Pergamon

Prog. Biophys. molec. Biol., Vol. 66, No. 2, pp. 113–139, 1996
© 1997 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0079-6107/96 $32.00 + 0.00

PII: S0079-6107(96)00018-1

PRINCIPLES OF MEMBRANE PROTEIN ASSEMBLY AND STRUCTURE

GUNNAR VON HEIJNE

Department of Biochemistry, Stockholm University, S-106 91 Stockholm, Sweden

I. INTRODUCTION

Membrane proteins come in a variety of sizes and shapes, though the basic architectural principles apparent from the few available three-dimensional (3D) structures are far less diverse than those of the globular proteins. This is a consequence of the requirement that membrane proteins must either bind to one leaflet of, or completely span, a lipid bilayer while at the same time maintaining contact with the surrounding aqueous medium. Peripheral binding to one side of the bilayer can be mediated by amphiphilic α-helices oriented parallel to the bilayer surface, and membrane-spanning structures can be built either from β-barrels or from bundles of transmembrane α-helices. Large extramembraneous domains do not seem to be much influenced by their attachment to the bilayer, and can essentially be viewed as tethered globular proteins.

The structural differences between membrane and globular proteins do not only result from the different physico-chemical requirements imposed by their different environments, but also reflect fundamental differences in the mechanisms responsible for their folding and assembly. While globular proteins encounter only one kind of environment from the time of their synthesis until they are degraded in old age, membrane proteins are first exposed to aqueous surroundings and only later gain access to the milieu for which they are designed. The problem of how the controlled integration of a membrane protein into the lipid bilayer takes place is still not fully worked out, and there are certainly aspects of membrane protein structure that will only be fully appreciated once this has been accomplished.

Both the structural and biogenetic aspects of membrane proteins have been reviewed before (Cowan and Rosenbusch, 1994; Lemmon and Engelman, 1994; von Heijne, 1994a; Manoil and Traxler, 1995; Popot and Saraste, 1995; Reithmeier, 1995; Spiess, 1995), and my ambition here is mainly to provide an update of what has happened in the field during the last few years. This includes some highly interesting new structures and some pioneering studies on helix–helix packing in a non-aqueous environment. The basic aspects of the initial step of insertion into the lipid bilayer are becoming increasingly clear, and there have also been some definite advances in the area of membrane protein structure prediction. For fuller treatments of many facets of the membrane protein assembly problem, the reader is referred to a recent book (von Heijne, 1997).

II. WHAT THE STRUCTURES SAY

Membrane proteins are not only tricky to handle biochemically, they are also notoriously difficult to overexpress (Grisshammer and Tate, 1995) and seldom yield high quality 2D or 3D crystals. So far, all the known structures are of proteins that can be isolated in sufficient quantities from natural sources, though recent work on overexpression in *E. coli* (Miroux and Walker, 1996) and other systems (Tucker and Grisshammer, 1996) looks promising.

Good 3D crystals seem to require the presence of rather extensive globular domains protruding from the membrane that can form the necessary crystal contacts. In the case of the bacterial cytochrome *c* oxidase, the globular domain was artificially enlarged by binding a monoclonal F_v fragment to the protein complex prior to crystallization (Iwata *et al.*, 1995). The formation of 2D crystals, on the other hand, can be induced by protein–protein contacts in the plane of the membrane (Kühlbrandt, 1992), and such crystals may thus be easier to obtain when the globular domains are small. However, structure determination by electron crystallography of 2D crystals rarely gives sufficient resolution for structure modeling on an atomic scale.

In this section, all the known high-resolution membrane protein structures as of August, 1996, will be briefly reviewed. The focus will be on structural rather than functional aspects.

2.1. *Bacteriorhodopsin: the Tour de Force of Electron Crystallography*

Bacteriorhodopsin, a light-driven proton pump from *Halobacterium halobium* (Lanyi, 1995), was the first integral membrane protein to be shown to consist of a bundle of transmembrane helices oriented roughly perpendicular to the membrane plane (Henderson and Unwin, 1975). Fifteen years later, the structure had been determined to sufficient resolution to allow the bulky aromatic side chains of the transmembrane helices to be built into the electron density (Henderson *et al.*, 1990), and another six years of data collection and analysis finally has produced a density map where most side chains, both in the helices and the connecting loops, can be fitted (Grigorieff *et al.*, 1996, Fig. 1A). Bacteriorhodopsin is not only one of the best structurally and functionally characterized integral membrane proteins, but has also served as the test-bed for the development of both hardware and software for electron crystallography.

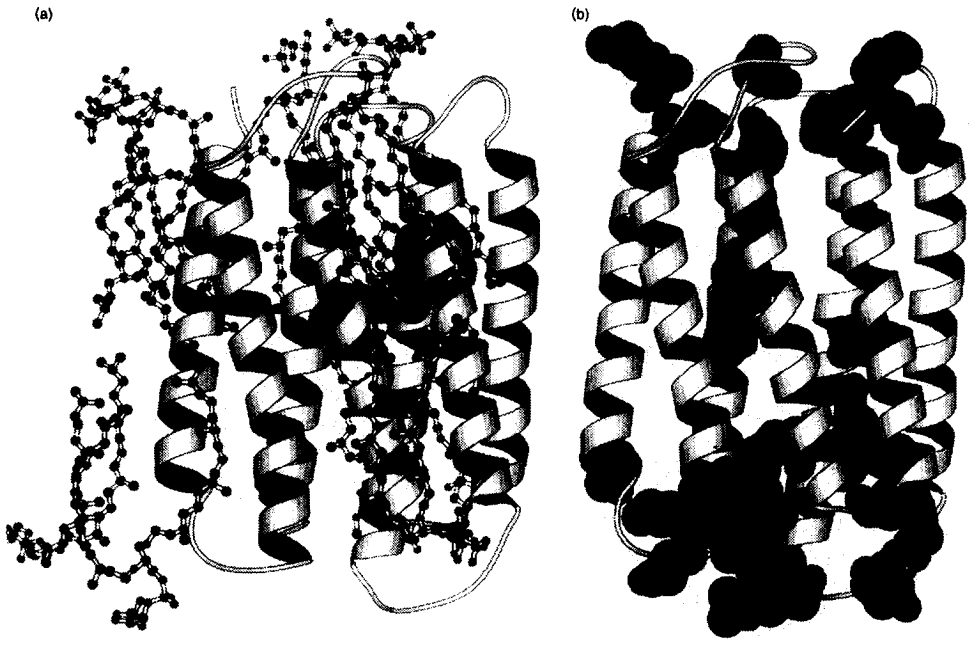

Fig. 1. Bacteriorhodopsin. (A) 3D structure (PDB file 2BRD). The retinal is shown as a space-filling model, and lipids visible in the electron density are shown as ball and stick models. The cytoplasmic side is down. (B) Same as in A, but with charged residues shown as space-filling models. All 3D structure illustrations have been made using MOLSCRIPT (Kraulis, 1991).

The transmembrane helices are composed mainly of hydrophobic residues. A few charged residues are buried within the structure, serving important functions in the proton-relay system (Fig. 1B). Bacteriorhodopsin has been widely used as a starting structure for modeling many medically important G-protein-coupled receptors, though it now appears that the relative disposition of the helices in these eukaryotic so-called 7TM proteins is sufficiently different from that in bacteriorhodopsin to make such modeling very difficult (Schertler *et al.*, 1993).

2.2. *Photosynthetic Reaction Center: a Helix-lined Electron Conductor*

The structure of the photosynthetic reaction center from *Rhodopseudomonas viridis* was solved by X-ray crystallography in 1985 (Deisenhofer *et al.*, 1985), followed two years later by the closely related one from *Rhodobacter sphaeroides* (Allen *et al.*, 1987; Yeates *et al.*, 1987). The structure is considerably larger than bacteriorhodopsin, with a total of three membrane-bound and one peripheral protein subunits, 11 transmembrane helices, four bacteriochlorophylls, two bacteriopheophytins, two quinones, and a non-heme iron. The heart of the protein is the electron-conducting chain of co-factors held in place by the protein scaffold (Fig. 2A).

Since this was the first high-resolution structure of a membrane protein available, it provided the first clear indications of how transmembrane helices pack (Rees *et al.*, 1989a,b). The atomic packing density in the interior of the protein is similar to that found in globular proteins, as is the average hydrophobicity of the internal residues. Residues exposed to the lipid environment are more hydrophobic than the internal ones, just the opposite of what is the case for globular proteins. The exposed surface residues are more variable between homologous proteins from different species than are the internal residues.

Many of the transmembrane helices protrude quite a distance from the bilayer, and consequently have a central hydrophobic part flanked by more polar residues.

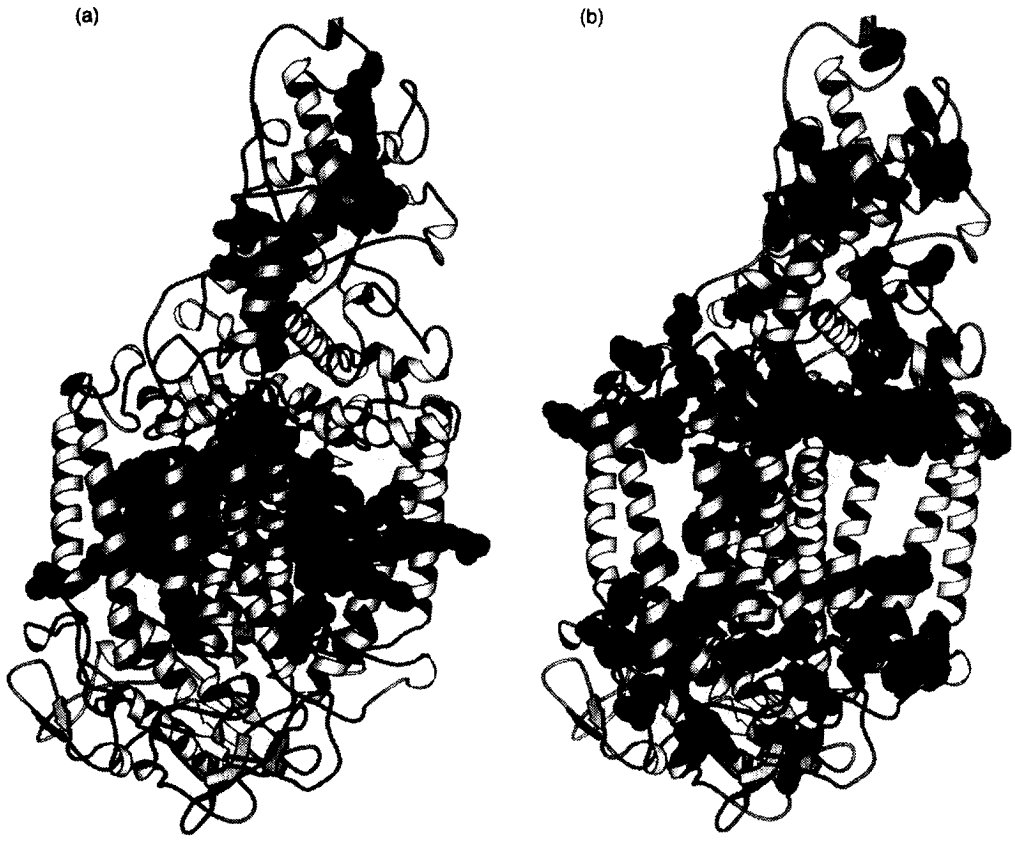

Fig. 2. *Rhodopseudomonas viridis* photosynthetic reaction center. (A) 3D structure (PDB file 1PRC). The cytoplasmic side is down. The extension of the lipid bilayer corresponds roughly to the large helix bundle in the lower part of the structure. Co-factors are shown as space-filling models. (B) Same as in A, but with Trp and Tyr residues shown as space-filling models.

Interestingly, Tyr and Trp residues tend to be concentrated near the presumed position of the lipid headgroup region (Fig. 2B).

2.3. *Plant Light-harvesting Complex: Ion-pairs and Crossbraces*

Another protein involved in photosynthesis, the plant light-harvesting chlorophyll a/b–protein complex, has been solved to high resolution, in this case by electron crystallography (Kühlbrandt and Wang, 1991; Kühlbrandt *et al.*, 1994). This is a rather small protein with three transmembrane helices that provide binding sites for a large number of chlorophylls and two carotenoids (Fig. 3). Two of the helices form a pair and are highly tilted (32°) relative to the membrane normal, whereas the third helix makes no direct contacts with the other two. The helix pair appears to be stabilized by two salt-bridges at one end, and by the two carotenoids that serve as crossbraces, almost spanning the entire bilayer.

