle D, Lattman EE. 1993. *Proc. Natl. Acad. Sci. USA* 90:3275–79
49. Li S-C, Deber CM. 1994. *Nat. Struct. Biol.* 1:368–73
50. Samatey FA, Xu CB, Popot J-L. 1995. *Proc. Natl. Acad. Sci. USA* 92:4577–81
51. Belrhali H, Nollert P, Royant A, Menzel C, Rosenbusch JP, et al. 1999. *Structure* 7:909–17
52. Luecke H, Schobert B, Richter H-T, Cartailler J-P, Lanyi JK. 1999. *J. Mol. Biol.* 291:899–911
53. Duneau J-P, Garnier N, Genest M. 1997. *J. Biomol. Struct. Dyn.* 15:555–72
54. Ostermeier C, Harrenga A, Ermler U, Michel H. 1997. *Proc. Natl. Acad. Sci. USA* 94:10547–53
55. MacKenzie KR, Engelman DM. 1998. *Proc. Natl. Acad. Sci. USA* 95:3583–90
56. Deisenhofer J, Epp O, Sinning I, Michel H. 1995. *J. Mol. Biol.* 246:429–57
57. Ovchinnikov YA, Abdulaev NG, Feigina MY, Kiselev AV, Lobanov NA. 1979. *FEBS Lett.* 100:219–24
58. Khorana HG, Gerber GE, Herlihy WC, Gray CP, Anderegg RJ, et al. 1979. *Proc. Natl. Acad. Sci. USA* 76:5046–50
59. Tomita M, Marchesi VT. 1975. *Proc. Natl. Acad. Sci. USA* 72:2964–68
60. MacKenzie KR, Prestegard JH, Engelman DM. 1997. *Science* 276:131–33

60a. Russ WP, Engelman DM. 2000. *J. Mol. Biol.*In press

60b. Senes A, Gerstein M, Engelman DM. 2000. *J. Mol. Biol.* 25:921–36

61. Rosenbusch JP, Garavito RM, Dorset DL, Engel A. 1982. *Protides of the Biological Fluids*, ed. E Peeters, pp. 171–74. Oxford: Pergamon
62. Pautsch A, Schulz GE. 1998. *Nat. Struct. Biol.* 5:1013–17
63. Yau W-M, Wimley WC, Gawrisch K, White SH. 1998. *Biochemistry* 37:14713–18
64. Landolt-Marticorena C, Williams KA, Deber CM, Reithmeier RAF. 1993. *J. Mol. Biol.* 229:602–8
65. Reithmeier RAF. 1995. *Curr. Opin. Struct. Biol.* 5:491–500
66. von Heijne G. 1994. *Membrane Protein Structure: Experimental Approaches*, ed. SH White, pp. 27–40. Oxford: Oxford Univ. Press
67. Hol WGJ. 1985. *Prog. Biophys. Mol. Biol.* 45:149–95
68. Ben-Tal N, Honig B. 1996. *Biophys. J.* 71:3046–50
69. Crick F. 1953. *Acta Crystallogr.* 6:689–97

70. Chothia C, Levitt M, Richardson D. 1977. *Proc. Natl. Acad. Sci. USA* 74:4130–34
71. Langosch D, Heringa J. 1998. *Proteins: Struct. Funct. Genet.* 31:1–10
71a. Zhang P, Toyoshima C, Yonekuza K, Green NM, Stokes DL. 1998. *Nature* 392:835–39
71b. Aver M, Scarborough GA, Kühlbrandt W. 1998. *Nature* 392:840–43
71c. Williams KA. 2000. *Nature* 403:112–15
71d. Johnson AE, van Waes MA. 1999. *Annu. Rev. Cell. Dev. Biol.* 15:799–842
72. Huang K-S, Bayley H, Liao M-J, London E, Khorana HG. 1981. *J. Biol. Chem.* 256:3802–9
73. Popot J-L, Gerchman S-E, Engelman DM. 1987. *J. Mol. Biol.* 198:655–76
74. Booth PJ, Farooq A. 1997. *Eur. J. Biochem.* 246:674–80
75. Sturgis JN, Robert B. 1994. *J. Mol. Biol.* 238:445–54
76. Surrey T, Schmid A, Jähnig F. 1996. *Biochemistry* 35:2283–88
76a. Tribet C, Audebert R, Popot J-L. 1996. *Proc. Natl. Acad. Sci. USA* 93:15047–50
77. Kikuchi J, Asakura T, Loach PA, Parkes-Loach PS, Shimada K, et al. 1999. *Biopolymers* 49:361–72
78. Weik M, Zaccaï G, Dencher NA, Oesterhelt D, Hauss T. 1998. *J. Mol. Biol.* 275:625–34
79. Ferguson AD, Hofmann E, Coulton JW, Diederichs K, Welte W. 1998. *Science* 282:2215–20
