

University of Groningen

Membrane topology and insertion of membrane proteins

Geest, Marleen van; Lolkema, Julius

Published in:
Microbiology and Molecular Biology Reviews

DOI:
[10.1128/MMBR.64.1.13-33.2000](https://doi.org/10.1128/MMBR.64.1.13-33.2000)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2000

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Geest, M. V., & Lolkema, J. S. (2000). Membrane topology and insertion of membrane proteins: Search for topogenic signals. *Microbiology and Molecular Biology Reviews*, 64(1), 13-+. DOI: 10.1128/MMBR.64.1.13-33.2000

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Membrane Topology and Insertion of Membrane Proteins: Search for Topogenic Signals

MARLEEN VAN GEEST AND JUKE S. LOKKEMA*

*Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute,
University of Groningen, 9751 NN Haren, The Netherlands*

INTRODUCTION	14
MEMBRANE TOPOLOGY	14
Principles	14
Enzyme Tags	15
The enzymes: alkaline phosphatase, β -galactosidase, and β -lactamase	15
Complementarity	15
Expression levels and specific activity	16
Sandwich fusions	16
Assumptions and rules	16
Glycosylation Tags	17
Glycosylation of membrane proteins	17
Distance between OST and the membrane	17
Glycosylation scanning	18
Glycosylation fusions	18
Chemical Modification	19
Cysteine residues in membrane proteins	19
Cysteine-scanning mutagenesis	19
Additional advantages	19
Conclusions	19
The BAD Tag	20
BAD	20
BAD in topology studies	21
Folding kinetics	22
Proteolysis	22
MEMBRANE INSERTION	22
Topology and Insertion	22
Membrane Protein Biogenesis	23
Insertion machinery in the ER membrane	23
Insertion into the bacterial cytoplasmic membrane	23
(i) Bacterial translocation machinery	23
(ii) Sec-dependent and Sec-independent insertion of membrane proteins	23
Insertion mechanism	24
Insertion of Isolated Hydrophobic Segments	25
Principles	25
Insertion vehicles	25
(i) M0 and M1 system	25
(ii) Lep system	26
(iii) Prolactin system	26
SA and ST activities of isolated transmembrane segments	26
Similar studies in <i>E. coli</i>	27
Interaction between Multiple Insertion Signals	28
Charge distribution and hydrophobic segments	28
Interaction between different hydrophobic segments	28
(i) Insertion depending on upstream segments	28
(ii) Insertion depending on downstream segments	28
(iii) Folding intermediates	29
PROSPECTS	29
REFERENCES	30

* Corresponding author. Mailing address: Department of Microbiology, University of Groningen Biological Center, Kerklaan 30, 9751 NN Haren, The Netherlands. Phone: 3150-3532155. Fax: 3150-3632154. E-mail: j.s.lokkema@biol.rug.nl.

INTRODUCTION

Integral membrane proteins represent an important class of proteins that are involved in a wide variety of cellular functions. Knowledge of the structure of proteins is crucial to understanding their function. Unfortunately, there are no general and reliable methods for forming three-dimensional crystals of membrane proteins suitable for crystallographic analysis, and to date, only a handful of high-resolution membrane protein structures have been solved whereas several thousands of three-dimensional structures of globular proteins are known. Because of this, biochemical and prediction methods were needed to obtain structural information about membrane proteins.

Although integral membrane proteins come in a variety of sizes and shapes, they have common basic architectural principles, most probably due to the lipid environment in which they are embedded. The membrane-spanning portions of the so-called α -helix bundle proteins, which are the subjects of this review, contain one or more transmembrane α -helices, each of which is a stretch of approximately 20 amino acids with largely hydrophobic side chains. The α -helices are oriented more or less perpendicular to the plane of the membrane. In bitopic membrane proteins, a single helix connects two domains of the protein on either side of the membrane. In polytopic membrane proteins, the membrane-spanning portion of the protein consists of multiple α -helices that are connected by extramembranous domains, i.e., the loops. Three-dimensional structures show that the helices of polytopic membrane proteins are packed intimately in the membrane. Analysis of the hydrophobic moments of transmembrane α -helices in polytopic membrane proteins of known structures indicates that the most polar face of each helix is buried in the interior of the molecule while the least polar face is exposed to the lipids (148, 149).

A fundamental aspect of the structure of polytopic membrane proteins is the membrane topology, i.e. the number of transmembrane segments and their orientation in the membrane. Fortunately, despite the difficulties encountered in obtaining high-resolution structures, the physicochemical constraints imposed by the lipid environment provide a simple method to predict the topology of a membrane protein. The predicted topology can be verified by a variety of molecular and biochemical techniques. Membrane protein topology predictions are based on the observations that (i) the transmembrane α -helices have a high overall hydrophobicity and (ii) the charge distribution of the hydrophilic loops that connect the transmembrane segments follows the "positive inside" rule, which states that nontranslocated loops are enriched in positively charged residues compared to translocated loops (191). The first observation is used to identify the transmembrane segments in the amino acid sequence by analyzing the hydrophobic properties of the amino acid sequence (39, 103, 175, 191), and the second observation is used to predict the overall orientation of the protein in the membrane. The biochemical techniques used to verify the predicted membrane topology are, without exception, based on modifications of the proteins by engineering the structural genes coding for the proteins. These techniques are reviewed and evaluated in the first part of this review.