2.4. *Bacterial Light-harvesting Complex II: the Importance of Chlorophyll*

The structure of the light-harvesting complex II, a part of the photosynthetic antennae, from *Rhodospeudomonas acidophila* was recently solved by X-ray crystallography (McDermott *et al.*, 1995). This remarkable ring structure is composed of nine α- and nine β-subunits, each only contributing a single transmembrane helix, and 27 chlorophylls (Fig. 4). The α-subunits form an inner ring, separated from the outer ring of β-subunits by a layer of chlorophylls. There is thus very little contact between the α- and β-subunits, and the chlorophylls seem to be as important for keeping the structure together as are the

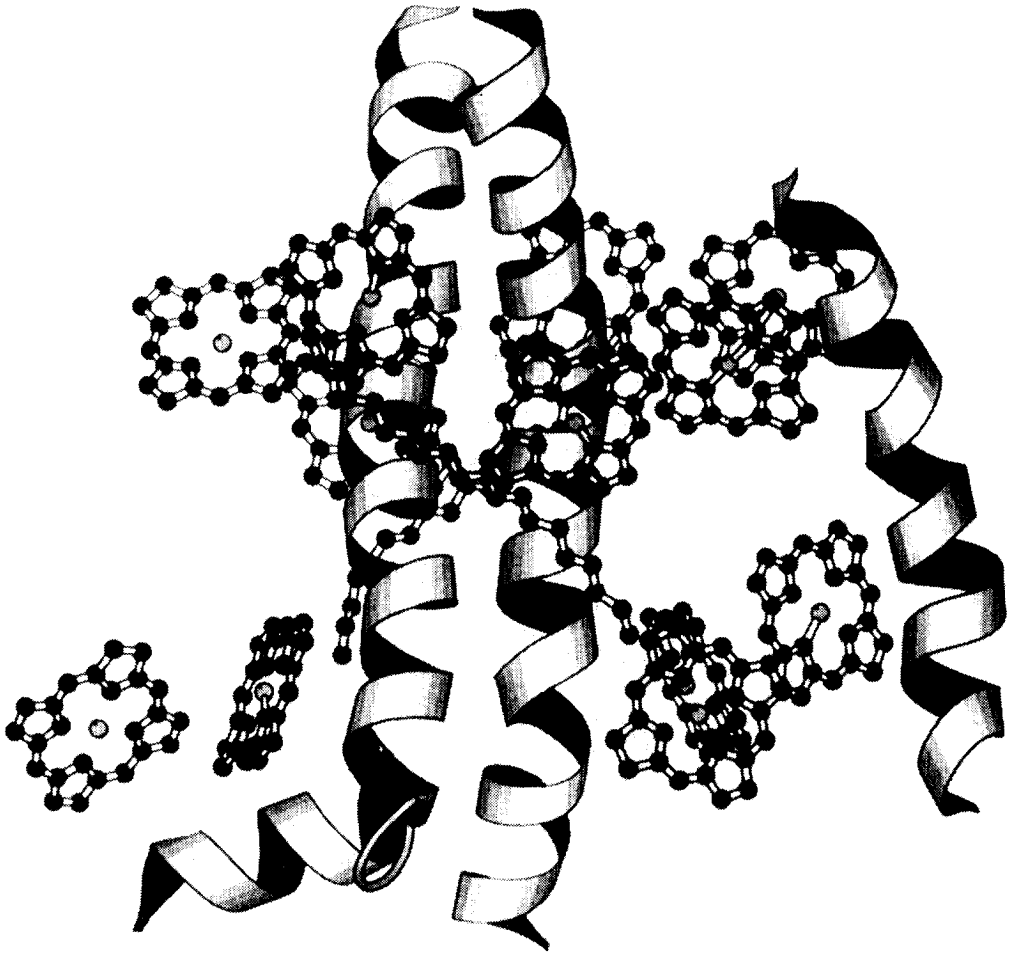

Fig. 3. Plant light-harvesting chlorophyll a/b-protein (courtesy of Dr. W. Kühlbrandt, Heidelberg). The stromal side is down. The long N- and C-terminal helices are tilted by about 30° in the membrane whereas the shorter middle helix is less tilted. Chlorophylls and carotenoid molecules are shown as ball-and-stick models. Since the loops connecting the helices are not well defined, they have been omitted from the structure.

proteins. The tightly interacting chlorophylls provide an ideal structure for charge delocalization and serve to funnel the electron resulting from light excitation towards the reaction center.

2.5. *Cytochrome* c *Oxidase: the Biggest Yet*

Given the paucity of membrane protein structures, the almost simultaneous appearance of the structures of the bacterial and mitochondrial cytochrome *c* oxidase complexes was a delightful coincidence (Iwata *et al.*, 1995; Tsukihara *et al.*, 1995; 1996). The bacterial complex contains three central subunits, while the mitochondrial one has no less than 13 subunits, 10 of which have one or more transmembrane helices, pushing the total number of helices to 28 (Fig. 5). In addition, the two catalytically important Cu-sites, two hemes, and eight phospholipids have been located in the structure.

Analysis of the structure of the mitochondrial oxidase has suggested a rather detailed consensus structure that seems to be shared between the cytochrome *c* oxidase, the bacterial photosynthetic reaction center, and bacteriorhodopsin, and that may thus be generally valid for helix bundle membrane proteins (Elofsson *et al.*, 1997): a central, helical region of about 20 Å rich in aliphatic residues and Phe, an "aromatic belt" composed of Trp and Tyr similar to but less prominent than that seen in the β-barrel

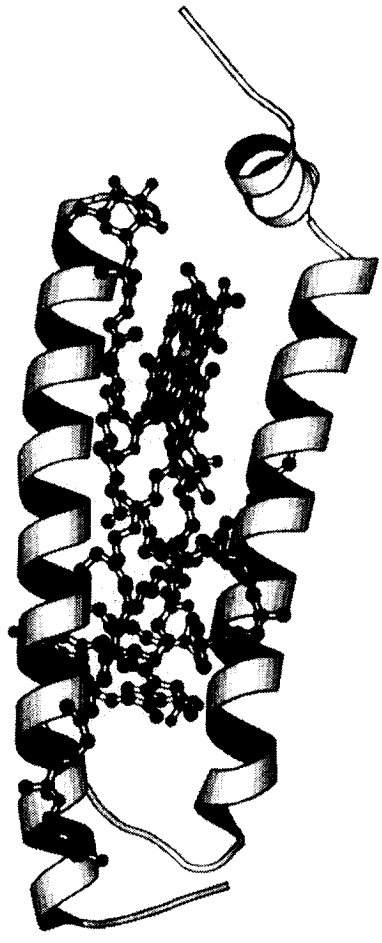

Fig. 4. One α- and one β-subunit from the *Rhodopseudomonas acidophila* light-harvesting antenna complex (courtesy of Dr. R.J. Cogdell, Glasgow). In the full complex, nine α-subunits form and inner ring separated from an outer ring of nine β-subunits by bacteriochlorophylls and carotenoid molecules (shown as ball-and-stick models). The cytoplasmic side is down.

porin structures (see next section), a region near the helix ends with amidated and charged residues, and helix-breaking Pro residues just outside the helices.

2.6. *Porins: β-Barrels and Aromatic Belts*

The porins form passive diffusion pores in the outer membrane of Gram-negative bacteria, and are built on a β-barrel architecture. The structures of five porins are known: those from *Rhodobacter capsulatus* (Weiss *et al.*, 1991b), *Rhodopseudomonas blastica* (Kreusch *et al.*, 1994), and the *Escherichia coli* OmpF, PhoE, and LamB porins (Cowan *et al.*, 1992; Schirmer *et al.*, 1995). In all cases, the structure is composed of a large anti-parallel β-barrel with short loops facing the periplasm and larger loops protruding outside the outer membrane (Fig. 6A). Substrate specificity and translocation kinetics depend on a segment of the chain that loops into the barrel (Fig. 6B), thus modulating its properties (Weiss *et al.*, 1991a; Schirmer *et al.*, 1995).

Not surprisingly, the outer surface of the barrel is composed of hydrophobic residues, while residues on the inside may be both hydrophobic and hydrophilic. Since the membrane-spanning segments of the chain are β-strands, this means that only every second residue is facing the lipids and hence needs to be hydrophobic. The strands in porin from *R. capsulatus* are tilted by about 30°–60° relative to the membrane normal, and are 6–17 residues long (Weiss *et al.*, 1991a). The long hydrophobic segments typical of

Fig. 5. Bovine mitochondrial cytochrome *c* oxidase (courtesy of Dr. Tomitake Tsukihara, Osaka, and Dr. Shinya Yoshikawa, Hyogo). The matrix side is down.

the helix bundle membrane proteins are absent from the porins, and in fact such segments, if introduced by molecular genetic techniques into an outer membrane protein, become trapped in the inner membrane during translocation (MacIntyre *et al.*, 1988), thus preventing assembly into the outer membrane.

A notable feature of the porin structures are the belts of aromatic residues that point out towards the lipid headgroup region on either side of the bilayer (Fig. 6A). The functional significance of these aromatic belts is not known, though it has been speculated that they may serve to keep the molecule in a fixed orientation relative to the bilayer (Weiss *et al.*, 1991a). As noted above, a similar clustering of aromatic residues near the lipid headgroup region, Trp and Tyr in particular, is also seen in the helix bundles proteins.

2.7. *Prostaglandin H_2 Synthase: Surface-anchoring Amphiphilic Helices*

Prostaglandin H_2 synthase binds to the lumenal leaflet of the endoplasmic reticulum (ER) membrane, and is thus classified as a peripheral membrane protein. In contrast to the peripherally bound globular subunits of, e.g., the photosynthetic reaction center and cytochrome *c* oxidase, its membrane binding is not mediated by other proteins but is the result of direct protein–lipid interactions.

The structure of prostaglandin synthase strongly suggests that binding to the bilayer is through four amphiphilic α-helices which, in contrast to most helices in globular proteins, has a series of solvent-exposed hydrophobic residues flanked by basic residues (Picot and

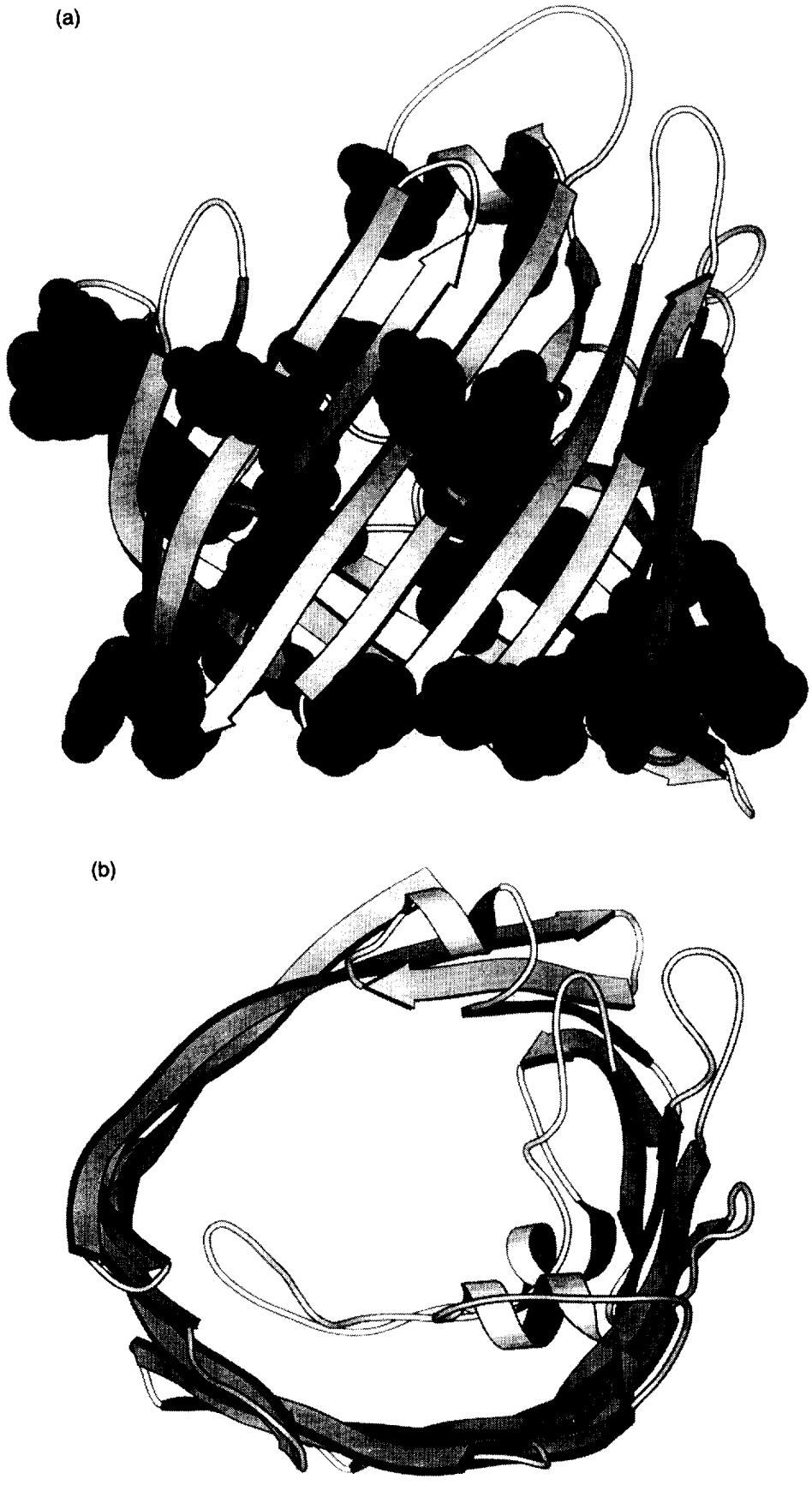

Fig. 6. *Rhodobacter capsulata* outer membrane porin. (A) 3D structure (PDB file 2POR) viewed in the plane of the membrane. The periplasmic side is down. Aromatic residues (Phe, Trp, Tyr) are shown as space-filling models. (B) View from outside the membrane. Note the internal loop that controls the pore properties.

Garavito, 1994; Picot *et al.*, 1994, Fig. 7). These hydrophobic residues are postulated to insert into the bilayer, while the flanking charged residues may interact with the phospholipid headgroups. The amphiphilic helices are thus thought to lie more or less flat on the membrane surface, a mode of binding that has also been seen in some small lytic peptides such as magainin (Bechinger *et al.*, 1993).

2.8. *Pore-forming Toxins: Inside-out Proteins*

An interesting class of integral membrane proteins are the pore-forming toxins (Parker and Pattus, 1993). These molecules manage the remarkable feat of first folding into a globular, water-soluble structure that can convert into a membrane-bound form with ion channel activity upon exposure to a lipid bilayer under the proper conditions.

The known structures are all of the water-soluble form, and are only suggestive of how the membrane-inserted structure may look. Nevertheless, there is good biophysical and biochemical evidence for both helix bundle and β-barrel structures (Parker and Pattus, 1993; Parker *et al.*, 1994), though just how similar these structures may be to the membrane proteins described above is not known. As an example, the structure of the soluble form of colicin A (Parker *et al.*, 1989) is shown in Fig. 8. The molecule consists of two layers of α-helices surrounding a central pair of very hydrophobic helices—a "helical hairpin". The initial binding to the lipid bilayer is thought to involve surface-exposed basic residues near the tight turn between the two hydrophobic helices (Lakey *et al.*, 1994). Upon binding, the whole structure converts into a molten-globule-like state (van der Goot *et al.*, 1991) allowing the central helical hairpin to associate with the lipids (Lakey *et al.*, 1993). How the ion channel is formed after these initial events is not known.