80. Kahn TW, Sturtevant JM, Engelman DM. 1992. *Biochemistry* 31:8829–39
81. Popot J-L, Trewhella J, Engelman DM. 1986. *EMBO J.* 5:3039–44
82. Popot J-L, Engelman DM, Gurel O, Zaccaï G. 1989. *J. Mol. Biol.* 210:829–47
83. Samatey FA, Zaccaï G, Engelman DM, Etchebest C, Popot J-L. 1994. *J. Mol. Biol.* 236:1093–104
84. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, Engelman DM. 1992. *Biochemistry* 31:12719–25
85. Fleming KG, Ackerman AL, Engelman DM. 1997. *J. Mol. Biol.* 272:266–75
86. Bu Z, Engelman DM. 1999. *Biophys. J.* 77:1064–73
86a. Snijder HJ, Ubarretxena-Belandia I, Blaauw M, Kalk KH, Verheji HM, et al. 1999. *Nature* 401:717–21
86b. Smith SO, Smith CS, Bormann BJ. 1996. *Nat. Struct. Biol.* 3:252–73
86c. Zhou FX, Cocco MJ, Russ WP, Brunger AT, Engelman DM. 2000. *Nat. Struct. Biol.* 7:154–60
86d. Choma C, Gratkowski H, Lear JD, DeGrado WF. 2000. *Nat. Struct. Biol.* 7:161–66
86e. Weiner DB, Liu J, Cohen JA, Williams WV, Greene MI. 1989. *Nature* 339:230–31
86f. Hynes NE, Stern DF. 1994. *Biochim. Biophys. Acta* 1198:165–84
86g. Ellgaard L, Molinari M, Helenius A. 1999. *Science* 286:1882–88
87. Engelman DM, Zaccaï G. 1980. *Proc. Natl. Acad. Sci. USA* 77:5894–98
88. Rees DC, DeAntonio L, Eisenberg D. 1989. *Science* 245:510–13
89. Cornette JL, Cease KB, Margalit H, Spouge JL, Berzofsky JA, DeLisi C. 1987. *J. Mol. Biol.* 195:659–85
90. Deleted in proof
91. Deleted in proof
92. Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA. 1997. *Annu. Rev. Cell Dev. Biol.* 13:457–512
93. Subramaniam S, Lindahl M, Bullough P, Faruqi AR, Tittor J, et al. 1999. *J. Mol. Biol.* 287:145–61
94. Adams PD, Arkin IT, Engelman DM, Brünger AT. 1995. *Nat. Struct. Biol.* 2:154–62
95. Adams PD, Engelman DM, Brünger AT. 1996. *Proteins* 26:257–61
96. Kühlbrandt W, Wang DN, Fujiyoshi Y. 1994. *Nature* 367:614–21
97. Xia D, Yu C-A, Kim H, Xia J-Z, Kachurin AM, et al. 1997. *Science* 277:60–66

98. Bibi E. 1998. *Trends Biochem. Sci.* 23:51–55
99. Doyle DA, Cabral JM, Pfuetzner RA, Kuo A, Gulbis JM, et al. 1998. *Science* 280:69–77
100. Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, et al. 1992. *Nature* 358:727–33
101. Weiss MS, Kreusch A, Schiltz E, Nestel U, Welte W, et al. 1991. *FEBS Lett.* 280:379–82
102. Locher KP, Rees B, Koebnik R, Mitschler A, Moulinier L, et al. 1998. *Cell* 95:771–78
103. Buchanan SK, Smith BS, Venkatramani L, Xia D, Esser L, et al. 1999. *Nat. Struct. Biol.* 6:56–63
104. Unger VM, Kumar NM, Gilula NB, Yeager M. 1997. *Nat. Struct. Biol.* 4:39–43
105. Cheng A, van Hoek AN, Yeager M, Verkman AS, Mitra AK. 1997. *Nature* 387:627–30
106. Walz T, Hirai T, Murata K, Heymann JB, Mitsuoka K, et al. 1997. *Nature* 387:624–27

106a. Tlitsuoka K, Murata K, Walz T, Hirai T, Agre P, et al. 1999. *J. Struct. Biol.* 128:34–43

107. Li HL, Lee S, Jap BK. 1997. *Nat. Struct. Biol.* 4:263–65
108. Boyd D, Manoil C, Froshauer S, San Millan JL, Green N, et al. 1990. *Protein Folding. Deciphering the Second Half of the Genetic Code*, ed. LM Gierash, J King, pp. 314–22. Washington, DC: Am. Assoc. Adv. Sci.