The success of biochemical approaches to determining membrane protein topologies will increase dramatically with the understanding of the dynamics of the biosynthetic pathway leading to the folded protein in the membrane. Thus, as well as knowledge of membrane protein synthesis and targeting to the appropriate membrane, understanding of the process of insertion into the membrane and formation of the final three-di-

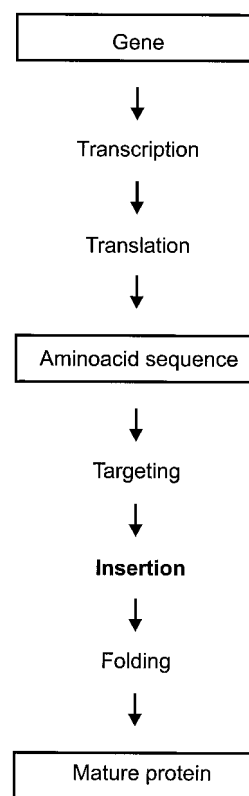


FIG. 1. Membrane protein biogenesis. In many membrane topology studies, the membrane protein is modified by engineering of the structural gene, while the results are analyzed at the level of the folded protein. The structure of a membrane protein can be deduced from its amino acid sequence only if the different steps leading to the structure are understood.

mensional structure is necessary to determine and understand the topology of membrane proteins (Fig. 1). Many important aspects of membrane protein biosynthesis seem to rely on rather well-defined signals encoded in the polypeptide chain, such as targeting signals and topogenic signals. The membrane topology is formed in a process in which the topogenic signals in the nascent polypeptide chain are recognized and translated by the insertion machinery. Topology studies and prediction methods will become much more reliable when all the topogenic signals present in the amino acid sequence of a membrane protein are recognized and understood and when it is known how the insertion machinery deals with them. Although several targeting and topological signals have been investigated extensively, a lot of topogenic information is still hidden in the polypeptide chains of membrane proteins as unknown signals. The second part of this review focuses on studies that specifically search for topogenic signals by using similar techniques to those used in the studies of membrane topology.

MEMBRANE TOPOLOGY

Principles

Plasma membrane proteins cross the membrane in a zigzag fashion and expose their hydrophilic loops alternately in the two compartments that are separated by the membrane. To verify predicted membrane protein topology models, the existence of all the putative transmembrane domains must be verified and the hydrophilic loops must be localized to one of the compartments. The simple structure of membrane pro-

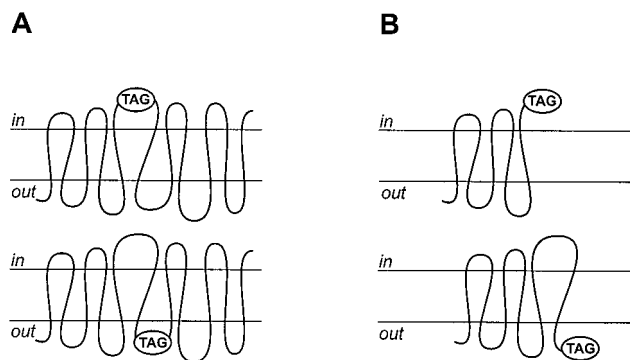


FIG. 2. Biochemical approaches to determination of membrane protein topology. (A) Insertions. (B) Fusions.

teins, the nature of the membrane, and the possibility of genetically modifying proteins allowed the development of a variety of biochemical and molecular techniques for the elucidation of the membrane topology of integral membrane proteins. In vivo, the membrane represents a barrier that separates different intracellular compartments from each other or separates the interior of the cell from the external environment. Due to the impermeability of the membrane to hydrophilic molecules, parts of a membrane protein that lie on opposite sides of the membrane are differently accessible to various agents. Easily identified target sites are inserted in the polypeptide, and membrane-impenetrable reagents are used to determine their accessibility at one side of the membrane (Fig. 2A). Examples of target sites include N-glycosylation sites, Cys residues, iodinated sites, antibody epitopes, and proteolytic sites that are introduced at specific positions in the protein by site directed mutagenesis. By inserting the tag at different positions in the protein, the complete topology can be determined.

A second approach to generate topological information makes use of gene fusions. A reporter molecule is attached to a hydrophilic domain of a membrane protein, and the cellular location of the reporter is determined by the topogenic information in the membrane protein (Fig. 2B). The reporters are typically molecules whose properties (for example, enzymatic activity) depend on their subcellular location. The gene encoding the reporter protein is fused at different sites in the gene encoding the membrane protein, and the properties of the resulting fusion proteins reveal at which side of the membrane the fusion sites reside. In the resulting fusion proteins, the C-terminal part of the membrane protein is deleted and replaced by the reporter moiety. Therefore, the fusions are termed C-terminal deletion fusions.

Enzyme Tags

The enzymes: alkaline phosphatase, β -galactosidase, and β -lactamase. The most commonly used reporter protein for gene fusion studies in prokaryotes is alkaline phosphatase, encoded by the *Escherichia coli* *phoA* gene (24, 118, 175). Alkaline phosphatase (PhoA) is a metalloenzyme that consists of two identical subunits and resides in the bacterial periplasm (30). Each of the subunits contains two intramolecular disulfide bridges. PhoA is initially synthesized as a precursor with an N-terminal signal sequence that is removed during translocation across the membrane. In the periplasm, the mature part of PhoA is oxidized. The cysteine residues form disulfide bridges, triggering the correct folding of PhoA, a partially protease

resistant conformation. The folding process is completed by dimer formation, yielding the active enzyme complex (5). The process of folding and assembly of PhoA occurs only after export to the periplasmic space, because various factors prevent formation of the disulfide bonds in the cytoplasm (48). This location-specific character of PhoA allowed its wide use as a reporter molecule in membrane protein topology studies in prokaryotic cells (47) and is applicable to heterologously expressed eukaryotic membrane proteins as well (21, 67, 108). In these studies, the mature part of PhoA, lacking its signal sequence, is fused behind various C-terminal truncated parts of a membrane protein. The truncated protein takes over the function of the signal sequence. If PhoA is fused to a domain of a membrane protein that normally is located in the cytoplasm, the PhoA molecule also remains in the cytoplasm and is inactive. In contrast, when PhoA is fused to a domain that normally is exported across the membrane, the PhoA moiety of the fusion protein is also exported across the membrane and is folded and assembled into the active state. Thus, the enzymatic activity of the fusion protein reveals the cellular location of the fusion site. The alkaline phosphatase activity of the fusion proteins is most easily determined by a plate assay using a chromogenic substrate such as XP (5-bromo-4-chloro-3-indolyl-phosphate) or, more quantitatively, by spectroscopic enzyme assays (124). Commercially available antibodies directed against PhoA enable the detection and quantitation of the fusion proteins by immunoblotting.