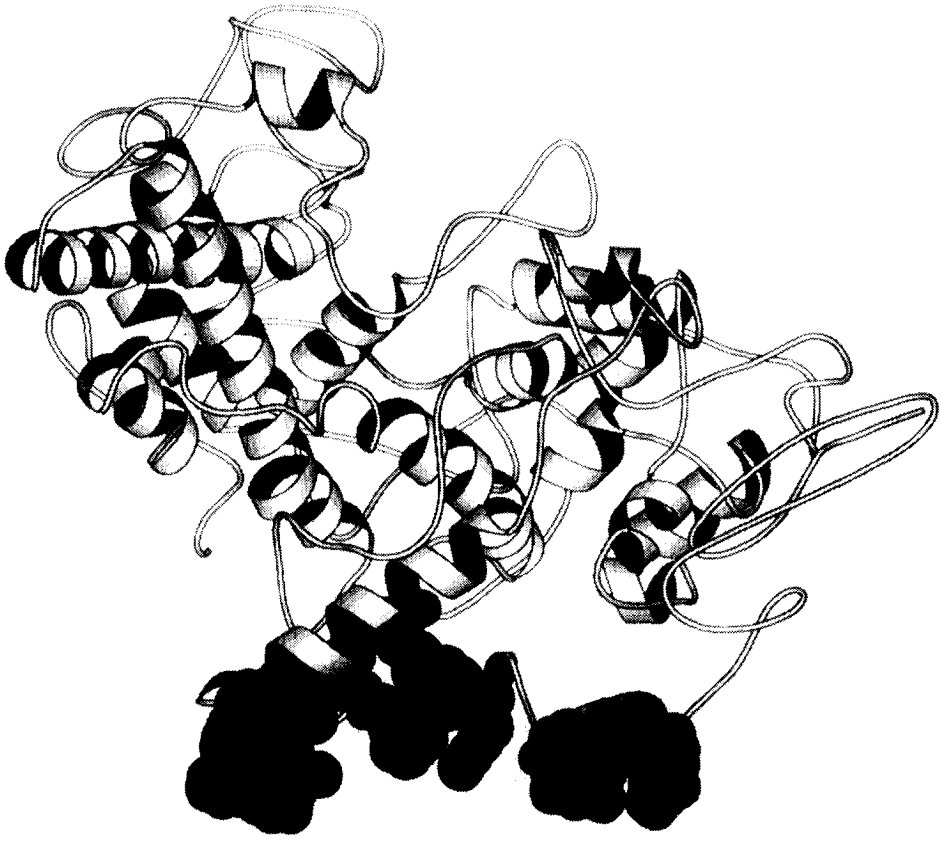

Fig. 7. Ovine prostaglandin H_2 synthase-1 (PDB file 1PRH). Note the cluster of exposed hydrophobic residues on the bottom α-helices (shown as space-filling models) believed to anchor the molecule in the lumenal leaflet of the ER membrane.

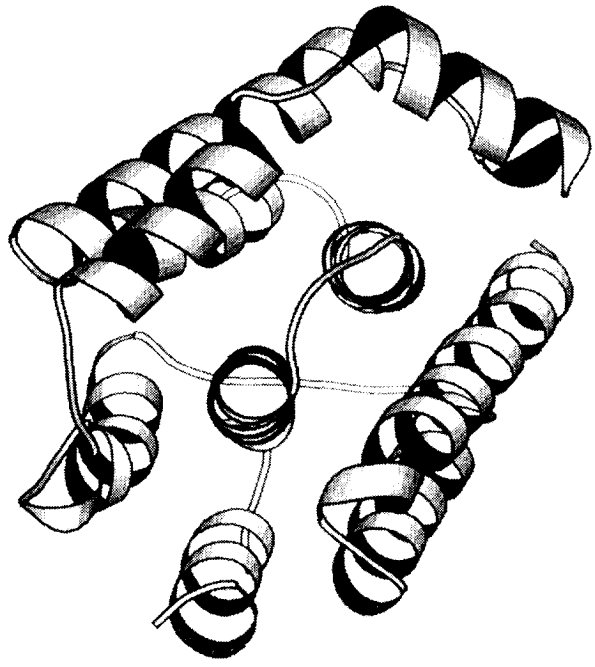

Fig. 8. The soluble form of colicin A (PDB file 1COL). The molecule is viewed from the side believed to bind to the bacterial membrane, i.e. the view is as seen by the victim. The hydrophobic helical hairpin (viewed on end) is sandwiched between two layers of helices that protect it from contact with water. Upon membrane binding, the helical hairpin is though to insert into the lipid bilayer.

III. POOR MAN'S X-RAY CRYSTALLOGRAPHY: STRUCTURE DETERMINATION THROUGH MUTAGENESIS

The difficulties encountered with high-resolution structure determination of membrane proteins has prompted the development of alternative approaches based mainly on molecular genetics, sometimes coupled with various biophysical techniques. The basic idea is to obtain bits and pieces of structural information related to helix–helix packing that can then be used to constrain model-building exercises.

3.1. *Glycophorin A and Phospholamban: Helix Packing Motifs*

Glycophorin A (GpA) is a single-spanning membrane protein found on red blood cells. It forms unusually stable homo-dimers that will not dissociate even in SDS, and that can thus be readily identified by denaturing SDS-PAGE. Competition of dimer formation with synthetic peptides (Lemmon *et al.*, 1992a) and the construction of various chimeric proteins have shown that the transmembrane helix is both necessary and sufficient for dimerization.

These properties have made it possible to study helix packing in the GpA dimer by saturation and site-directed mutagenesis. In a first series of experiments, some 280 random mutations throughout the transmembrane helix were analysed, and it was found that only seven residues were critical for dimer formation (Fig. 9A). All these residues line up on one side of the helix, suggesting that the helix dimer is a right-handed supercoil (Lemmon *et al.*, 1992b; Treutlein *et al.*, 1992; Fig. 9B). Further studies demonstrated that all the other residues in the helix could be simultaneously changed to leucine with minimal effects on dimer formation (Lemmon *et al.*, 1994).

The central feature of the helix–helix interface is a Gly-Val-Xxx-Xxx-Gly-Val motif in each helix. Apparently, the glycines allow a very close approach between the two helices, while the valines provide a large contact area on either side of the Gly–Gly contact. The

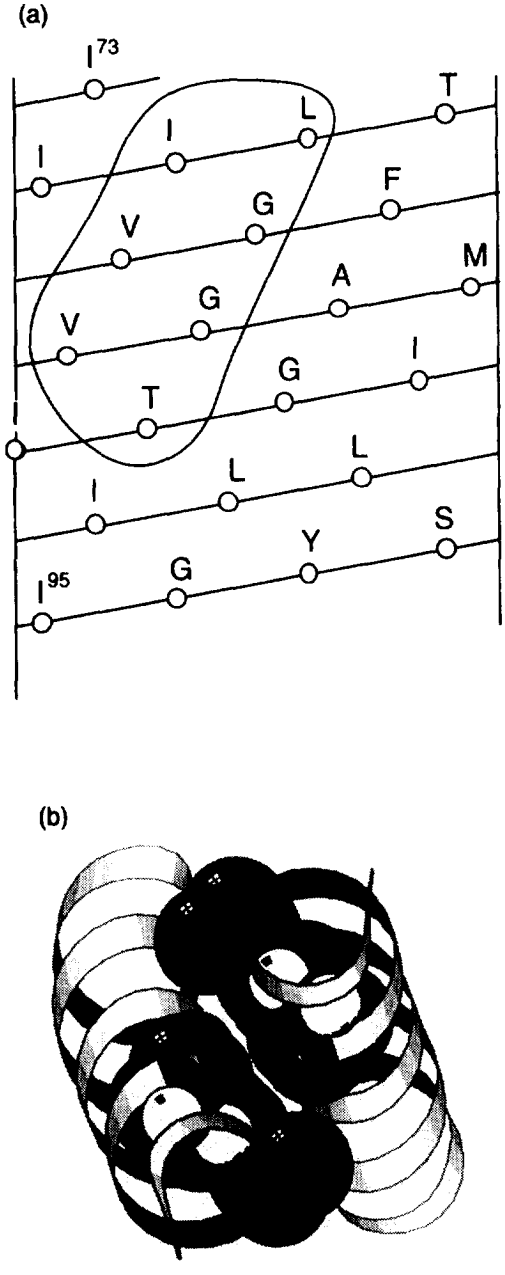

Fig. 9. Glycophorin A. (a) Helical net plot of the transmembrane helix. The packing interface is encircled. (b) 3D model for the transmembrane helix dimer (courtesy of Dr. Axel Brünger, Yale). One of the two central Gly-Val motifs is shown as a space-filling model. Note how the two glycines pack across the helix–helix interface, and how the flanking valines cover the Gly–Gly contact area.

N-terminal Ile–Leu pair is somewhat less critical, and can be replace by certain other combinations of hydrophobic residues (Mingarro *et al.*, 1996).

The initial mutational analysis of the GpA dimer was based on a very large number of random mutations, and it has recently been shown that a more rapid way to identify the critical interface residues is to make a first screen using alanine insertions rather than substitutions (Mingarro *et al.*, 1996). The basis for this approach is that the insertion of a residue into a transmembrane helix will displace the residues on the N-terminal side of the insertion by 100° relative to those on the C-terminal side, thus effectively disrupting helix–helix packing interfaces. Insertions that have no effect on GpA dimerization will thus

allow the ends of the interface motif to be identified before saturation mutagenesis is applied. An Ala-insertion scan of helix III in lactose permease (see next section) suggests that insertion mutagenesis may be a generally efficient way to obtain a first, rough idea of where the structurally and functionally critical residues reside in transmembrane helices (Braun *et al.*, 1997).

Mutational analysis of the homo-pentameric channel protein phospholamban that forms SDS-resistant pentamers has recently been carried out (Arkin *et al.*, 1994; 1995; 1996; Adams *et al.*, 1995; Ludlam *et al.*, 1996; Simmerman *et al.*, 1996), much along the lines established for GpA, and a 3D model has been proposed (Fig. 10). In this case, the packing in the helix–helix interfaces is reminiscent of the well-characterized leucine zipper motif found in globular proteins.

3.2. *Lactose Permease: the Mutagenesis Paradigm*

Of all membrane proteins that have been subjected to mutational analysis, lactose permease (LacY) from *E. coli* has no competitor (Kaback *et al.*, 1994; Kaback, 1996). Extensive site-directed and cysteine-scanning mutagenesis studies in which all of the 417 residues in the permease have been mutagenized have revealed that only four residues—Glu_{269} (helix VIII), Arg_{302} (helix IX), His_{322} (helix X), and Glu_{325} (helix X)—are irreplaceable with respect to active lactose transport. Such a small number of essential residues may appear surprising at first sight, but is in fact similar to what has been found by extensive mutagenesis of another *E. coli* inner membrane protein, diacylglycerol kinase (Wen *et al.*, 1996).

This library of single and double Cys mutants has made possible a number of sophisticated biophysical approaches aimed at obtaining structural information. Attachment to defined cysteines of fluorescent probes capable of exhibiting excimer fluorescence if located within 3.5 Å of each other has allowed neighboring residues in the structure to be identified (Jung *et al.*, 1993). Neighboring residues have also been identified by engineering divalent metal-binding sites (bis- or tris-His residues) within the permease (He *et al.*, 1995a,b; Jung *et al.*, 1995) and by using engineered Cu-binding sites to effect site-directed chemical cleavage (Wu *et al.*, 1995). More long-range distance

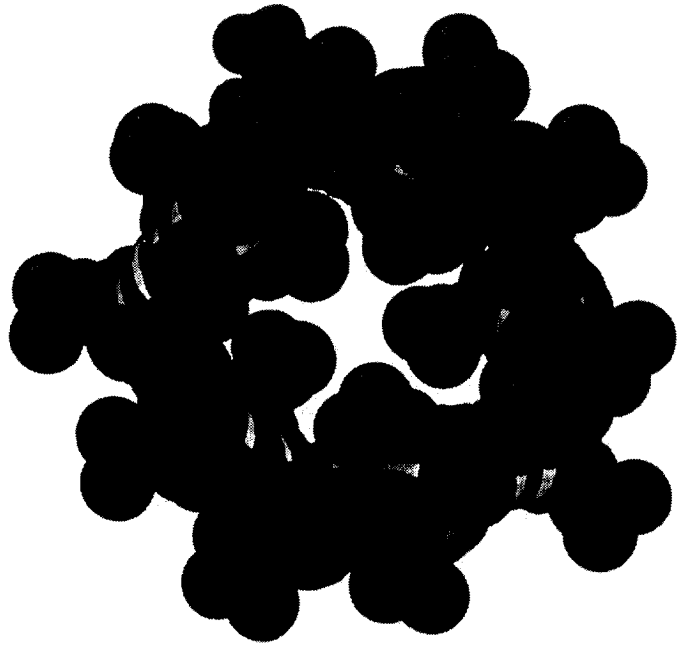

Fig. 10. Model for the phospholamban pentameric ion channel viewed from outside the membrane (PDB file 1PSL). All leucine residues are shown as space-filling models.

measurements have been made by a combination of engineered metal-binding sites and spin labeling of defined cysteines (Voss *et al.*, 1995). Using site-directed mutagenesis of functionally critical residues followed by genetic selection of second-site revertants, specific ion-pairs have been identified that also provide structural constraints (Lee *et al.*, 1992a; 1993; Dunten *et al.*, 1993). Collectively, these studies have allowed a rough model of the packing of helices VII to XI to be proposed (Kaback *et al.*, 1994; Kaback, 1996).

Related methods that have been used to obtain structural information on other membrane proteins include the engineering of double-Cys mutants that may be oxidized to form disulfide bonds if properly located (Pakula and Simon, 1992; Whitley *et al.*, 1993; Lee *et al.*, 1994; Lee and Hazelbauer, 1995), and the use of lipid- and water-soluble probes to quench the signal from ESR spin-labels attached to defined cysteines (Altenbach *et al.*, 1989; 1990; 1994). Also, a large body of data related to ligand binding residues in the family of G-protein-coupled receptors have been used to extract structural information (Schwartz *et al.*, 1995).