109. Ehrmann M, Beckwith J. 1991. *J. Biol. Chem.* 266:16530–33
110. Kaback HR. 1992. *Int. Rev. Cytol.* 137A:97–125
111. Whitley P, Nilsson L, von Heijne G. 1993. *Biochemistry* 32:8534–39
112. von Heijne G, Wickner W, Dalbey RE. 1988. *Proc. Natl. Acad. Sci. USA* 85:3363–66
113. Yamane K, Akiyama Y, Ito K, Mizushima S. 1990. *J. Biol. Chem.* 265:21166–71
114. McGovern K, Ehrmann M, Beckwith J. 1991. *EMBO J.* 10:2773–82
115. Bibi E, Verner G, Chang C-Y, Kaback HR. 1991. *Proc. Natl. Acad. Sci. USA* 88:7271–75
116. Popot J-L, de Vitry C, Atteia A. 1994. *Membrane Protein Structure: Experimental Approaches*, ed. SH White, pp. 41–96. New York: Oxford Univ. Press
117. Köster W. 1990. *Mol. Gen. Genet.* 223:379–84
118. Liao M-J, Huang K-S, Khorana HG. 1984. *J. Biol. Chem.* 259:4200–4
119. Marti T. 1998. *J. Biol. Chem.* 273:9312–22
120. Martin NP, Leawitt LM, Sommers CM, Dumont ME. 1999. *Biochemistry* 38:682–95
121. Groves JD, Wang L, Tanner MJA. 1998. *FEBS Lett.* 433:223–27
122. Popot J-L, Saraste M. 1995. *Curr. Opin. Biotechnol.* 6:394–402
123. Saier MH Jr, Reizer J. 1991. *Curr. Biol.* 1:362–68
124. Shuman HA. 1987. *Annu. Rev. Genet.* 21:155–77
125. Dassa E, Hofnung M. 1985. *Ann. Inst. Pasteur/Immunol.* 136A:281–88
126. Köster W. 1991. *Biol. Met.* 4:23–32
127. Rubin RA, Levy SB, Heinrikson RL, Kezdy FJ. 1990. *Gene* 87:7–13
128. Rubin RA, Levy SB. 1991. *J. Bacteriol.* 173:4503–9
129. Stühmer W, Conti F, Suzuki H, Wand X, Noda M, et al. 1989. *Nature* 339:597–603
130. Isacoff EY, Jan YN. 1991. *Nature* 353:86–90
131. Anemüller S, Shäfer G. 1990. *Eur. J. Biochem.* 191:292–305
132. Castresana J, Lübben M, Saraste M, Higgins DG. 1994. *EMBO J.* 13:2516–25
133. Castresana J, Lübben M, Saraste M. 1995. *J. Mol. Biol.* 250:202–10

134. Iwata S, Ostermeier C, Ludwig B, Michel H. 1995. *Nature* 376:660–69
135. Ishizuka M, Machida K, Shimada S, Mogi A, Tsuchiya T, et al. 1990. *J. Biochem.* 108:866–73
136. Saraste M, Metso T, Nakari T, Jalli T, Lauraeus M, Van Der Oost J. 1991. *Eur. J. Biochem.* 195:517–25
137. Chepuri V, Gennis RB. 1990. *J. Biol. Chem.* 265:12978–86
138. Pierson BK, Olson JM. 1989. *Microbial Mats*, ed. Y Cohen, E Rosenberg, pp. 402–6. Washington, DC: Am. Soc. Microbiol.