Fusions of the *E. coli* enzyme β -galactosidase (LacZ), a large tetrameric cytoplasmic enzyme, are used in a complementary fashion to the PhoA fusions. In contrast to PhoA, LacZ exhibits enzymatic activity in the cytoplasm. When attached downstream of an export signal, the enzyme is trapped in the membrane, which prevents proper folding; therefore, the enzyme is inactive (164). Thus, fusions of LacZ to a periplasmic fusion site of a membrane protein are inactive, while fusions to a cytoplasmic domain are active. Similar to PhoA fusions, β -galactosidase is visualized by a plate assay using the chromogenic substrate X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside) or is measured quantitatively using the spectroscopic enzyme assay (63). Antibodies directed against LacZ to quantify the expression level are commercially available as well.

The third commonly used reporter molecule, which has been used as an alternative to PhoA, is the mature form of the monomeric periplasmic enzyme β -lactamase, encoded by the *bla* gene (28). Periplasmic β -lactamase protects cells against lysis by β -lactam antibiotics such as ampicillin, which otherwise inactivate enzymes essential for cell wall biosynthesis that are anchored to the outer surface of the cytoplasmic membrane. Cells expressing cytoplasmic β -lactamase are sensitive to the antibiotics because cytoplasmic β -lactamase cannot intervene between the incoming antibiotic and their periplasmic targets. Thus, cells expressing fusion proteins in which the mature form of β -lactamase is fused to a cytoplasmic domain of a membrane protein fail to grow on plates with a high ampicillin concentration while cells expressing fusions to periplasmic domains are able to grow on these plates. Plating the cells at a high density identifies in-frame cytoplasmic fusions. Under these conditions, lysis of a fraction of the cell population liberates enough of the cytoplasmic β -lactamase to permit growth of the cells. Plates with different ampicillin concentrations are used to determine the localization of the β -lactamase moiety (27).

Complementarity. Various topology models have been tested using a combination of PhoA and LacZ fusions (45, 57, 76, 91, 92, 106, 170, 202, 206). Random fusions to both reporter molecules are often constructed, and this is followed by screen-

ing for positive clones on indicator plates. Positive LacZ fusions represent fusions to cytoplasmic domains, while positive PhoA fusions represent fusions to periplasmic domains. Genetic tricks have been developed which make it possible to exchange the LacZ reporter molecules in positive fusions for the PhoA molecule and vice versa (117, 200), resulting in fusions with complementary enzymatic properties (76, 170). The use of PhoA as a reporter molecule in these studies is advantageous in principle because a positive result (i.e., a high-activity fusion) requires the active export of the mature reporter enzyme moiety into the periplasm. In contrast, the use of LacZ as a complementary cytoplasmic reporter may give a positive phenotype as a result of many artifacts. Saturation of the export machinery, especially when using overexpression systems, and the liberation of an initially membrane-embedded, inactive portion of β -galactosidase may explain the unexpectedly high β -galactosidase activity. Conflicting results, i.e., high LacZ and PhoA activity at the same fusion site, have been reported in a number of membrane topology studies (45, 63, 68, 70, 77, 184). Furthermore, in a few cases, the truncated membrane proteins adopt alternate conformations depending on which reporter protein is present. For unknown reasons, the correct folding in the membrane is more often achieved with PhoA as the reporter (57, 106). In recent years, most topology studies have been done with just PhoA or β -lactamase.

Expression levels and specific activity. It is important to demonstrate that the low activity of an alkaline phosphatase fusion protein is due to a cytoplasmic location of the PhoA moiety and not to a lower level of expression of the fusion protein. Thus, the level of expression is usually determined by Western blot analysis using antibodies raised against PhoA to quantitatively relate the enzymatic activity to the protein concentration. If all full-length fusion proteins in a set show more or less the same level of expression, irrespective of their localization and activity, the activities of the fusion proteins can easily be corrected for the level of expression or may be taken to indicate topology without correction (152, 169). On the other hand, if different amounts of fusion protein are detected within a set, this may indicate differential stability or differential synthesis, which can be resolved by pulse-chase immunoblot experiments (156). Instability of the membrane protein moiety results in the liberation of alkaline phosphatase from the fusion protein. The fate of the alkaline phosphatase then depends on the side of the membrane at which it is released. Release into the periplasm results in stable and active PhoA. Immunoblots show mixtures of full-length fusion proteins and free PhoA in ratios that depend on the stability of the membrane-embedded part of the proteins. In contrast, due to improper folding of the PhoA moiety in the cytoplasm, in cytoplasmic fusion proteins the free PhoA moiety is rapidly degraded and is not or poorly detected by Western blotting. As a consequence, fusions with a periplasmic PhoA moiety show larger amounts of immunoreactive products on a Western blot than do cytoplasmic fusions (45, 106, 143, 184). This systematic behavior can be used as an additional method to discriminate between fusions with a periplasmic or cytoplasmic PhoA moiety, but it is important that it is not used to normalize the activities. To demonstrate the synthesis of the fusion proteins, pulse and/or pulse-chase experiments are generally used (26, 45, 151, 152, 168, 170, 206, 208).

Sandwich fusions. In the PhoA sandwich approach (55), the alkaline phosphatase sequence, lacking the signal sequence, is inserted into the membrane protein rather than replacing different C-terminal membrane protein sequences. Thus, the entire membrane protein is present and the approach does not suffer from the drawbacks pertinent to the C-terminal deletion

fusion approach. This approach may give a more accurate picture of the topology if the final topology of a protein is affected by interactions between amino- and carboxyl-terminal sequences. Similar to the fusion approach, high alkaline phosphatase activity of a sandwich construct reveals a periplasmic location of the domain bearing the insertion while low activity corresponds to a cytoplasmic location.