3.3. *Glycosylation Mapping: a Molecular Ruler for Transmembrane Helices*

Segments of eukaryotic membrane proteins that are translocated across the ER membrane (see below) can become modified by N-linked glycosylation on Asn-X-Thr/Ser acceptor sites. The transfer of the oligosaccharide moiety onto the nascent chain is catalysed by oligosaccharyl transferase (OST), itself an oligomeric integral membrane protein with its catalytic site located in the lumen of the ER (Kelleher *et al.*, 1992; Silberstein *et al.*, 1992).

Detailed mutagenesis studies demonstrated that potential Asn-X-Thr/Ser acceptor sites can only be glycosylated if they are located at least 10–12 residues away from the nearest transmembrane segment (Nilsson and von Heijne, 1993), suggesting that the OST active site is located some distance away from the membrane surface. Further studies using artificial poly-Leu transmembrane segments showed that the location of the end of a hydrophobic transmembrane segment relative to the OST active site can be determined with high precision by measuring the degree of glycosylation of potential acceptor sites located at different distances form the hydrophobic segment (Nilsson *et al.*, 1994). The OST active site can thus be used as a point of reference, in essence providing a "molecular ruler" that can be used to study transmembrane helices.

This approach has been used to determine the point at which a transmembrane helix exits the membrane, i.e., the approximate location of the helix relative to the lipid headgroup region (our unpublished data). So far, the number of residues between the known end of the transmembrane segment in the H-subunit of the photosynthetic reaction center and a potential acceptor site required for half-maximal glycosylation of that site has been determined. The location of other transmembrane helices relative to the lipid headgroup region can now be determined from the position of the half-maximally glycosylated acceptor site by comparison to the reference H-subunit helix.

In summary, techniques based on mutagenesis, although labor-intensive and yielding only bits and pieces of structural information, may prove to be the only way to guide modeling efforts when good crystals cannot be obtained.

IV. MEMBRANE PROTEIN TOPOLOGY: THE INS AND OUTS OF TRANSMEMBRANE PROTEINS

A fundamental aspect of membrane protein structure is topology, i.e. the number of transmembrane segments and their orientation in the membrane. For most proteins, this is the only structural information available, and it provides a basis for site-directed mutagenesis and for all attempts to model the structure. Topology is often predicted from sequence (see below), but many experimental methods for determining topology have also been developed.

In *E. coli*, the most widely used approach is to make a series of fusions between parts of the protein under study and a reporter protein such as PhoA, β-lactamase (Bla), or LacZ

(Manoil, 1991; Hennessey and Broome-Smith, 1993; Prinz and Beckwith, 1994). PhoA and Bla are active only in the periplasm whereas LacZ is active in the cytoplasm, and the respective activities can be easily determined both on plates and in liquid culture. PhoA or Bla fusions that display high activity suggest a periplasmic location of the loop to which the reporter has been fused, while active LacZ fusions signal a cytoplasmic location. Although there are possible pitfalls, these can usually be avoided by a judicious selection of fusion points (Boyd *et al.*, 1993), or, if necessary, by the construction of sandwich fusions where the reporter is inserted in-frame into the membrane protein (Ehrmann *et al.*, 1990).

A more direct method is to probe the topology of an inner membrane protein by protease-treatment of spheroplasts. An extension of this method is to engineer specific protease-sensitive sites into different loops (Whitley *et al.*, 1994a; Sahin-Toth *et al.*, 1995).

In eukaryotic cells, sites carrying N-linked oligosaccharides arguably provide the most reliable topological marker, since Asn-X-Thr/Ser acceptor sites can only be modified in the lumen of the ER. Such sites can also easily be inserted into the protein, and will almost always be modified if they are located in a lumenal loop and not too close to the neighboring transmembrane segments (see above). Proteins modified in this way can either be analysed *in vivo* or, more conveniently, *in vitro* in the presence of microsomes. This approach is now widely used to study topology (Zhang and Ling, 1991; Zhang *et al.*, 1995).

Other methods include chemical labeling with membrane-impenetrable reagents, proteolysis of various membrane preparations, and the localization of epitopes for peptide-specific antibodies by ELISA or electron microscopy (Jennings, 1989; Contifine *et al.*, 1996). In general, however, such methods rarely give clear-cut results and are often hard to interpret.

V. MEMBRANE PROTEIN INSERTION IN BACTERIA: TWO DIFFERENT MECHANISMS?

Certain aspects of membrane protein structure can only be properly understood in the light of how the proteins are inserted into the lipid bilayer. For many prokaryotic and eukaryotic membrane proteins, the cellular machineries responsible for translocating proteins through membranes are needed for membrane insertion; no great surprise, perhaps, since integral membrane proteins contain domains that need to be translocated through the bilayer.

However, at least in bacteria, there are also inner membrane proteins that do not appear to need help from the translocation machinery for insertion. There thus seem to exist at least two different insertion mechanisms—generally referred to as Sec-dependent and Sec-independent—that may apply to different proteins, or even to different domains within the same protein (Lee *et al.*, 1992b; von Heijne, 1994b).

5.1. *The Sec Machinery: Moving Proteins through Membranes*

Protein translocation across the inner membrane of *E. coli* is carried out by a machinery composed of cytoplasmic components—Ffh, FtsY, SecB, and GroEL/ES—and the inner membrane components SecA, SecY, SecE, SecG, SecD, and SecF (Ito, 1995). In addition, the leader peptidase enzyme (Lep) located in the inner membrane is necessary for removing signal peptides from secretory proteins, while the lipoprotein signal peptidase (Lsp) does the same for lipoproteins (Dalbey and von Heijne, 1992; von Heijne, 1994c).

Chaperones such as Ffh, SecB, and GroEL/ES ensure that nascent secretory proteins do not fold or aggregate prematurely, and also appear to help in targeting the pre-proteins to the Sec machinery proper (Hartl *et al.*, 1990; Hardy and Randall, 1991; Randall, 1992; Luirink and Dobberstein, 1994; Diamond *et al.*, 1995). SecA is peripherally bound to the inner membrane, yet can carry out an ATP-driven cycle of insertion and de-insertion into the membrane sector of the translocase during which some 20–30 residues of the nascent chain are moved through the membrane (Economou and Wickner, 1994; Uchida *et al.*, 1995). The integral membrane components SecY/E/G are thought to form a channel

through which the nascent chain is extruded, and may move in concert with SecA during its insertion/de-insertion cycle (Nishiyama *et al.*, 1996). The functions of the likewise membrane-bound SecD and SecF components are less clear, but they are thought to regulate the SecA cycle (Economou *et al.*, 1995).

In addition to secretory proteins, all outer membrane and certain inner membrane proteins also require the Sec machinery for translocation across or insertion into the inner membrane. Inner membrane proteins with large periplasmic domains, either made with an N-terminal signal-anchor sequence such as the chemoreceptors (Tar, Tap, Tsr), or with internal, signal-anchor sequences (Lep, MalF) are Sec-dependent (Wolfe *et al.*, 1985; Gebert *et al.*, 1988; Sääf *et al.*, 1995; Traxler and Murphy, 1996), whereas proteins with only short periplasmic loops or tails in most cases seem to be Sec-independent (Wolfe *et al.*, 1985; von Heijne, 1989; Andersson and von Heijne, 1993). Besides length, the number of charged residues in a translocated loop also seems to affect its degree of Sec-dependence (Andersson and von Heijne, 1994b; Cao *et al.*, 1995).

For Sec-dependent inner membrane proteins, translocation through the Sec translocase is initiated either by a cleavable, N-terminal signal peptide or by an N-terminal or internal signal-anchor sequence (Fig. 11). Translocation of the downstream polar region then continues until a sufficiently hydrophobic stretch of chain—a stop-transfer sequence—enters the translocase. At this point translocation is aborted, and the part of the chain immediately following the stop-transfer sequence becomes localized to the cytoplasm. Membrane translocation may be re-initiated by a second signal-anchor sequence further downstream, again be aborted by a second stop-transfer sequence, etc., until the end of the chain is reached. For Sec-dependent proteins, membrane insertion is thus thought to occur in a sequential, N-to-C-terminal fashion. The hydrophobic, membrane-spanning signal-anchor and stop-transfer sequences must at some point be expelled laterally out of the translocase into the lipid bilayer; the timing of this event is unknown.

Compared to Sec-dependent secretion of proteins, the new elements in the membrane protein insertion process are the recognition of the stop-transfer sequence and its expulsion into the bilayer. Very little is known about how this happens, but it is clear that the overall hydrophobicity of the stop-transfer sequence is the most important functional determinant (Chen and Kendall, 1995). Positively charged residues on the C-terminal end

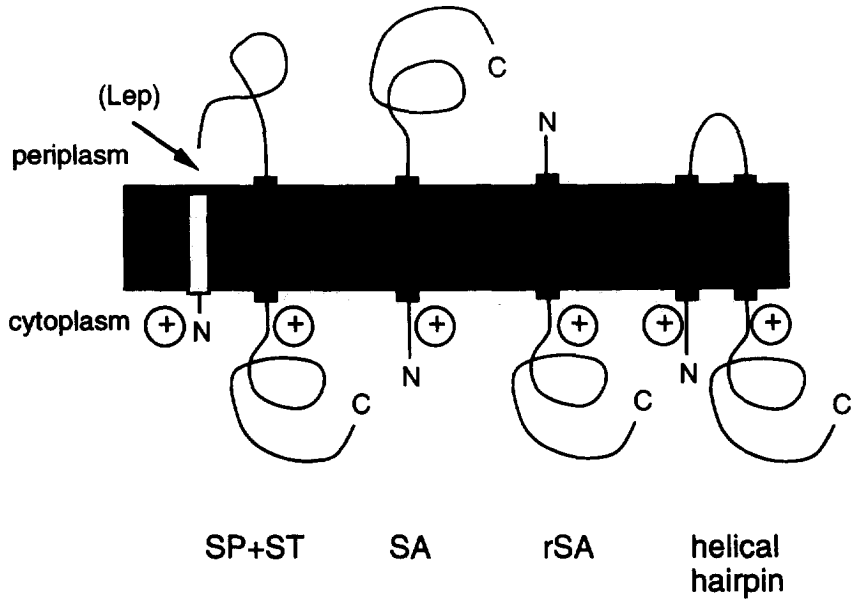

Fig. 11. Topogenic signals in membrane proteins. SP = signal peptide cleaved by signal peptidase (Lep in *E. coli*), ST = stop-transfer sequence, SA = signal-anchor sequence, rSA = reverse signal-anchor sequence. Note that all topogenic signals consist of a hydrophobic segment flanked on its cytoplasmic side by positively charged residues.

of the stop-transfer sequence may also contribute to its function, although this has so far only been demonstrated for stop-transfer function during translocation across the ER membrane (Kuroiwa *et al.*, 1991).

5.2. *Sec-independent Membrane Insertion: Positive Charges and Helical Hairpins*

Sec-independent insertion of a membrane protein is operationally defined by the observation that expression under non-permissive conditions in strains carrying conditionally lethal defects in one or other of the Sec components (SecA, SecY, SecE) or in the presence of sodium azide (known to block the ATPase function of SecA; Oliver *et al.*, 1990) has no measurable effect on the efficiency of insertion (von Heijne, 1994b). This in itself does not prove that the Sec machinery is not at all involved; however, the recent finding that yeast mitochondria do not have a Sec machinery, yet can insert proteins into the inner membrane (see below) strongly suggests that at least some bacterial inner membrane proteins are truly Sec-independent.

So far, Sec-independent translocation has been seen for two kinds of periplasmic regions: short loops (Andersson and von Heijne, 1993) and both short and long N-terminal tails (Cao and Dalbey, 1994; Whitley *et al.*, 1994b). In all cases, the translocated region contains only few positively charged amino acids but may contain many negatively charged ones; in contrast, the cytoplasmic loops in these proteins are enriched in positively charged residues (von Heijne, 1986). This bias in the distribution of positively charged residues between the cytoplasmic and periplasmic regions is often referred to as the "positive inside" rule (von Heijne and Gavel, 1988).

It has been shown that the membrane electrochemical potential $\Delta\mu_H^+$ (acidic and positive in the periplasm) influences the topological effects of charged residues. A short periplasmic segment in the phage M13 coat protein needs the $\Delta\mu_H^+$ for translocation only when it carries four or more acidic residues, but not otherwise (Cao *et al.*, 1995), and studies on derivatives of Lep have shown that negatively charged residues become equally strong topological determinants as positively charged ones in the absence of $\Delta\mu_H^+$ (Andersson and von Heijne, 1994a). Although the individual roles of ΔpH and $\Delta\psi$ have not been studied, these results suggest that there may be an electrophoretic effect behind the positive inside rule.

In contrast to the Sec-dependent insertion mechanism, Sec-independent insertion does not seem to necessarily proceed in an N-to-C-terminal direction; rather, it appears to be based on helical hairpins (Engelman and Steitz, 1981) composed of two neighboring hydrophobic segments with flanking positively charged residues and a connecting loop containing few such residues (Fig. 11). Helical hairpins are thought to compete with each other for insertion into the bilayer, as shown by the behavior of so-called "frustrated" inner membrane proteins: artificial constructs with four hydrophobic segments where the distribution of positively charged residues is such that any topology with four transmembrane segments will have at least one highly charged loop in the periplasm (Gafvelin and von Heijne, 1994).

In summary, Sec-independent translocation across the inner membrane has been seen both for N-terminal tails and short periplasmic loops of inner membrane proteins. Presumably, short C-terminal tails may also translocate independently of the Sec machinery, though this has not been demonstrated. Positively charged residues prevent Sec-independent translocation, and this effect is at least in part dependent on the $\Delta\mu_H^+$ across the inner membrane. The basic unit in Sec-independent insertion appears to be the helical hairpin. It may be that Sec-independent proteins insert spontaneously into the lipid bilayer, though this has not been formally proven. It seems likely that cytoplasmic chaperones are needed to prevent inner membrane proteins from aggregating prior to membrane insertion, though again a clear demonstration is lacking.