139. Nitschke W, Rutherford AW. 1991. *Trends Biochem. Sci.* 16:241–45
140. Nitschke W, Mühlenhoff U, Liebl U. 1998. *Photosynthesis, a Comprehensive Treatise*, ed. AS Raghavendra, pp. 285–304. Cambridge: Cambridge Univ. Press
141. Vermaas WFJ. 1994. *Photosynth. Res.* 41:285–94
142. Krauss N, Schubert W-D, Klukas O, Fromme P, Witt HT, Saenger W. 1996. *Nat. Struct. Biol.* 3:965–73
143. Fromme P, Witt HT, Schubert W-D, Klukas O, Saenger W, Krauss N. 1996. *Biochim. Biophys. Acta* 1275:76–83
144. Schubert W-D, Klukas O, Krauss N, Saenger W, Fromme P, Witt HT. 1997. *J. Mol. Biol.* 272:741–69
145. Rhee K-H, Morris EP, Barber J, Kühlbrandt W. 1998. *Nature* 396:283–86
146. Michel H, Deisenhofer J. 1988. *Biochemistry* 27:1–7
147. Hankamer B, Morris EP, Barber J. 1999. *Nat. Struct. Biol.* 6:560–64
148. Jennings ML. 1989. *Annu. Rev. Biochem.* 58:999–1027
149. Pao GM, Wu L-F, Johnson KD, Höfte H, Chrispeels MJ, et al. 1991. *Mol. Microbiol.* 5:33–37
150. Wistow GJ, Pisano MM, Chepelinsky AB. 1991. *Trends Biochem. Sci.* 16:170–71
151. Hurtley SM, Helenius A. 1989. *Annu. Rev. Cell Biol.* 5:277–307
152. Manoil C, Traxler B. 1995. *Annu. Rev. Genet.* 29:131–50
153. Deleted in proof
154. Hauska G, Nitschke W, Herrmann RG. 1988. *J. Bioenerg. Biomembr.* 20:211–28
155. Schoepp B, Chabaud E, Breyton C, Verméglio A, Popot J-L. 2000. *J. Biol. Chem.* 275:5275–83
156. Lear JD, Wasserman ZR, DeGrado WF. 1988. *Science* 240:1177–81
157. Oiki S, Danho W, Madison V, Montal M. 1988. *Proc. Natl. Acad. Sci. USA* 85:8703–7
158. Fujii J, Ueno A, Kitano K, Tanaka S, Tada M. 1986. *J. Clin. Invest.* 79:301–4
159. Simmerman HKB, Lovelace DE. 1989. *Biochim. Biophys. Acta* 997:322–29
160. Arkin IT, Rothman M, Ludlam CFC, Aimoto S, Engelman DM, et al. 1995. *J. Mol. Biol.* 248:824–34
161. McDermott G, Prince SM, Freer AA, Hawthornthwaite-Lawless AM, Papiz MZ, et al. 1995. *Nature* 374:517–21
162. Koepke J, Hu XC, Muenke C, Schulten K, Michel H. 1996. *Structure* 4:581–97
163. Herrmann RG, Alt J, Schiller B, Widger WR, Cramer WA. 1984. *FEBS Lett.* 176:239–44
164. Deisenhofer J, Epp O, Miki K, Huber R, Michel H. 1984. *J. Mol. Biol.* 180:385–98
165. Chang C-H, Tiede D, Tang J, Smith U, Norris J, Schiffer M. 1986. *FEBS Lett.* 205:82–86
166. Allen JP, Feher G, Yeates TO, Komiya H, Rees DC. 1987. *Proc. Natl. Acad. Sci. USA* 84:6162–66
167. Arnoux B, Ducruix A, Reiss-Husson F, Lutz M, Norris J, et al. 1989. *FEBS Lett.* 258:47–50
167a. Chang CH, El-Kabbani D, Tiede D, Norris J, Schiffer M. 1991. *Biochemistry* 30:5352–60
168. Ermler U, Fritzsch G, Buchanan SK, Michel H. 1994. *Structure* 2:925–36

168a. Stowell MHB, McPhillips TM, Rees DC, Soltis SM, Abresch E, Feher G. 1997. *Science* 276:812–16
168b. Lancaster CRD, Michel H. 1997. *Structure* 5:1339–59
168c. Lancaster CRD, Michel H. 1999. *J. Mol. Biol.* 286:883–98
169. Grigorieff N, Ceska TA, Downing KH, Baldwin JM, Henderson R. 1996. *J. Mol. Biol.* 259:393–421