The sandwich approach was used solely or in addition to PhoA fusions to determine the topology of a number of bacterial proteins including the maltose transporter subunit MalF (55), the aromatic amino acid permease AroP (42), the phenylalanine-specific permease PheP (137), the tryptophan transporter Mtr (157), and the eukaryotic multidrug resistance protein MDR (21). To minimize the disruption of known topological signals, such as positively charged residues in cytoplasmic loops, the fusions were positioned in the C-terminal portion of each hydrophilic loop (see also the next section). While in principle sandwich fusions may be the better approach, in practice many insertion constructs are inactive (like C-terminal deletion fusions). Only 3 of 25 AroP-PhoA sandwich fusions (42) and 5 of 25 PheP-PhoA sandwich-fusions (137) retained some transport activity. As was also observed with the Mtr permease (157), the active constructs involved insertions into periplasmic domains. Since it is possible that insertion of the large PhoA domain alters the topology of the protein, inactive sandwich proteins are risky for use in topology studies and do not provide an advantage relative to C-terminal deletion fusions.

Assumptions and rules. Experimental support for a topological model requires a minimum of one fusion in each extra membranous domain. If the protein contains stretches of residues of intermediate hydrophobicity that cannot unambiguously be identified as membrane spanning, fusions should be made approximately every 30 residues (26).

The major assumption in the fusion approach to membrane protein topology is that truncation of a membrane protein does not affect its native topology. The fusion technique, especially using the reporter molecule alkaline phosphatase, has been used in a large number of topology studies (for a list of proteins until 1994, see reference 24), and several of the models that were obtained have been confirmed by other biochemical data (29, 143, 180) or even by crystallographic data (36, 85, 86, 177, 208), indicating that the assumption is valid for at least a number of proteins. On the other hand, different types of anomalous behavior of fusion proteins have been reported, which has led to the formulation of rules to optimize the analysis and to avoid pitfalls (24, 108, 175). One type of anomalous behavior is the unexpectedly high activity that is observed when alkaline phosphatase is fused near the N-terminal end of a cytoplasmic loop. An extensive study of MalF in which PhoA was fused at different sites in a cytoplasmic loop suggested that the high activity of fusions near the N-terminal end may be due to loss of C-terminal hydrophilic sequences containing positively charged residues that would lock the loop in the cytoplasm (25). It was proposed that fusions should be constructed to the C-terminal end of cytoplasmic loops or at least downstream of basic amino acid residues (26). A comparative study using β -lactamase as the reporter domain demonstrated that the equivalent MalF- β -lactamase fusions did not confer anomalous high resistance to ampicillin, indicating that the β -lactamase domain was correctly retained in the cytoplasm. The difference in behavior of PhoA and β -lactamase as the reporter domain may be the result of differences in the ability of the reporter proteins to fold into export-incompetent states in the cytoplasm; β -lactamase would fold more rapidly than PhoA (144). Similar to the β -lactamase fusions, the equivalent fu-

sions of MalF and the biotin acceptor domain of the *Propionibacterium shermanii* transcarboxylase (see "The BAD tag" below) as the reporter did not result in anomalous behavior (87). The data obtained with truncated MalF fusions suggested that a truncated membrane protein may adopt different topologies depending on which reporter protein is present. However, the membrane protein also plays a role. β -Lactamase fusions in the first cytoplasmic domain of the rhamnose/H⁺ symporter of *Salmonella enterica* serovar Typhimurium (174) conferred an anomalous high resistance to ampicillin, indicating that with this protein β -lactamase is not retained in the cytoplasm. Also, PhoA fusions in two different positively charged cytoplasmic loops of the melibiose permease MelB resulted, independent of the positions of the fusion sites relative to the positive charges, in low-activity fusions, demonstrating that the positive charges were not essential for locking the reporter in the cytoplasm (143). Apparently, the behavior of fusion proteins is dependent on the combination of membrane protein and reporter protein.

A second type of mislocation of the reporter molecule was observed when PhoA was fused in periplasmic loops following a transmembrane segment containing one or more charged residues. The low activity of the fusions indicated that the membrane segment is retained in the cytoplasm (6, 29). The export ability of the segment may be reduced because of an overall lower hydrophobicity or because C-terminal sequences of the membrane protein are required to stabilize the segment in the membrane through the formation of interhelical salt bridges. Replacing the PhoA moiety in these fusions with β -lactamase resulted in the same anomalous behavior (144).

An artefact occurs with membrane proteins whose N termini face the periplasm, such as the *E. coli* inner membrane protein ProW (76), which is a component of the ProU osmoregulatory system (71). Fusions of the reporter molecule to the N-terminal domain lack transmembrane sequences and are bound to remain in the cytoplasm (76, 109, 144, 198).

An extensive PhoA fusion analysis study with the lactose permease, LacY, and the melibiose permease, MelB, of *E. coli* led to remarkable new insights. Fusions constructed at intervals of 3 to 5 residues in each predicted transmembrane segment showed a sharp change from high to low PhoA activity. The hypothesis was proposed that the cytoplasmic half of an outgoing segment is enough to promote translocation of the alkaline phosphatase moiety to the periplasmic space while the periplasmic half of an ingoing segment is enough to prevent translocation (29, 143, 180). This rule was used to determine the approximate middle of transmembrane segments VII and XI in LacY by constructing PhoA fusions at alternating residues. The fusion point of the chimera that showed a sharp increase in alkaline phosphatase activity was set at the middle of the transmembrane segment (180). A similar sharp transition in low and high alkaline phosphatase activity for fusions that were 6 to 10 residues apart was observed with the glutamate transporter of *Bacillus stearothermophilus*, GluT (169). Clearly, these observations and explanations are in marked contrast to those discussed above.