5.3. *Outer Membrane Proteins: Concerted Folding and Insertion?*

Outer membrane proteins in Gram-negative bacteria are made with an N-terminal signal peptide, and are initially translocated across the inner membrane by the Sec

machinery. Mutant versions of the outer membrane protein OmpA are found as soluble species in the periplasm or loosely bound to the outer membrane (Bosch *et al.*, 1986; Klose *et al.*, 1988; 1993) and the periplasmic chaperones Skp and SurA have been found to be involved in the folding of certain outer membrane proteins (Chen and Henning, 1996; Lazar and Kolter, 1996), suggesting that integration into the outer membrane proceeds via a soluble periplasmic intermediate.

Folding and insertion of outer membrane proteins into lipid vesicles can be reconstituted *in vitro* in the presence of low levels of detergent, and the current view is that the formation of the β-barrel with its large hydrophobic surface happens just prior to, or in concert with, membrane insertion (Surrey and Jähnig, 1992; van Gelder *et al.*, 1994; Dekker *et al.*, 1995; De Cock *et al.*, 1996; Surrey *et al.*, 1996). How the membrane orientation of outer membrane proteins is controlled is not known, though it is striking that the loops protruding into the periplasm are much shorter than the extracellular ones in the known 3D structures.

VI. MITOCHONDRIAL MEMBRANE PROTEIN BIOGENESIS: MAKING DO WITHOUT SEC

From evolutionary considerations, protein insertion into the membranes of organelles such as mitochondria and chloroplasts would be expected to have conserved many features of the bacterial system. Indeed, Sec-homologues have been found in chloroplasts (Douglas, 1992; Scaramuzzi *et al.*, 1992a,b; Nakai *et al.*, 1993; 1994; Berghöfer *et al.*, 1995; Laidler *et al.*, 1995), and both Sec-dependent and Sec-independent thylakoid membrane proteins have been shown to exist (Mant *et al.*, 1994; van Wijk *et al.*, 1995). Thylakoid membrane proteins also have more positively charged residues in the loops facing the stroma than in those facing the lumen (Gavel *et al.*, 1991), consistent with the existence of a Sec-independent insertion mechanism.

Initially, it was expected that there would also be a mitochondrial Sec machinery (Hartl and Neupert, 1990), but no Sec components have been found in any of the fully sequenced mitochondrial genomes (except for *Reclinomonas americana*) or in the fully sequenced nuclear genome of *Saccharomyces cerevisiae* (Glick and von Heijne, 1996). Two proteases distantly related to Lep have been found in the mitochondrial intermembrane space, though. Unless the putative Sec components have diverged beyond the point of recognition by sequence similarity (which is unlikely, considering the good conservation of Sec components between eubacteria, archaebacteria, and chloroplasts), the implication is that all inner membrane proteins encoded in the mitochondrial genome must insert by a Sec-independent mechanism. Indeed, just as the thylakoid membrane proteins, mitochondrially encoded inner membrane proteins also conform to the positive inside rule, with more positively charged residues in their matrix-facing loops. This is not generally true for the nuclearly encoded, imported inner membrane proteins, suggesting that at least some of these proteins may be inserted into the bilayer using a different mechanism (Gavel and von Heijne, 1992).

The absence of a mitochondrial Sec machinery suggests that there must be strong constraints on the kinds of proteins that can be inserted into or translocated across the inner membrane from the matrix side: in essence, only Sec-independent proteins would qualify. Indeed, membrane assembly of the mitochondrially encoded COXII protein has been found to have characteristics reminiscent of Sec-independent assembly in *E. coli* (Herrmann *et al.*, 1995). Beyond the mitochondrially encoded proteins, only imported proteins with short segments containing few positively charged residues in the intermembrane space (IMS) would be expected to be able to insert from the matrix side. This seems to be the case for proteins such as subunit 9 from the F_1F_0 ATPase (Rojo *et al.*, 1995) and the Rieske FeS protein (Ostermann *et al.*, 1989), which are encoded in the nuclear genome, translated in the cytoplasm, imported through both mitochondrial membranes, and then inserted back into the inner membrane. These proteins thus use a "conservative sorting" pathway (Hartl and Neupert, 1990), where the final membrane

insertion step appears to be equivalent to Sec-independent membrane insertion in bacteria.

How nuclearly encoded inner membrane proteins that do not conform to the positive inside rule (e.g., many of the small subunits of the mitochondrial cytochrome *c* oxidase complex; Tsukihara *et al.*, 1996) are inserted is unclear at present, though it appears that their hydrophobic transmembrane segments serve as stop-transfer sequences that cannot be pulled through the inner membrane translocation machinery and move sideways out of the translocase into the bilayer (Glick *et al.*, 1992; Rospert *et al.*, 1994; Gärtner *et al.*, 1995).

VII. MEMBRANE PROTEINS IN THE ER: ESCAPING FROM THE PROTEIN-CONDUCTING CHANNEL

In eukaryotic cells, most integral membrane proteins destined for one or other of the membranes along the secretory pathway are initially inserted into the ER membrane, from which they are further routed by vesicular transport (Rothman, 1996). A machinery related to the bacterial Sec complex is responsible for translocating secretory proteins into the lumen of the ER (Jungnickel *et al.*, 1994; Schatz and Dobberstein, 1996), and is also utilized by the integral membrane proteins (High, 1995; Oliver *et al.*, 1995). The only clear example of "Sec-independent" insertion into the ER is for the tail-anchored protein synaptobrevin, in which a short C-terminal tail at the end of its only transmembrane segment is translocated independently of the normal targeting and translocation machineries (Kutay *et al.*, 1995; Whitley *et al.*, 1996).

7.1. *Targeting and Translocation Across the ER Membrane: a GTP-controlled Machine*

Both secretory and membrane proteins are targeted to the ER through an initial interaction with the cytoplasmic signal recognition particle (SRP) that primes the ribosome–nascent chain complex for further interactions with the SRP receptor on the cytoplasmic face of the ER, followed by transfer to the Sec61p translocation complex (Rapoport, 1992). The ordered binding and recognition steps along this pathway are controlled by GTP binding, hydrolysis, and GDP–GTP exchange (Connolly and Gilmore, 1993; Miller *et al.*, 1993; Rapiejko and Gilmore, 1994; Bacher *et al.*, 1996).

Once the ribosome–nascent chain complex is bound to the Sec61p translocase, the chain is translocated co-translationally to the lumen where it is received by chaperones such as the hsp70 protein BiP that may help to pull the chain across the membrane (Schatz and Dobberstein, 1996). In addition to BiP, folding and oligomerization in the ER often depend on other chaperones such as the calnexin/calreticulin system (Helenius, 1994; Pind *et al.*, 1994; Gelman *et al.*, 1995; Peterson *et al.*, 1995).

7.2. *Weaving Proteins into a Lipid Bilayer: Signal-anchors and Stop-transfer Signals*

As for Sec-dependent membrane protein assembly in *E. coli*, the insertion of integral membrane proteins into the ER is controlled by four kinds of topogenic signals: signal peptides, signal-anchor sequences, reverse signal-anchor sequences, and stop-transfer sequences (Fig. 11). Signal peptides target proteins to the ER and initiate translocation of the following portion of the nascent chain, signal-anchor sequences have the same function but are not cleaved and end up anchoring the protein to the bilayer, reverse signal-anchor sequences also target to the ER but become oriented with their N-terminus in the lumen, and stop-transfer sequences abort ongoing translocation and then become embedded in the bilayer. Proteins with more than one transmembrane segment are thought to be weaved into the membrane starting from the most N-terminal topogenic signal in a co-translational, sequential process (Wessels and Spiess, 1988; Lipp *et al.*, 1989), at least when the loops connecting the transmembrane segments are long (Gafvelin *et al.*, 1997). Hydrophobic segments and positively charged flanking residues are the major features also of the eukaryotic topogenic signals, and the topological effects of

charged residues are similar though not identical in the ER and in *E. coli* (von Heijne and Manoil, 1990; Gafvelin *et al.*, 1997).

Recently, the steps through which a stop-transfer sequence is released from the translocase into the surrounding lipid have been studied in some detail using site-directed photochemical crosslinking (Do *et al.*, 1996). An ordered sequence of steps has been postulated, starting with the hydrophobic stretch entering the putative Sec61α-channel, then being progressively moved out of the channel into a location where it contacts the TRAM protein (and possibly also phospholipids), and finally being released into the bilayer upon termination of translation. The possibility of extracting ribosome-bound nascent chains but not fully synthesized proteins containing multiple transmembrane segments from the ER with urea has also been taken as an indication that integral membrane proteins are not fully released into the lipid bilayer until the ribosome disassembles from the translocase (Borel and Simon, 1996).

Generally speaking, insertion of membrane proteins into the ER is thus in many ways analogous to Sec-dependent insertion into the inner membrane of *E. coli*. In fact, sequences that as act as stop-transfer signals in one system often do so in the other (Sääf *et al.*, 1997), and heterologously expressed membrane proteins often adopt the same topology in *E. coli* and in the ER (Hennessey *et al.*, 1993). Whether there is a distinction in eukaryotic cells similar to the one between Sec-dependent and Sec-independent insertion in *E. coli* is not clear, except for the tail-anchored proteins mentioned above. One difference between prokaryotic and eukaryotic membrane proteins that may be relevant to this question is the observation that positively charged residues in lumenal loops in eukaryotic proteins are not as strongly selected against as in bacterial proteins (Wallin and von Heijne, 1995), suggesting that lumenal loops may depend on the Sec61p machinery for translocation.

VIII. THE PREDICTION GAME: FROM SEQUENCE TO TOPOLOGY TO STRUCTURE

A good test of the current level of knowledge in a field is the ability to make quantitative predictions. For membrane proteins, this means predictions of topology and structure from the amino acid sequence. Due to the rather strict structural constraints imposed by the lipid environment (basically the requirement that all hydrogen bonds in the lipid-embedded part of the molecule must be satisfied internally), it has long been thought that structural predictions will be simpler for membrane proteins than for globular ones, and, as the following discussion shows, this is probably true. A second simplification possible for membrane proteins of the helix-bundle class but not for β-barrel membrane proteins or for globular proteins is the assumption that membrane insertion and helix–helix packing are strictly separable events: according to the "two stage model" (Popot and de Vitry, 1990; Popot and Engelman, 1990), all transmembrane segments first insert into the bilayer as individually stable helices and these pre-formed helices then pack together.

8.1. *Predicting Topology: Hydrophobicity Analysis and the Positive Inside Rule*

Topology predictions start from two basic observations: (I) transmembrane helices are 20–30 residues long and have a high overall hydrophobicity, and (II) short non-translocated loops contain many positively charged residues whereas short translocated loops contain few such residues (or, more generally, the average amino acid composition is different between translocated and non-translocated loops). Observation (I) is the basis for the identification of transmembrane segments from hydrophobicity plots (Kyte and Doolittle, 1982), and observation (II) makes it possible to predict the orientation of a protein in the membrane, and even to choose the most likely topology when the identification of the transmembrane segments from the hydrophobicity plot is uncertain (von Heijne, 1992).

Topology prediction methods can be more or less sophisticated in terms of how the underlying hydrophobicity scale has been derived (Cornette *et al.*, 1987; Cserzo *et al.*,

1994; Samatey *et al.*, 1995), in terms of the algorithm used (von Heijne, 1992; Jones *et al.*, 1994; Rost *et al.*, 1995; 1996), and in terms of whether or not information from homologous sequences is taken into account (Rost *et al.*, 1995; Persson and Argos, 1996).

The best current methods claim that >90% of all transmembrane segments can be correctly identified and that the full topology is correctly predicted for >80% of all proteins (von Heijne, 1992; Claros and von Heijne, 1994; Persson and Argos, 1996; Rost *et al.*, 1996). It seems that predictions are slightly better for prokaryotic than for eukaryotic proteins, probably in part because better experimental methods to determine topology such as fusion protein analysis have made the database of known topologies larger, and in part because prokaryotic membrane proteins tend to have more hydrophobic transmembrane helices and shorter extra-membraneous loops that conform better to the positive inside rule.

Topology prediction of β-barrel membrane proteins is more difficult since the membrane-spanning β-strands are short and thus hard to detect in the sequence. Reasonable predictions can be made when the protein can be aligned with the sequence of a protein of known structure (Welte *et al.*, 1991). If this is not possible, one can look for short stretches of chain where every second residue is hydrophobic, where the turn potential is low, and that end with aromatic residues (Schirmer and Cowan, 1993), though this only works if one already has a good idea of which part of the protein forms the β-barrel (Nakai and Kanehisa, 1991).

8.2. *3D Structure Prediction: Packing Helices in the Computer*

A more demanding task is to predict helix–helix packing and 3D structure *ab initio*. Two different approaches have been tried so far: simulated annealing/energy minimization starting from a large number of systematically generated starting conformations (Treutlein *et al.*, 1992; Arkin *et al.*, 1994; Lemmon *et al.*, 1994; Tuffery *et al.*, 1994; Adams *et al.*, 1995), or an initial prediction of the most likely lipid-exposed and buried helix faces (Baldwin, 1993; Suwa *et al.*, 1995; Efremov and Vergoten, 1996) followed by a combinatorial packing algorithm (Taylor *et al.*, 1994).

As the examples of the glycophorin A dimer and the phospholamban pentamer show (Treutlein *et al.*, 1992; Adams *et al.*, 1995), structure prediction *ab initio* is still not possible even for such relatively simple and highly symmetrical systems, and experimental information from, e.g., mutagenesis data are needed to identify the most likely model among the set of calculated low-energy structures. An interesting procedure for the systematic use of the available experimental data has recently been proposed in conjunction with the modeling of G-protein-coupled receptors (Herzyk and Hubbard, 1995).