170. Kimura Y, Vassylyev DG, Miyazawa A, Kidera A, Matsushima M, et al. 1997. *Nature* 389:206–11
171. Pebay-Peyroula E, Rummel G, Rosenbusch JP, Landau EM. 1997. *Science* 277:1676–81
171a. Essen LO, Siegert R, Lehmann WD, Oesterhelt D. 1998. *Proc. Natl. Acad. Sci. USA* 95:11673–78
172. Luecke H, Richter H-T, Lanyi JK. 1998. *Science* 280:1934–37
173. Mitsuoka K, Hirai T, Miyazawa A, Kidera A, Kimura Y, Fujiyoshi Y. 1999. *J. Mol. Biol.* 286:861–82
174. Weiss MS, Wacker T, Weckesser J, Welte W, Schulz GE. 1990. *FEBS Lett.* 267:268–72
175. Weiss MS, Abele U, Weckesser J, Welte W, Schiltz E, Schulz GE. 1991. *Science* 254:1627–30
176. Weiss MS, Schulz GE. 1992. *J. Mol. Biol.* 227:493–509
177. Kreusch A, Neubüser A, Schiltz E, Weckesser J, Schulz GE. 1994. *Protein Sci.* 3:58–63
178. Schirmer T, Keller TA, Wang Y-F, Rosenbusch J. 1995. *Science* 267:512–14
179. Hirsch A, Breed J, Saxena K, Richter O-MH, Ludwig B, et al. 1997. *FEBS Lett.* 404:208–10
180. Meyer JEW, Hofnung M, Schulz GE. 1997. *J. Mol. Biol.* 266:761–75
180a. Forst D, Welte W, Wacker T, Diederichs K. 1998. *Nat. Struct. Biol.* 5:37–46
180b. Dutzler R, Rummel G, Alberti S, Hernandez-Alles S, Phale PS, et al. 1999. *Structure* 7:425–34
180c. Vogt J, Schulz GE. 1999. *Structure* 7:1301–9
180d. Snijder HJ, Ubarretxena-Belandia I, Blaauw M, Kalk KH, Verheij HM, et al. 1999. *Nature* 401:717–21
181. Picot D, Loll PJ, Garavito RM. 1994. *Nature* 367:243–49
182. Luong C, Miller A, Barnett J, Chow J, Ramesha C, Browner MF. 1996. *Nat. Struct. Biol.* 3:927–33
183. Kurumbail R, Stevens AM, Gierse JK, McDonald JJ, Stegeman R, et al. 1996. *Nature* 384:644–48
184. Tsukihara T, Aoyama H, Yamashita E, Tomikazi T, Yamaguchi H, et al. 1995. *Science* 269:1069–74
185. Yoshikawa S, Shinzawa-Itoh K, Nakashima R, Yaono R, Yamashita E, et al. 1998. *Science* 280:1723–29
185a. Yoshikawa S, Shinzawa-Itoh K, Tsukihara T. 1998. *J. Bioeng. Biomembr.* 30:7–14
185b. Harrenga A, Michel H. 1999. *J. Biol. Chem.* 274:33296–99
186. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE. 1996. *Science* 274:1859–66
187. Zhang Z, Huang L, Shulmeister VM, Chi Y-I, Kim KK, et al. 1998. *Nature* 392:677–84
188. Iwata S, Lee JW, Okada K, Lee JK, Iwata M, et al. 1998. *Science* 281:64–71
188a. Kim H, Xia D, Yu CA, Xia JZ, Kachurin AM, et al. 1998. *Proc. Natl. Acad. Sci. USA* 95:8026–33
189. Chang G, Spencer RH, Lee AT, Barclay MT, Rees DC. 1998. *Science* 282:2220–26
190. Wendt KU, Poralla K, Schulz GE. 1997. *Science* 277:1811–15
191. Wendt KU, Lenhart A, Schulz GE. 1999. *J. Mol. Biol.* 286:175–87
192. Iverson TM, Luna-Chavez C, Cecchini G, Rees DC. 1999. *Science* 284:1961–66
192a. Lancaster CRD, Kröger A, Auer M, Michel H. 1999. *Nature* 402:377–85
192b. Stock D, Leslie AGW, Walker JE. 1999. *Science* 286:1700–5

193. Goodsell DS. 1991. *Trends Biochem. Sci.* 16:203–6
194. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. 1994. *Molecular Biology of the Cell.* New York: Garland. 3rd ed. 1294 pp.