Proper assembly of the truncated membrane protein part of a fusion protein in the membrane seems to depend on the membrane protein, the reporter protein, and the location of the fusion site. Therefore, the fusion approach is successful for certain membrane proteins but not for others. The folding of the fusion proteins depends on features of the nascent chain such as charges and hydrophobicity, but it may also depend on the rate of folding of extramembraneous domains and, clearly, on other still unknown features. It should be noted that the fusions that behave anomalously are only a minor fraction of

all the fusions. They may be especially helpful in understanding additional features that give a particular stretch of residues a transmembrane disposition (see "Membrane insertion" below). When we understand all these features, we may be able to deduce the complete membrane topology from the results of the reporter fusion technique.

Glycosylation Tags

Glycosylation of membrane proteins. In eukaryotic cells, glycosylation activity is found in the lumen of the endoplasmic reticulum (ER) and is accomplished by the enzyme oligosaccharyl transferase (OST), which adds oligosaccharides to the amino group of asparagine (Asn) residues of the consensus sequence Asn-X-Thr/Ser (196). N glycosylation is a common feature of eukaryotic membrane proteins, and the consensus sequence is usually found in the largest luminal exposed loop of the protein (104). The presence of endogenous sites in multiple loops is possible as well (171). Since modification of the glycosylation site occurs in a compartment-specific manner, the presence of glycosylation provides topological information. The loop(s) carrying the oligosaccharide is identified by successive elimination of potential glycosylation sites by site-directed mutagenesis (83). The state of glycosylation of native or mutant proteins is assayed in an *in vivo* expression system or *in vitro* in cell-free systems consisting of reticulocyte lysate or wheat germ lysate supplemented with microsomes, membranes derived from the ER. Addition of the oligosaccharide chain to a single site results in an increase in the apparent molecular mass of approximately 2.5 kDa on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Glycosylation can be manipulated by treatment with glycosidases (for example, Endo H [134, 187] and N-glycosidase F [132, 189]) or by expression of the protein in the presence of glycosylation inhibitors (for example, tunicamycin [18]) or competitive acceptor peptides (for example, Ac-Asn-Tyr-Thr-NH₂ [140] and N-benzoyl-Asn-Leu-Thr-N-methylamide [130, 187]).

Glycosylation scanning mutagenesis and glycosylation fusions make use of the glycosylation properties in the ER lumen to study the complete topology of membrane proteins in the ER membrane. They have become the counterpart of the PhoA fusion and sandwich techniques to analyze eukaryotic membrane proteins. In the glycosylation scanning mutagenesis approach, consensus glycosylation sites or domains bearing a glycosylation site are introduced into membrane proteins devoid of glycosylation sites. Localization of the insertion site on the luminal or cytoplasmic side of the membrane is inferred from the presence or absence of glycosylation of the engineered protein, respectively. The technique has proven very useful in defining the folding pattern of various membrane proteins (18, 32, 34, 84, 131, 132, 140, 167, 178). In the glycosylation fusion approach, a domain bearing one or more glycosylation sites is fused behind different C-terminal deletion mutants of a membrane protein. The fused glycosylation site functions as a reporter molecule in which glycosylation of the domain indicates a luminal position of the fusion point while no glycosylation indicates a cytoplasmic location (13, 132, 187).

Distance between OST and the membrane. The OST enzyme, which is responsible for the glycosylation of eukaryotic membrane proteins, is embedded in the ER membrane and has its catalytic site facing the luminal compartment (196). Successful N glycosylation by the active site of OST requires adequate spacing of the acceptor site in the luminal loop of the membrane protein from the membrane surface (131, 197). The exact distance has been determined using *E. coli* leader peptidase as a model protein (130). It was found that there is a

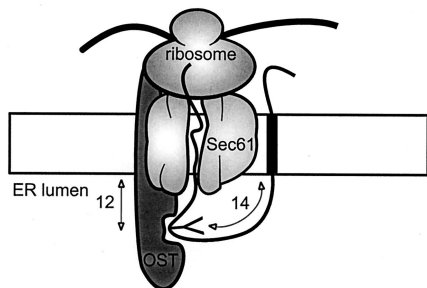


FIG. 3. ER translocase complex. The luminal OST is responsible for the glycosylation of membrane proteins in the ER. The active site of OST is located ~14 residues away from the end of a downstream transmembrane segment and ~12 residues away from an upstream hydrophobic segment.

precise distance constraint that allows glycosylation of acceptor sites only when the Asn residue is located a minimum of 12 residues upstream or 14 residues downstream of a transmembrane segment (Fig. 3). A survey of polytopic membrane proteins revealed that luminal loops containing N-linked oligosaccharides have a minimum size of about 30 residues, with the N-glycosylation acceptor site at least 10 residues away from the predicted transmembrane segment (104). N-glycosylation scanning mutagenesis of extracytosolic loops in the anion exchanger of the erythrocyte membrane Band 3 demonstrated that the distance requirements inferred from Leader peptidase applied to this protein (140). Introduced acceptor sites were glycosylated only in loops larger than 25 residues, and a sharp cutoff was observed in glycosylation of sites positioned too close to the membrane. Furthermore, insertion of a small glycosylation tag in short loops did not result in N glycosylation, whereas these loops were efficiently glycosylated when a large domain bearing a glycosylation site was inserted. The results account for the lack of glycosylation of endogenous (104) or introduced (140, 178) sites that are close to the membrane. The distance constraint between OST and the glycosylation site was used to map the ends of the transmembrane segments of Band 3, resulting in a refined topology model (140).