IX. THE FUTURE: CRYSTALLIZATION, CRYSTALLIZATION, CRYSTALLIZATION ...

Where, finally, is the membrane protein field moving? The basic rules of membrane protein topology seem to be in place, although many interesting mechanistic questions relating to the insertion of membrane proteins into lipid bilayers *in vivo* remain. Three-dimensional structure prediction is still not possible unless there is a fair amount of experimental information to guide the search, and there is an acute need for more high-resolution structures to provide a database for detailed theoretical studies of helix–helix packing.

High-resolution structures are of course also indispensable for understanding the functions of membrane proteins—many of which are of central importance both for cell biology and for the pharmaceutical industry—and overexpression, purification, and crystallization of membrane proteins are now perceived as one of the major challenges in structural biology. It seems a fair bet that crystallization of membrane proteins, although still a very risky undertaking, will be on many people's agenda during the next decade.

ACKNOWLEDGEMENTS

This work was supported by grants from the Swedish Natural Sciences Research Council (NFR), the Swedish Technical Sciences Research Council (TFR), the Swedish Cancer Foundation, and the Göran Gustafsson Foundation.

REFERENCES

Adams, P. D., Arkin, I. T., Engelman, D. M. and Brünger, A. T. (1995) Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nature struct. Biol.* **2,** 154–162.

Allen, J., Feher, G., Yeates, T., Komiya, H. and Rees, D. (1987) Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The protein subunits. *Proc. natn. Acad. Sci. U.S.A.* **84,** 6162–6166.

Altenbach, C., Flitsch, S., Khorana, H. and Hubbell, W. (1989) Structural studies in transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry* **28,** 7806–7812.

Altenbach, C., Greenhalg, D. A., Khorana, H. G. and Hubbell, W. L. (1994) A collision gradient method to determine the immersion depth of nitroxides in lipid bilayers: Application to spin-labeled mutants of bacteriorhodopsin. *Proc. natn. Acad. Sci. U.S.A.* **91,** 1667–1671.

Altenbach, C., Marti, T., Khorana, H. G. and Hubbell, W. L. (1990) Transmembrane protein structure—spin labeling of bacteriorhodopsin mutants. *Science* **248,** 1088–1092.

Andersson, H. and von Heijne, G. (1993) *Sec*-dependent and *sec*-independent assembly of *E. coli* inner membrane proteins—the topological rules depend on chain length. *Eur. molec. Biol. Org. J.* **12,** 683–691.

Andersson, H. and von Heijne, G. (1994) Membrane protein topology: Effects of DmH^+ on the translocation of charged residues explain the "positive inside" rule. *Eur. molec. Biol. Org. J.* **13,** 2267–2272.

Andersson, H. and von Heijne, G. (1994) Positively charged residues influence the degree of sec-dependence in protein translocation across the *E. coli* inner membrane. *Fedn Eur. biochem. Socs Lett.* **347,** 169–172.

Arkin, I. T., Adams, P. D., MacKenzie, K. R., Lemmon, M. A., Brünger, A. T. and Engelman, D. M. (1994) Structural organization of the pentameric transmembrane α-helices of phospholamban, a cardiac ion channel. *Eur. molec. Biol. Org. J.* **13,** 4757–4764.

Arkin, I. T., MacKenzie, K. R., Fisher, L., Aimoto, S., Engelman, D. M. and Smith, S. O. (1996) Mapping the lipid-exposed surfaces of membrane proteins. *Nature struct. Biol.* **3,** 240–243.

Arkin, I. T., Rothman, M., Ludlam, C. F. C., Aimoto, S., Engelman, D. M., Rothschild, K. J. and Smith, S. O. (1995) Structural model of the phospholamban ion channel complex in phospholipid membranes. *J. molec. Biol.* **248,** 824–834.

Bacher, G., Lutcke, H., Jungnickel, B., Rapoport, T. A. and Dobberstein, B. (1996) Regulation by the ribosome of the GTPase of the signal- recognition particle during protein targeting. *Nature* **381,** 248–251.

Baldwin, J. M. (1993) The probable arrangement of the helices in G protein-coupled receptors. *Eur. molec. Biol. Org. J.* **12,** 1693–1703.

Bechinger, B., Zasloff, M. and Opella, S. J. (1993) Structure and orientation of the antibiotic peptide magainin in membranes by solid-state nuclear magnetic resonance spectroscopy. *Protein Sci.* **2,** 2077–2084.

Berghöfer, J., Karnauchov, I., Herrmann, R. G. and Klösgen, R. B. (1995) Isolation and characterization of a cDNA encoding the SecA protein from spinach chloroplasts—Evidence for azide resistance of Sec-dependent protein translocation across thylakoid membranes in spinach. *J. biol. Chem.* **270,** 18,341–18,346.

Borel, A C. and Simon, S. M. (1996) Biogenesis of polytopic membrane proteins: Membrane segments assemble with translocation channels prior to membrane integration. *Cell* **85,** 379–389.

Bosch, D., Leunissen, J., Verbakel, J., de Jong, M., van Erp, H. and Tommassen, J. (1986) Periplasmic accumulation of truncated forms of outer-membrane PhoE protein of *Escherichia coli* K-12. *J. molec. Biol.* **189,** 449–455.

Boyd, D., Traxler, B. and Beckwith, J. (1993) Analysis of the topology of a membrane protein by using a minimum number of alkaline phosphatase fusions. *J. Bacteriol.* **175,** 553–556.

Braun, P., Persson, B., Kaback, H. R. and von Heijne, G. (1997) Alanine insertion scanning mutagenesis of lactose permease transmembrane helices. Identification of a structurally and functionally important segment in helix III, (*submitted*).

Cao, G. Q. and Dalbey, R. E. (1994) Translocation of N-terminal tails across the plasma membrane. *Eur. molec. Biol. Org. J.* **13,** 4662–4669.

Cao, G. Q., Kuhn, A. and Dalbey, R. E. (1995) The translocation of negatively charged residues across the membrane is driven by the electrochemical potential: Evidence for an electrophoresis-like membrane transfer mechanism. *Eur. molec. Biol. Org. J.* **14,** 866–875.

Chen, H. F. and Kendall, D. A. (1995) Artificial transmembrane segments—Requirements for stop transfer and polypeptide orientation. *J. biol. Chem.* **270,** 14,115–14,122.

Chen, R. and Henning, U. (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Molec. Microbiol.* **19,** 1287–1294.

Claros, M. G. and von Heijne, G. (1994) TopPred II: An improved software for membrane protein structure prediction. *CABIOS* **10,** 685–686.

Connolly, T. and Gilmore, R. (1993) GTP hydrolysis by complexes of the signal recognition particle and the signal recognition particle receptor. *J. Cell Biol.* **123,** 799–807.

Contifine, B. M., Lei, S. J. and McLane, K. E. (1996) Antibodies as tools to study the structure of membrane proteins: The case of the nicotinic acetylcholine receptor. *A. Rev. Biophys. biomolec. Struct.* **25,** 197–229.

Cornette, J. L., Cease, K. B., Margalit, H., Spouge, J. L., Berzofsky, J. A. and De Lisi, C. (1987) Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J. molec. Biol.* **195,** 659–685.

Cowan, S. W. and Rosenbusch, J. P. (1994) Folding pattern diversity of integral membrane proteins. *Science* **264,** 914–916.

Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Joansonius, J. N. and Rosenbusch, J. P. (1992) Crystal structures explain functional properties of two *E. coli* porins. *Nature* **358,** 727–733.

Cserzo, M., Bernassau, J. M., Simon, I. and Maigret, B. (1994) New alignment strategy for transmembrane proteins. *J. molec. Biol.* **243,** 388–396.

Dalbey, R. E. and von Heijne, G. (1992) Signal peptidases in prokaryotes and eukaryotes—A new protease family. *Trends biochem. Sci.* **17,** 474–478.

De Cock, H., Van Blokland, S. and Tommassen, J. (1996) *In vitro* insertion and assembly of outer membrane protein PhoE of *Escherichia coli* K-12 into the outer membrane—Role of Triton X-100. *J. biol. Chem.* **271,** 12,885–12,890.

Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. (1985) Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution. *Nature* **318,** 618–624.

Dekker, N., Merck, K., Tommassen, J. and Verheij, H. M. (1995) *In vitro* folding of *Escherichia coli* outer-membrane phospholipase A. *Eur. J. Biochem.* **232,** 214–219.

Diamond, D. L., Strobel, S., Chun, S. Y. and Randall, L. L. (1995) Interaction of SecB with intermediates along the folding pathway of maltose-binding protein. *Protein Sci.* **4,** 1118–1123.

Do, H., Falcone, D., Lin, J., Andrews, D. W. and Johnson, A. E. (1996) The cotransational integration of membrane proteins into the phospholipid bilayer is multistep process. *Cell* **85,** 369–378.

Douglas, S. E. (1992) A secY homologue is found in the plastid genome of *Cryptomonas φ*. *Fedn Eur. biochem. Socs Lett.* **298,** 93–96.

Dunten, R. L., Sahintoth, M. and Kaback, H. R. (1993) Role of the charge pair aspartic acid-237-lysine-358 in the lactose permease of *Escherichia coli*. *Biochemistry* **32,** 3139–3145.

Economou, A., Pogliano, J. A., Beckwith, J., Oliver, D. B. and Wickner, W. (1995) SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell* **83,** 1171–1181.

Economou, A. and Wickner, W. (1994) SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* **78,** 835–843.

Efremov, R. G. and Vergoten, G. (1996) Hydrophobic organization of alpha-helix membrane bundle in bacteriorhodopsin. *J. Protein Chem.* **15,** 63–76.

Ehrmann, M., Boyd, D. and Beckwith, J. (1990) Genetic analysis of membrane protein topology by a sandwich gene fusion approach. *Proc. natn. Acad. Sci. U.S.A.* **87,** 7574–7578.

Elofsson, A., Wallin, E., Tsukihara, T., Yoshikawa, S. and von Heijne, G. (1997) Anatomy of helix bundle membrane proteins, *Protein Sci., in press.*

Engelman, D. M. and Steitz, T. A. (1981) The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis. *Cell* **23,** 411–422.

Gafvelin, G., Sakaguchi, M., Andersson, H. and von Heijne, G. (1997) Topological rules for membrane protein assembly in eukaryotic cells, *J. biol. Chem., in press.*

Gafvelin, G. and von Heijne, G. (1994) Topological "frustration" in multi-spanning *E. coli* inner membrane proteins. *Cell* **77,** 401–412.

Gärtner, F., Bömer, U., Guiard, B. and Pfanner, N. (1995) The sorting signal of cytochrome b(2) promotes early divergence from the general mitochondrial import pathway and restricts the unfoldase activity of matrix Hsp70. *Eur. molec. Biol. Org. J.* **14,** 6043–6057.

Gavel, Y., Steppuhn, J., Herrmann, R. and von Heijne, G. (1991) The positive-inside rule applies to thylakoid membrane proteins. *Fedn Eur. biochem. Socs Lett.* **282,** 41–46.

Gavel, Y. and von Heijne, G. (1992) The distribution of charged amino acids in mitochondrial inner membrane proteins suggests different modes of membrane integration for nuclearly and mitochondrially encoded proteins. *Eur. J. Biochem.* **205,** 1207–1215.

Gebert, J. F., Overhoff, B., Manson, M. and Boos, W. (1988) The Tsr chemosensory transducer of *Escherichia coli* assembles into the cytoplasmic membrane via a Seca-dependent process. *J. biol. Chem.* **263,** 16,652–16,660.

Gelman, M. S., Chang, W. S., Thomas, D. Y., Bergeron, J. J. M. and Prives, J. M. (1995) Role of the endoplasmic reticulum chaperone calnexin in subunit folding and assembly of nicotinic acetylcholine receptors. *J. biol. Chem.* **270,** 15,085–15,092.

Glick, B. S., Brandt, A., Cunningham, K., Muller, S., Hallberg, R. L. and Schatz, G. (1992) Cytochromes-c1 and cytochromes-b2 are sorted to the intermembrane space of yeast mitochondria by a stop-transfer mechanism. *Cell* **69,** 809–822.

Glick, B. S. and von Heijne. G. (1996) *Saccharomyces cerevisiae* mitochondria lack a bacterial-type Sec machinery. *Protein Sci.* **5,** 2651–2652.

Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. and Henderson, R. (1996) Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. molec. Biol.* **259,** 393–421.

Grisshammer, R. and Tate, C. G. (1995) Overexpression of integral membrane proteins for structural studies. *Q. Rev. Biophys.* **28,** 315–422.

Hardy, S. J. S. and Randall, L. L. (1991) A kinetic partitioning model of selective binding of nonnative proteins by the bacterial chaperone SecB. *Science* **251,** 439–443.

Hartl, F. U., Lecker, S., Schiebel, E., Hendrick, J. P. and Wickner, W. (1990) The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. Coli* plasma membrane. *Cell* **63,** 269–279.

Hartl, F. U. and Neupert, W. (1990) Protein sorting to mitochondria—Evolutionary conservations of folding and assembly. *Science* **247,** 930–938.

He, M. M., Voss, J., Hubbell, W. L. and Kaback, H. R. (1995) Use of designed metal-binding sites to study helix proximity in the lactose permease of *Escherichia coli*. 1. Proximity of helix VII (Asp237 and Asp240) with helices X (Lys319) and XI (Lys358). *Biochemistry* **34,** 15,661–15,666.

He, M. M., Voss, J., Hubbell, W. L. and Kaback, H. R. (1995) Use of designed metal-binding sites to study helix proximity in the lactose permease of *Escherichia coli*. 2. Proximity of helix IX (Arg302) with helix X (His322 and Glu325). *Biochemistry* **34,** 15667–15670.

Helenius, A. (1994) How N-Linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Molec. Biol. Cell* **5,** 253–265.

Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E. and Downing, K. H. (1990) A model for the structure of bacteriorhodopsin based on high resolution electron cryo-microscopy. *J. molec. Biol.* **213,** 899–929.

Henderson, R. and Unwin, P. N. T. (1975) Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257,** 28–32.

Hennessey, E. S. and Broome-Smith, J. K. (1993) Gene-fusion techniques for determining membrane-protein topology. *Curr. Opin. struct. Biol.* **3,** 524–531.