195. Saxton MJ. 1989. *Biophys. J.* 56:615–22
196. Lavergne J, Bouchaud J-P, Joliot P. 1992. *Biochim. Biophys. Acta* 1101:13–22
197. Peters R. 1986. *Biochim. Biophys. Acta* 864:305–59
198. Dilger JP, Fisher LR, Haydon DA. 1982. *Chem. Phys. Lipids* 30:159–76
199. Allen LC. 1975. *Proc. Natl. Acad. Sci. USA* 72:4701–5
200. Reynolds JA, Gilbert DB, Tanford C. 1974. *Proc. Natl. Acad. Sci. USA* 71:2925–27
201. Amzel LM. 1997. *Proteins* 28:144–49
202. Bibi E, Kaback HR. 1992. *Proc. Natl. Acad. Sci. USA* 89:1524–28
203. Schmidt-Rose T, Jentsch TJ. 1997. *J. Biol. Chem.* 272:20515–21
204. Nielsen SM, Elling CE, Schwartz TW. 1998. *Eur. J. Biochem.* 251:217–26
205. Koebnik R. 1996. *EMBO J.* 15:3529–37
206. Liao M-J, London E, Khorana HG. 1983. *J. Biol. Chem.* 258:9949–55
207. Sigrist H, Wenger RH, Kislig E, Wuthrich M. 1988. *Eur. J. Biochem.* 177:125–33
208. Hansen OK, Pompejus M, Fritz H-J. 1994. *Biol. Chem. Hoppe-Seyler* 375:715–19
209. Kobilka BK, Kobilka TS, Daniel K, Regan JW, Caron MG, Lefkowitz RJ. 1988. *Science* 240:1310–16
210. Maggio R, Vogel Z, Wess J. 1993. *FEBS Lett.* 319:195–200
211. Ridge KD, Lee SSJ, Yao LL. 1995. *Proc. Natl. Acad. Sci. USA* 92:3204–8
212. Yu H, Kono M, McKee TD, Oprian DD. 1995. *Biochemistry* 34:14963–69
213. Wrubel W, Stochaj U, Sonnewald U, Theres C, Ehring R. 1990. *J. Bacteriol.* 172:5374–81
214. Wrubel W, Stochaj U, Ehring R. 1994. *FEBS Lett.* 349:433–38
215. Zen KH, McKenna E, Bibi E, Hardy D, Kaback HR. 1994. *Biochemistry* 33:8198–206
216. Bibi E, Kaback HR. 1990. *Proc. Natl. Acad. Sci. USA* 87:4325–29
217. Groves JD, Tanner MJA. 1995. *J. Biol. Chem.* 270:9097–105
218. Wang L, Groves JD, Mawby WJ, Tanner MJA. 1997. *J. Biol. Chem.* 272:10631–38
219. Groves JD, Wang L, Tanner MJA. 1998. *Biochem. J.* 332:161–71
220. Wilkinson BM, Esnault Y, Craven RA, Skiba F, Fieschi J, et al. 1997. *EMBO J.* 16:4549–59
221. Heinemeyer W, Alt J, Herrmann RG. 1984. *Curr. Genet.* 8:543–49
222. Yamaguchi M, Hatefi Y, Trach K, Hoch JA. 1988. *J. Biol. Chem.* 263:2761–67
223. Jakobsson E. 1997. *Trends Biochem. Sci.* 22:339–44
224. Stowell MHB, McPhillips TM, Rees DC, Soltis SM, Abresch E, Feher G. 1997. *Science* 276:812–16
225. Prince SM, Papiz MZ, Freer AA, McDermott G, Hawthornthwaite-Lawless AM, et al. 1997. *J. Mol. Biol.* 268:412–23
226. Sharp KA, Nicholls A, Friedman R, Honig B. 1991. *Biochemistry* 30:9686–97