Glycosylation scanning. To study the topology of membrane proteins by using glycosylation scanning mutagenesis, the endogenous glycosylation sites in the protein must be removed and novel sites must be introduced at various positions. The mutations necessary to remove the endogenous sites and the resulting absence of glycosylation are not expected to have profound effects on the topology of membrane proteins. Indeed, the activities of several deglycosylated proteins, such as Band 3 (74), the erythrocyte glucose transporter GLUT1 (84), the glycine transporter GLYT1 (132), and the human Na^+ /glucose cotransporter SGLT1 (178), were comparable to those of the wild-type proteins, implying that glycosylation is not necessary for function and that the mutations to remove the glycosylation sites do not affect activity. The deglycosylated Na^+ / Cl^- -coupled γ -aminobutyric acid (GABA) transporter from rat brain (GAT-1) exhibited only 38% of wild-type activity, but it was shown that the reduction in activity was due at least in part to reduced trafficking to the plasma membrane (18).

Insertion of novel glycosylation sites in the hydrophilic domains of membrane proteins has a dramatic effect on activity. A 4-amino-acid insertion (NNSS) in the human Na^+ /glucose cotransporter (SGLT1) (178) at 13 different positions resulted in 10 insertion mutants with abolished transport activity. Transport was completely abolished upon insertion of one or two glycosylation sites in 6 of the 11 hydrophilic loops of the

GABA transporter GAT-1, and activity in the remaining loops was strictly dependent on the position in the loop (18). In contrast, all four cystic fibrosis transmembrane conductance regulator (CFTR) glycosylation insertion mutants were fully functional (32) and 10 of 15 mutants of the erythrocyte glucose transporter GLUT1 carrying a 41-amino-acid insert bearing a glycosylation site retained 2.5 to 30% of the deoxyglucose transport activity in oocytes (84). Since it cannot be ruled out that the inserted sequences scramble the topology of a protein, the use of inactive mutants for topology mapping is risky. On the other hand, it is generally believed that misfolded proteins are retained in biosynthetic compartments and rapidly degraded, implying that a construct which reaches the plasma membrane has a good chance of having the correct topology (84). In N-glycosylation scanning mutagenesis, internal loops score negative, but the accessibility of an external loop to the glycosylation enzymes may also be limited because of steric hindrance. Consequently, absence of glycosylation does not always disprove an extracellular localization. For example, insertion of several glycosylation sites at different positions in the most C-terminal loop of the glycine transporter (GLYT1) did not result in glycosylated proteins even though data obtained with other biochemical methods supported an extracellular location of this loop (132). Creation of a glycosylation site by replacement of two amino acids in the third extracellular loop of the GABA transporter resulted in a glycosylated protein, while creation of a double glycosylation site by insertion of 4 amino acids (NNSS) at the same position did not result in glycosylation (18). The results demonstrate that the ability to undergo glycosylation can be critically dependent on the position of the site in the loop and on the nature of the insert.

Steric factors, such as membrane proximity, that prevent glycosylation can be overcome by using longer insertion epitopes. The native glycosylation domain of SGLT1 or the entire first glycosylated luminal domain of GLUT4, consisting of 48 and 41 amino acids, respectively, resulted in significantly higher levels of glycosylation when inserted in SGLT1 (178) and GLUT1 (84) than was observed with smaller tags.

Glycosylation fusions. Domains bearing a glycosylation site have been used as reporter molecules in topology studies in very much the same way as alkaline phosphatase in PhoA fusions: the reporter molecule containing one or more glycosylation sites is fused behind different C-terminal deletions of a membrane protein. The domain will be glycosylated only when it is translocated to the luminal side of the ER membrane. Often, the technique is used in combination with *in vitro* translation and insertion in the presence or absence of insertion-competent microsomes derived from the ER membrane. Translation in the absence of microsomes results in a protein with the apparent molecular mass of the full-length fusion protein, while translation in the presence of microsomes results in the same or a higher apparent molecular mass depending on the cytoplasmic or luminal localization of the domain, respectively (see "Glycosylation of membrane proteins" above). The β -subunit of the gastric H^+ , K^+ -ATPase contains five N-linked glycosylation sites and has been used as a reporter molecule for the topological analysis of a number of membrane proteins, including the gastric H^+ , K^+ -ATPase (13), two rat Ca^{2+} ATPases (16) the rabbit gastric cholecystokinin A (CCK-A) receptor (17), and the glycine transporter GLYT1 (132). The membrane topology of the glutamate carrier GLT1 was studied by fusing a glycosylation consensus sequence behind a series of C-terminally truncated proteins (194). The *in vitro* translation-transcription system was also used to study the topology of the P-type ATPase from *Helicobacter pylori* (123) and the sodium ion-dependent citrate carrier of *Kleb-*

siella pneumoniae CitS (187), demonstrating that bacterial proteins can be expressed and analyzed in the ER system as well.

Chemical Modification

Cysteine residues in membrane proteins. The cysteine residue is a relatively hydrophobic, nonbulky residue, and its introduction at most positions in a membrane protein is likely to be tolerated. This feature and the ease of specific chemical modification with sulfhydryl reagents are the basis of several methods aiming at topological and structure-function-related information on membrane proteins. In cysteine scanning mutagenesis, a series of single cysteine mutants is created and the reactivity of the single cysteine mutant to various sulfhydryl reagents is assessed under different conditions. To obtain single Cys mutants, the native cysteine residues present in the membrane protein are removed and novel cysteine residues are introduced at specific positions in the Cys-less protein. The native cysteine residues are generally removed by replacement by serine residues, which usually renders an active protein. Substitution of amino acids with cysteine residues appears to be tolerated at most positions as well. The eight native cysteine residues of the lactose permease (LacY) of *E. coli* have been replaced simultaneously to yield a Cys-less permease that retained at least 50% of its wild-type activity (188). Furthermore, it was demonstrated that except for four amino acids, all the amino acids of LacY could be replaced by a cysteine residue without disrupting the function of the transporter (for reviews, see references 62 and 93). Thus, cysteine-scanning mutagenesis permits topology assessment and other structure-function studies under conditions in which the proteins are active and therefore, presumably, the protein structure is not seriously altered.