Hennessey, E. S., Hashemzadeh-Bonehi, L., Hunt, L. A. and Broome-Smith, J. K. (1993) Assembly of eukaryotic class-III (N-out, C-in) membrane proteins into the *Escherichia coli* cytoplasmic membrane. *Fedn Eur. biochem. Socs Lett.* **331,** 159–161.

Herrmann, J. M., Koll, H., Cook, R. A., Neupert, W. and Stuart, R. A. (1995) Topogenesis of cytochrome oxidase subunit II—Mechanisms of protein export from the mitochondrial matrix. *J. biol. Chem.* **270,** 27079–27086.

Herzyk, P. and Hubbard, R. E. (1995) Automated method for modeling seven-helix transmembrane receptors from experimental data. *Biophys. J.* **69,** 2419–2442.

High, S. (1995) Protein translocation at the membrane of the endoplasmic reticulum. *Prog. Biophys. molec. Biol.* **63,** 233–250.

Ito, K. (1995) Protein translocation genetics. In *Membrane Organization and Biogenesis in Prokaryotic Organisms*, (ed. R. D. Dalbey), JAI Press.

Iwata, S., Ostermeier, C., Ludwig, B. and Michel, H. (1995) Structure at 2.8 Å resolution of cytochrome c oxidase from *Paracoccus denitrificans*. *Nature* **376,** 660–669.

Jennings, M. L. (1989) Topography of membrane proteins. *A. Rev. Biochem.* **58,** 999–1027.

Jones, D. T., Taylor, W. R. and Thornton, J. M. (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33,** 3038–3049.

Jung, K., Jung, H., Wu, J., Prive, G. G. and Kaback, H. R. (1993) Use of site-directed fluorescence labeling to study proximity relationships in the lactose permease of *Escherichia coli*. *Biochemistry* **32,** 12,273–12,278.

Jung, K., Voss, J., He, M., Hubbell, W. L. and Kaback, H. R. (1995) Engineering a metal binding site within a polytopic membrane protein, the lactose permease of *Escherichia coli*. *Biochemistry* **34,** 6272–6277.

Jungnickel, B., Rapoport, T. A. and Hartmann, E. (1994) Protein translocation: Common themes from bacteria to man. *Fedn Eur. biochem. Socs Lett.* **346,** 73–77.

Kaback, H. R. (1996) The lactose permease of *Escherichia coli*: Past, present, and future. In *Handbook of biological physics: Transport processes in eukaryotic and prokaryotic organisms*, (eds. W. N. Konings and H. R. Kaback), Elsevier, Amsterdam (*in press*).

Kaback, H. R., Frillingos, S., Jung, H., Jung, K., Prive, G. G., Ujwal, M. L., Weitzman, C., Wu, J. and Zen, K. (1994) The lactose permease meets Frankenstein. *J. exp. Biol.* **196,** 183–195.

Kelleher, D. J., Kreibich, G. and Gilmore, R. (1992) Oligosaccharyltransferase activity is associated with a protein complex composed of robophorins I and II and a 48 kd protein. *Cell* **69,** 55–65.

Klose, M., Schwarz, H., MacIntyre, S., Freudl, R., Eschbach, M. L. and Henning, U. (1988) Internal deletions in the gene for an *Escherichia coli* outer membrane protein define an area possibly important for recognition of the outer membrane by this polypeptide. *J. biol. Chem.* **263,** 13,291–13,296.

Klose, M., Storiko, A., Stierhof, Y. D., Hindennach, I., Mutschler, B. and Henning, U. (1993) Membrane assembly of the outer membrane protein OmpA of *Escherichia coli*. *J. biol. Chem.* **268,** 25,664–25,670.

Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. appl. Crystallogr.* **24,** 946–950.

Kreusch, A., Neubuser, A., Schiltz, E., Weckesser, J. and Schulz, G. E. (1994) Structure of the membrane channel porin from *Rhodopseudomonas blastica* at 2.0 Å resolution. *Protein Sci.* **3,** 58–63.

Kühlbrandt, W. (1992) Two-dimensional crystallization of membrane proteins. *Q. Rev. Biophys.* **25,** 1–49.

Kühlbrandt, W. and Wang, D. N. (1991) Three-dimensional structure of plant light-harvesting complex determined by electron crystallography. *Nature* **350,** 130–134.

Kühlbrandt, W., Wang, D. N. and Fujiyoshi, Y. (1994) Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367,** 614–621.

Kuroiwa, T., Sakaguchi, M., Mihara, K. and Omura, T. (1991) Systematic analysis of stop-transfer sequence for microsomal membrane. *J. biol. Chem.* **266,** 9251–9255.

Kutay, U., Ahnerthilger, G., Hartmann, E., Wiedenmann, B. and Rapoport, T. A. (1995) Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. *Eur. molec. Biol. Org. J.* **14,** 217–223.

Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J. molec. Biol.* **157,** 105–132.

Laidler, V., Chaddock, A. M., Knott, T. G., Walker, D. and Robinson, C. (1995) A SecY homolog in *Arabidopsis thaliana*—Sequence of a full-length cDNA clone and import of the precursor protein into chloroplasts. *J. biol. Chem.* **270,** 17,664–17,667.

Lakey, J. H., Duche, D., Gonzalezmanas, J. M., Baty, D. and Pattus, F. (1993) Fluorescence energy transfer distance measurements—The hydrophobic helical hairpin of Colicin-A in the membrane bound state. *J. molec. Biol.* **230,** 1055–1067.

Lakey, J. H., Parker, M. W., Gonzalezmanas, J. M., Duche, D., Vriend, G., Baty, D. and Pattus, F. (1994) The role of electrostatic charge in the membrane insertion of colicin a—Calculation and mutation. *Eur. J. Biochem.* **220,** 155–163.

Lanyi, J. K. (1995) Bacteriorhodopsin as a model for proton pumps. *Nature* **375,** 461–462.

Lazar, S. W. and Kolter, R. (1996) SurA assists the folding of *Escherichia coli* outer membrane proteins. *J. Bacteriol.* **178,** 1770–1773.

Lee, G. F., Burrows, G. G., Lebert, M. R., Dutton, D. P. and Hazelbauer, G. L. (1994) Deducing the organization of a transmembrane domain by disulfide crosslinking. The bacterial chemoreceptor Trg. *J. biol. Chem.* **269,** 29,920–29,927.

Lee, G. F. and Hazelbauer, G. L. (1995) Quantitative approaches to utilizing mutational analysis and disulfide crosslinking for modeling a transmembrane domain. *Protein Sci.* **4,** 1100–1107.

Lee, J.-I., Hwang, P. P., Hansen, C. and Wilson, T. H. (1992) Possible salt bridges between transmembrane α-helices of the lactose carrier of *Escherichia coli. J. biol. Chem.* **267,** 20,758–20,764.

Lee, J.-I., Hwang, P. P. and Wilson, T. H. (1993) Lysine 319 interacts with both glutamic acid 269 and aspartic acid 240 in the lactose carrier of *Escherichia coli. J. biol. Chem.* **268,** 20,007–20,015.

Lee, J. I., Kuhn, A. and Dalbey, R. E. (1992) Distinct domains of an oligotopic membrane protein are Sec-dependent and Sec-independent for membrane insertion. *J. biol. Chem.* **267,** 938–943.

Lemmon, M., Flanagan, J., Hunt, J., Adair, B., Bormann, B.-J., Dempsey, C. and Engelman, D. (1992) Glycophorin A dimerization is driven by specific interactions between transmembrane α-helices. *J. biol. Chem.* **267,** 7683–7689.

Lemmon, M. A. and Engelman, D. M. (1994) Specificity and promiscuity in membrane helix interactions. *Fedn Eur. biochem. Socs Lett.* **346,** 17–20.

Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J. and Engelman, D. M. (1992) Sequence specificity in the dimerization of transmembrane α-helices. *Biochemistry* **31,** 12,719–12,725.

Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brünger, A. T. and Engelman, D. M. (1994) A dimerization motif for transmembrane α-helices. *Nature struct. Biol.* **1,** 157–163.

Lipp, J., Flint, N., Haeuptle, M.-T. and Dobberstein, B. (1989) Structural requirements for membrane assembly of proteins spanning the membrane several times. *J. Cell Biol.* **109,** 2013–2022.

Ludlam, C. F. C., Arkin, I. T., Liu, X. M., Rothman, M. S., Rath, P., Aimoto, S., Smith, S. O., Engelman, D. M. and Rothschild, K. J. (1996) Fourier transform infrared spectroscopy and site-directed isotope labeling as a probe of local secondary structure in the transmembrane domain of phospholamban. *Biophys. J.* **70,** 1728–1736.

Luirink, J. and Dobberstein, B. (1994) Mammalian and *Escherichia coli* signal recognition particles. *Molec. Microbiol.* **11,** 9–13.

MacIntyre, S., Freudl, R., Eschbach, M.-L. and Henning, U. (1988) An artificial hydrophobic sequence functions as either an anchor or a signal sequence at only one of two positions within the *Escherichia coli* outer membrane protein OmpA. *J. biol. Chem.* **263,** 19,053–19,059.

Manoil, C. (1991) Analysis of membrane protein topology using alkaline phosphatase and β-galactosidase gene fusions. *Meth. Cell Biol.* **34,** 61–75.

Manoil, C. and Traxler, B. (1995) Membrane protein assembly: Genetic, evolutionary and medical perspectives. *A. Rev. Genet.* **29,** 131–150.

Mant, A., Skovgaard-Nielsen, V., Knott, T. G., Lindberg-Møller, B. and Robinson, C. (1994) Multiple mechanisms for the targeting of photosystem I subunits F, H, K, L, and N into and across the thylakoid membrane. *J. biol. Chem.* **269,** 27,303–27,309.

McDermott, G., Price, S. M., Freer, A. A., Hawthornthwaite-Lawless, A. M., Papiz, M. Z., Cogdell, R. J. and Isaacs, N. W. (1995) Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria. *Nature* **374,** 517–521.

Miller, J. D., Wilhelm, H., Gierasch, L., Gilmore, R. and Walter, P. (1993) GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. *Nature* **366,** 351–354.

Mingarro, I., Whitley, P., Lemmon, M. A. and von Heijne, G. (1996) Ala-insertion scanning mutagenesis of the glycophorin A transmembrane helix. A rapid way to map helix–helix interactions in integral membrane proteins. *Protein Sci.* **5,** 1339–1341.

Miroux, B. and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. molec. Biol.* **260,** 289–298.

Nakai, K. and Kanehisa, M. (1991) Expert system for predicting protein localization sites in Gram-negative bacteria. *Proteins: Struct. Funct. Genet.* **11,** 95–110.

Nakai, M., Goto, A., Nohara, T., Sugita, D. and Endo, T. (1994) Identification of the SecA protein homolog in pea chloroplasts and its possible involvement in thylakoidal protein transport. *J. biol. Chem.* **269,** 31,338–31,341.

Nakai, M., Sugita, D., Omata, T. and Endo, T. (1993) Sec-Y protein is localized in both the cytoplasmic and thylakoid membranes in the cyanobacterium *Synechococcus* PCC7942. *Biochem. Biophys. Res. Commun.* **193,** 228–234.

Nilsson, I. and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. biol. Chem.* **268,** 5798–5801.

Nilsson, I., Whitley, P. and von Heijne, G. (1994) The C-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J. Cell Biol.* **126,** 1127–1132.

Nishiyama, K., Suzuki, T. and Tokuda, H. (1996) Inversion of the membrane topology of SecG coupled with SecA-dependent preprotein translocation. *Cell* **85,** 71–81.

Oliver, D. B., Cabelli, R. J., Dolan, K. M. and Jarosik, G. P. (1990) Azide-resistant mutants of *Escherichia coli* alter the SecA-protein, an azide-sensitive component of the protein export machinery. *Proc. natn. Acad. Sci. U.S.A.* **87,** 8227–8231.

Oliver, J., Jungnickel, B., Görlich, D., Rapoport, T. and High, S. (1995) The sec61 complex is essential for the insertion of proteins into the membrane of the endoplasmic reticulum. *Fedn Eur. biochem. Socs Lett.* **362,** 126–130.

Ostermann, J., Horwich, A. L., Neupert, W. and Hartl, F. U. (1989) Protein folding in mitochondria requires complex formation with Hsp60 and ATP hydrolysis. *Nature* **341,** 125–130.

Pakula, A. A. and Simon, M. I. (1992) Determination of transmembrane protein structure by disulfide cross-linking—The *Escherichia coli* tar receptor. *Proc. natn. Acad. Sci. U.S.A.* **89,** 4144–4148.

Parker, M., Pattus, F., Tucker, A. and Tsernoglou, D. (1989) Structure of the membrane-pore-forming fragment of colicin A. *Nature* **337,** 93–96.

Parker, M. W., Buckley, J. T., Postma, J. P. M., Tucker, A. D., Leonard, K., Pattus, F. and Tsernoglou, D. (1994) Structure of the aeromonas toxin proaerolysin in its water-soluble and membrane-channel states. *Nature* **367,** 292–295.

Parker, M. W. and Pattus, F. (1993) Rendering a membrane protein soluble in water—A common packing motif in bacterial protein toxins. *Trends biochem. Sci.* **18,** 391–395.

Persson, B. and Argos, P. (1996) Topology prediction of membrane proteins. *Protein Sci.* **5,** 363–371.

Peterson, J. R., Ora, A., Van, P. N. and Helenius, A. (1995) Transient, lectin-like association of calreticulin with folding intermediates of cellular and viral glycoproteins. *Molec. Biol. Cell* **6,** 1173–1184.

Picot, D. and Garavito, R. M. (1994) Prostaglandin H synthase: Implications for membrane structure. *Fedn Eur. biochem. Socs Lett.* **346,** 21–25.

Picot, D., Loll, P. J. and Garavito, R. M. (1994) The X-ray crystal structure of the membrane protein prostaglandin-H(2) synthase-1. *Nature* **367,** 243–249.

Pind, S., Riordan, J. R. and Williams, D. B. (1994) Participation of the endoplasmic reticulum chaperone calnexin (p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator. *J. biol. Chem.* **269,** 12,784–12,788.