Cysteine-scanning mutagenesis. In topology studies, amino acids that are thought to reside in the putative extracellular or intracellular loops of a membrane protein are replaced by cysteine residues and the orientation with respect to the membrane is evaluated using membrane-permeable and -impermeable sulfhydryl reagents. Sulfhydryl reagents are available that contain either a biotin group (115), a fluorescent group (212), or a radiolabel (97), allowing detection of labelled proteins by streptavidin binding, fluorescence, or autoradiography, respectively. Labelling of a cysteine residue with a membrane-permeable reagent containing a detectable group after pretreatment of whole cells, spheroplasts, or right-side-out-oriented membrane vesicles with a membrane-impermeable reagent is indicative of a cytoplasmic cysteine residue, while the absence of labelling reveals a periplasmic location (Fig. 4). In addition to subcellular localization, labelling of a cysteine residue may depend on reactivity, possible steric constraints due to local structure, and, in cells and spheroplasts, scavenging of the label by thiol groups on cytoplasmic proteins. Therefore, it is essential to analyze labelling from both sides of the membrane either by performing the complementary experiment using inside-out membrane vesicles or by showing that the protein can be labelled in a noncompartmentalized system, i.e., in detergent solution. The sulfhydryl reagent reacts with cysteine residues present in all other proteins in the membrane, and therefore immunoprecipitation of the membrane protein or a purification step is necessary to show specific labelling.

Maleimide-based sulfhydryl reagents, which form a thioether bond with cysteine residues (Fig. 5), have been used in several topology studies including those of the MotA protein (212), P-glycoprotein (115), transposon 10-encoded metal-tetracycline/H⁺ antiporter (97), Na⁺/proline transporter (92), and subunit a of the *E. coli* F₀F₁-ATP synthase (114, 183). The

permeable maleimide compounds biotinmaleimide (92, 114, 115), N-[¹⁴C]ethylmaleimide (97), or fluoresceinmaleimide (183, 212) were used in combination with the membrane-impermeable maleimide compound acetomaleimide disulfonate.

A limitation of the cysteine-scanning technique is that not all residues are accessible to the sulfhydryl compounds. In particular, cysteine residues close to or within a putative transmembrane domain were found to be less reactive with biotin-containing compounds than those located closer to the middle of a hydrophilic loop (115). For this reason, several reagents should be tried. On the other hand, the use of small impermeable reagents allows the detection of residues in small loops that may not be accessible to macromolecular reagents such as antibodies, proteases, or the glycosylation machinery (35, 212). Native cysteine residues present in membrane proteins have been reported to be inaccessible to the sulfhydryl reagents as well. In those cases, it may not be necessary to construct a Cys-less protein prior to the introduction of new cysteine residues (188).

Additional advantages. An attractive feature of the use of cysteine scanning mutagenesis is that the library of functional single-cysteine mutants can be used for a more detailed structural analysis of the membrane protein. The *E. coli* LacY protein is the most complete example of cysteine-scanning mutagenesis to date, and its library of cysteine mutants (single and double Cys mutants) has been analyzed by a number of biophysical techniques aiming at obtaining both structural and dynamic information (62, 93). Helix packing, helix tilt, and ligand-induced conformational changes have been investigated using sulfhydryl-specific cross-linkers, spin labels, and fluorescent probes as well as site-directed peptide bond cleavage. Furthermore, cysteine-scanning mutagenesis in putative transmembrane segments has been used to explore the secondary structure of the segment in detail and to map channel-lining residues. Similar studies have been done for several other membrane proteins including the nicotinic acetylcholine receptor (1, 2, 4), the GABA receptor (203, 204), the cystic fibrosis transmembrane conductance regulator (CFTR) (3), the UhpT transporter (205), potassium channels (136), the tetracycline antiporter TetB (98), the erythrocyte anion exchanger AE1 (172), the GABA transporter GAT-1 (207), the oxalate/formate antiporter OxIT (64), and the serotonin transporter SERT (35). In these studies, sulfhydryl reagents that differ in charge, size, and hydrophilicity were used (88). These approaches are not discussed any further here.

Conclusions. Cysteine-scanning mutagenesis is the most useful technique developed for topology studies so far, at least for bacterial systems. The advantages are as follows: (i) the analysis is done on the complete and active protein; (ii) the modifications of the protein are minimal and restricted to a single amino acid residue change, and so they are not likely to disturb the membrane topology; (iii) the point mutations are well tolerated, usually leaving the proteins active; (iv) the actual analysis is done by chemical modification after insertion of the protein into the membrane; (v) the chemical modification can be done using whole cells, thereby avoiding problems related to the conversion of cells into membrane vesicles with a uniform orientation; and (vi) the library of single Cys mutants forms the basis for many other types of investigations. Disadvantages of the technique are as follows: (i) a Cys-less mutant of the protein has to be constructed, and (ii) the Cys residues in the single Cys mutants may not be accessible to the thiol reagents. The Cys-scanning mutagenesis approach has also been used for eukaryotic membrane proteins (18, 75, 93, 160) but has not yet been used in combination with *in vitro* translation/insertion using microsomes derived from the ER mem-