Popot, J.-L. and de Vitry, C. (1990) On the microassembly of integral membrane proteins. *A. Rev. Biophys. Chem.* **19,** 369–403.

Popot, J.-L. and Saraste, M. (1995) Engineering membrane proteins. *Curr. Opin. Biotech.* **6,** 394–402.

Popot, J. L. and Engelman, D. M. (1990) Membrane protein folding and oligomerization—The 2-stage model. *Biochemistry* **29,** 4031–4037.

Prinz, W. A. and Beckwith, J. (1994) Gene fusion analysis of membrane protein topology: A direct comparison of alkaline phosphatase and β-lactamase fusions. *J. Bacteriol.* **176,** 6410–6413.

Randall, L. L. (1992) Peptide binding by chaperone SecB—Implications for recognition of nonnative structure. *Science* **257,** 241–245.

Rapiejko, P. J. and Gilmore, R. (1994) Signal sequence recognition and targeting of ribosomes to the endoplasmic reticulum by the signal recognition particle do not require GTP. *Molec. Biol. Cell* **5,** 887–897.

Rapoport, T. A. (1992) Transport of proteins across the endoplasmic reticulum membrane. *Science* **258,** 931–936.

Rees, D. C., DeAntonio, L. and Eisenberg, D. (1989) Hydrophobic organization of membrane proteins. *Science* **245,** 510–513.

Rees, D. C., Komiya, H., Yeates, T. O., Allen, J. P. and Feher, G. (1989) The bacterial photosynthetic reaction center as a model for membrane proteins. *A. Rev. Biochem.* **58,** 607–633.

Reithmeier, R. A. F. (1995) Characterization and modelling of membrane proteins using sequence analysis. *Curr. Opin. struct. Biol.* **5,** 491–500.

Rojo, E. E., Stuart, R. A. and Neupert, W. (1995) Conservative sorting of F0-ATPase subunit 9: Export from matrix requires delta pH across inner membrane and matrix ATP. *Eur. molec. Biol. Org. J.* **14,** 3445–3451.

Rospert, S., Müller, S., Schatz, G. and Glick, B. S. (1994) Fusion proteins containing the cytochrome b2 presequence are sorted to the mitochondrial intermembrane space independently of hsp60. *J. biol. Chem.* **269,** 17,279–17,288.

Rost, B., Casadio, R., Fariselli, P. and Sander, C. (1995) Transmembrane helices predicted at 95% accuracy. *Protein Sci.* **4,** 521–533.

Rost, B., Fariselli, P. and Casadio, R. (1996) Topology prediction for helical transmembrane segments at 86% accuracy. *Protein Sci.* **5,** 1704–1718.

Rothman, J. E. (1996) The protein machinery of vesicle budding and fusion. *Protein Sci.* **5,** 185–194.

Sahin-Toth, M., Dunten, R. L. and Kaback, H. R. (1995) Design of a membrane protein for site-specific proteolysis: Properties of engineered factor Xa protease sites in the lactose permease of *Escherichia coli*. *Biochemistry* **34,** 1107–1112.

Samatey, F. A., Xu, C. and Popot, J.-C. (1995) On the distribution of amino acid residues in transmembrane alpha-helix bundles. *Proc. natn. Acad. Sci. U.S.A.* **92,** 4577–4581.

Scaramuzzi, C. D., Hiller, R. G. and Stokes, H. W. (1992) Identification of a chloroplast-encoded secA gene homologue in a chromophytic alga—possible role in chloroplast protein translocation. *Curr. Genet.* **22,** 421–427.

Scaramuzzi, C. D., Stokes, H. W. and Hiller, R. G. (1992) Characterisation of a chloroplast-encoded sec-Y homologue and atpH from a chromophytic alga—Evidence for a novel chloroplast genome organisation. *Fedn Eur. biochem. Socs Lett.* **304,** 119–123.

Schatz, G. and Dobberstein, B. (1996) Common principles of protein translocation across membranes. *Science* **271,** 1519–1526.

Schertler, G. F. X., Villa, C. and Henderson, R. (1993) Projection structure of rhodopsin. *Nature* **362,** 770–772.

Schirmer, T. and Cowan, S. W. (1993) Prediction of membrane-spanning β-strands and its application to maltoporin. *Protein Sci.* **2,** 1361–1363.

Schirmer, T., Keller, T. A., Wang, Y. F. and Rosenbusch, J. P. (1995) Structural basis for sugar translocation through maltoporin channels at 3.1 angstrom resolution. *Science* **267,** 512–514.

Schwartz, T. W., Gether, U., Schambye, H. T. and Hjort, S. A. (1995) Molecular mechanism of action of non-peptide ligands for peptide receptors. *Curr. pharmaceut. Design* **1,** 325–342.

Silberstein, S., Kelleher, D. and Gilmore, R. (1992) The 48-kDa subunit of the mammalian oligosaccharyl-transferase complex is homologous to the essential yeast protein WBP1. *J. biol. Chem.* **267,** 23,658–23,663.

Simmerman, H. K. B., Kobayashi, Y. M., Autry, J. M. and Jones, L. R. (1996) A leucine zipper stabilizes the pentameric membrane domain of phospholamban and forms a coiled-coil pore structure. *J. biol. Chem.* **271,** 5941–5946.

Spiess, M. (1995) Heads or tails—What determines the orientation of proteins in the membrane. *Fedn Eur. biochem. Socs Lett.* **369,** 76–79.

Surrey, T. and Jähnig, F. (1992) Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proc. natn. Acad. Sci. U.S.A.* **89,** 7457–7461.

Surrey, T., Schmid, A. and Jahnig, F. (1996) Folding and membrane insertion of the trimeric beta-barrel protein OmpF. *Biochemistry* **35,** 2283–2288.

Suwa, M., Hirokawa, T. and Mitaku, S. (1995) A continuum theory for the prediction of lateral and rotational positioning of alpha-helices in membrane proteins: Bacteriorhodopsin. *Proteins: Struct. Funct. Genet.* **22,** 363–377.

Sääf, A., Andersson, H., Gafvelin, G. and von Heijne, G. (1995) SecA-dependence of the translocation of a large periplasmic loop in the *E. coli* MalF inner membrane protein is a function of sequence context. *Molec. Membrane Biol.* **12,** 209–215.

Sääf, A., Wallin, E. and von Heijne, G. (1997) Translocation of marginally hydrophobic segments across prokarytoic and eukaryotic membranes, (*submitted*).

Taylor, W. R., Jones, D. T. and Green, N. M. (1994) A method for α-helical integral membrane protein fold prediction. *Proteins: Struct. Funct. Genet.* **18,** 281–294.

Traxler, B. and Murphy, C. (1996) Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery. *J. biol. Chem.* **271,** 12,394–12,400.

Treutlein, H. R., Lemmon, M. A., Engelman, D. M. and Brünger, A. T. (1992) The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* **31,** 12,726–12,732.

Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. and Yoshikawa, S. (1995) Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 Å. *Science* **269,** 1064–1069.

Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. and Yoshikawa, S. (1996) The whole stucture of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. *Science* **272,** 1136–1144.

Tucker, J. and Grisshammer, R. (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem. J.* **317,** 891–899.

Tuffery, P., Etchebest, C., Popot, J. L. and Lavery, R. (1994) Prediction of the positioning of the seven transmembrane α-helices of bacteriorhodopsin—A molecular simulation study. *J. molec. Biol.* **236,** 1105–1122.

Uchida, K., Mori, H. and Mizushima, S. (1995) Stepwise movement of preproteins in the process of translocation across the cytoplasmic membrane of *Escherichia coli*. *J. biol. Chem.* **270,** 30,862–30,868.

van der Goot, F. G., Gonzalez-Manas, J. M., Lakey, J. H. and Pattus, F. (1991) A molten-globule membrane-insertion intermediate of the pore-forming domain of colicin-A. *Nature* **354,** 408–410.

van Gelder, P., de Cock, H. and Tommassen, J. (1994) Detergent-induced folding of the outer-membrane protein PhoE, a pore protein induced by phosphate limitation. *Eur. J. Biochem.* **226,** 783–787.

van Wijk, K. J., Knott, T. G. and Robinson, C. (1995) Evidence for SecA- and delta pH-independent insertion of D1 into thylakoids. *Fedn Eur. biochem. Socs Lett.* **368,** 263–266.

von Heijne, G. (1986) The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *Eur. molec. Biol. Org. J.* **5,** 3021–3027.

von Heijne, G. (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* **341,** 456–458.

von Heijne, G. (1992) Membrane protein structure prediction—Hydrophobicity analysis and the positive-inside rule. *J. molec. Biol.* **225,** 487–494.

von Heijne, G. (1994) Membrane proteins: From sequence to structure. *A. Rev. Biophys. biomolec. Struct.* **23,** 167–192.

von Heijne, G. (1994) *Sec*-independent protein insertion into the inner *E. coli* membrane: A phenomenon in search of an explanation. *Fedn Eur. biochem. Socs. Lett.* **346,** 69–72.

von Heijne, G. (1994) *Signal Peptidases*, R.G. Landes, Austin, TX.

von Heijne, G. (1997) *Membrane Protein Assembly*, R.G. Landes, Austin, TX.

von Heijne, G. and Gavel, Y. (1988) Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* **174,** 671–678.

von Heijne, G. and Manoil, C. (1990) Membrane proteins—From sequence to structure. *Protein Eng.* **4,** 109–112.

Voss, J., Hubbell, W. L. and Kaback, H. R. (1995) Distance determination in proteins using designed metal ion binding sites and site-directed spin labeling: Application to the lactose permease of *Escherichia coli*. *Proc. natn. Acad. Sci. U.S.A.* **92,** 12,300–12,303.

Wallin, E. and von Heijne, G. (1995) Properties of N-terminal tails in G-protein coupled receptors—A statistical study. *Protein Eng.* **8,** 693–698.

Weiss, M. S., Abele, U., Weckesser, J., Welte, W., Schiltz, E. and Schulz, G. E. (1991) Molecular architecture and electrostatic properties of a bacterial porin. *Science* **254,** 1627–1630.

Weiss, M. S., Kreusch, A., Schiltz, E., Nestel, U., Welte, W., Weckesser, J. and Schulz, G. E. (1991) The structure of porin from *Rhodobacter capsulata* at 1.8 Å resolution. *Fedn Eur. biochem. Socs Lett.* **280,** 379–382.

Welte, W., Weiss, M. S., Nestel, U., Weckesser, J., Schiltz, E. and Schulz, G. E. (1991) Prediction of the general structure of OmpF and PhoE from the sequence and structure of porin from Rhodobacter-Capsulatus—Orientation of porin in the membrane. *Biochim. biophys. Acta* **1080,** 271–274.

Wen, J. A., Chen, X. and Bowie, J. U. (1996) Exploring the allowed sequence space of a membrane protein. *Nature struct. Biol.* **3,** 141–148.

Wessels, H. P. and Spiess, M. (1988) Insertion of a multispanning membrane protein occurs sequentially and requires only one signal sequence. *Cell* **55,** 61–70.

Whitley, P., Grahn, E., Kutay, U., Rapoport, T. and von Heijne, G. (1996) A 12 residues long poly-leucine tail is sufficient to anchor syntaptobrevin to the ER membrane. *J. biol. Chem.* **271,** 7583–7586.

Whitley, P., Nilsson, I. and von Heijne, G. (1994) *De novo* design of integral membrane proteins. *Nature struct. Biol.* **1,** 858–862.

Whitley, P., Nilsson, L. and von Heijne, G. (1993) 3-Dimensional model for the membrane domain of *Escherichia coli* leader peptidase based on disulfide mapping. *Biochemistry* **32,** 8534–8539.

Whitley, P., Zander, T., Ehrmann, M., Haardt, M., Bremer, E. and von Heijne, G. (1994) Sec-independent translocation of a 100-residues long periplasmic N-terminal tail in the *E. coli* inner membrane proteins ProW. *Eur. molec. Biol. Org. J.* **13,** 4653–4661.

Wolfe, P. B., Rice, M. and Wickner, W. (1985) Effects of two sec genes on protein assembly into the plasma membrane of *Escherichia coli*. *J. biol. Chem.* **260,** 1836–1841.

Wu, J. H., Perrin, D. M., Sigman, D. S. and Kaback, H. R. (1995) Helix packing of lactose permease in *Escherichia coli* studied by site-directed chemical cleavage. *Proc. natn. Acad. Sci. U.S.A.* **92,** 9186–9190.

Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P. and Feher, G. (1987) Structure of the reaction center from *Rhodobacter sphaeroides*: Membrane–protein interactions. *Proc. natn. Acad. Sci. U.S.A.* **84,** 6438–6442.

Zhang, J. T., Lee, C. H., Duthie, M. and Ling, V. (1995) Topological determinants of internal transmembrane segments in P-glycoprotein sequences. *J. biol. Chem.* **270,** 1742–1746.

Zhang, J. T. and Ling, V. (1991) Study of membrane orientation and glycosylated extracellular loops of mouse P-glycoprotein by *in vitro* translation. *J. biol. Chem.* **266,** 18224–18232.

NOTE ADDED IN PROOF

The structure of the toxin hemolysin has recently been solved [Song, L., Hobaugh, M. R., Shustak, C., Ceheley, S., Bayley, H. and Gouaux, J. E. (1996) Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science* **274,** 1859–1866], providing another example of a β-barrel membrane protein. The involvements of the *E. coli* SRP and SRP receptor homologues Ffh and FtsY in the assembly of certain inner membrane proteins has now been demonstrated [de Gier, J.-W. L., Mansournia, P., Valent, Q., Phillips, G. J., Luirink, J. and von Heijne, G. (1996) Assembly of a cytoplasmic membrane protein in *Escherichia coli* is dependent on the signal recognition particle. *Fedn Eur. biochem. Socs Lett.* **399,** 307–309; Seluanov, A. and Bibi, E. (1997) FtsY, the prokaryotic signal recognition particle, is essential for biogenesis of membrane proteins. *J. biol. Chem.* **272,** 2053–2055; Ulbrandt, N. D., Newitt, J. A. and Bernstein, H. D. (1997) The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell 88,* 187–196].