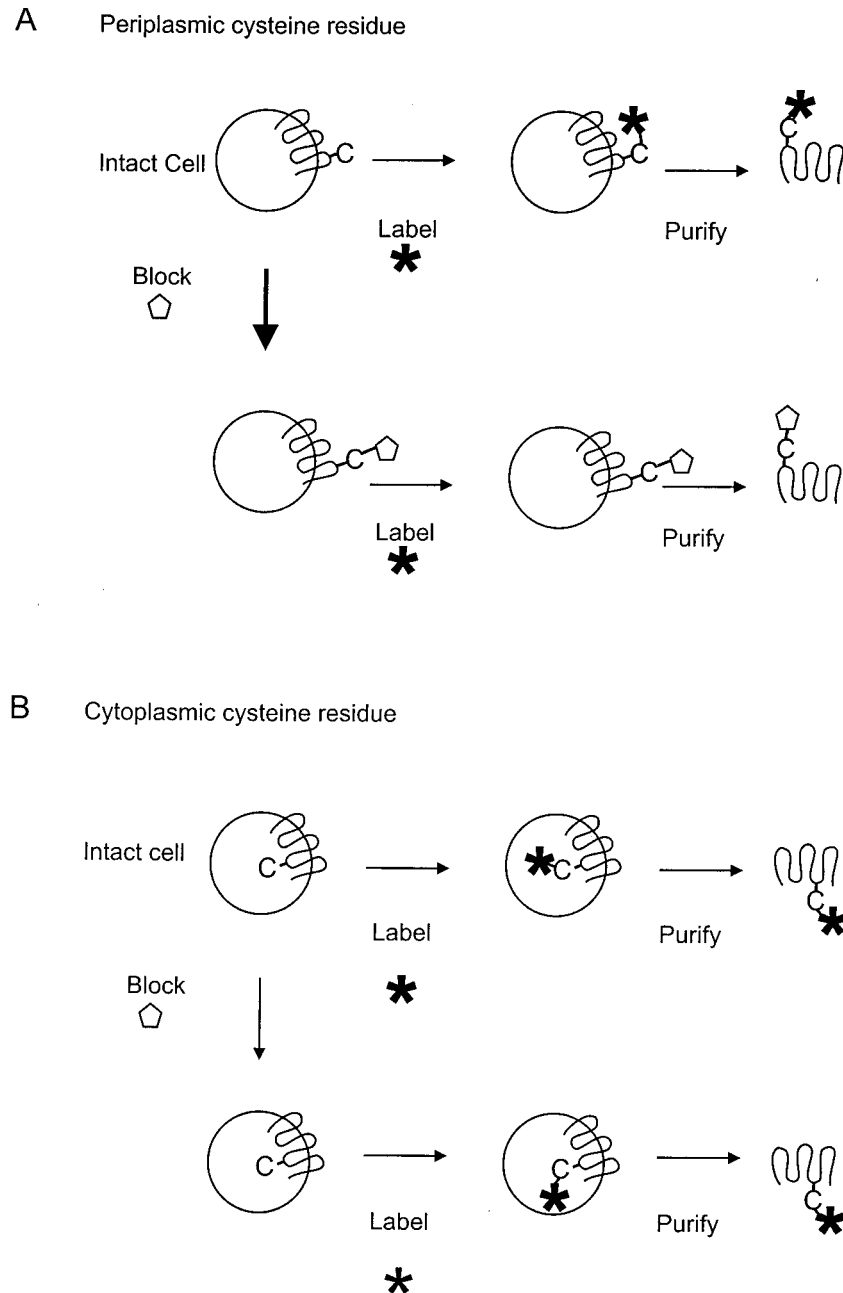


FIG. 4. Cysteine accessibility assay. Labeling of a periplasmic (A) and a cytoplasmic (B) cysteine residue in intact cells. The cells are treated with a detectable and membrane-permeant cysteine reagent (Label; from left to right) with or without pretreatment with a membrane-impermeable cysteine reagent (Block). Following the treatment, the protein is purified from the cells and assayed for labeling.

brane, which may have additional advantages. Besides cysteine residues, other amino acids may be useful. In a study to test the accessibility of predicted external loops of the serotonin neurotransmitter transporter, SERT, lysine residues were introduced as targets for the impermeant biotinylating reagent NHS-SS-biotin, which reacts specifically with lysine residues (35). A mutant SERT protein in which four native lysine residues in loop regions were replaced by arginine or glutamine residues showed normal transport activity. Similarly, a series of active mutant proteins containing novel single external lysine residues could be constructed, indicating that the replacement of lysine residues is tolerated, at least in external loops. A

disadvantage of using lysine residues is that most proteins contain many lysine residues, making it less easy to construct a functional protein devoid of lysine residues.

The BAD Tag

BAD. Biotin carboxylases are a family of enzymes that use a covalently attached biotin moiety as a mobile carboxy carrier to catalyze the transfer of CO_2 between metabolites (155). The biotin group is attached posttranslationally to a specific lysine residue of newly synthesized biotin carboxylases by the cytoplasmic enzyme biotin ligase, which uses ATP to activate the

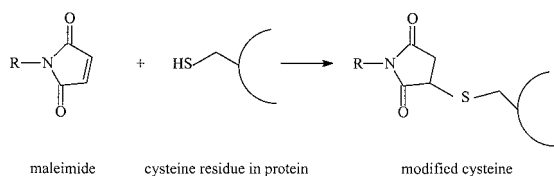


FIG. 5. Reaction of a thiol with a maleimide.

carboxy group of biotin during the reaction (Fig. 6) (163). Biotinylated proteins are rare in nature, varying from one to five in different organisms. The biotin carboxyl carrier protein of acetyl coenzyme A carboxylase is the only biotinylated protein found in *E. coli* (59). A high degree of sequence homology is observed around the biotin attachment site of all biotinylated proteins, and it was demonstrated that the biotin ligase enzyme from different sources is able to biotinylate acceptor proteins from various species both in vivo (43) and in vitro (121). The minimum size of the biotin domain necessary for biotinylation is about 75 to 80 residues (43, 110), and this domain, the biotin acceptor domain (BAD), exists as a stable entity. The crystal structure of the BAD of the *E. coli* biotin carboxyl carrier protein has been solved (11). Expression of gene fusions of the BAD of various carboxylases to the C terminus of different soluble proteins resulted in in vivo biotinylated proteins. Biotin binds to avidin with unusually high affinity (the dissociation constants are in the order of 10^{-15} M), yielding a complex that is extremely stable over a wide range of temperatures and pH values (72). The biotin-avidin system is used to detect biotinylated proteins with high sensitivity and to purify biotinylated proteins in one step by affinity chromatography using immobilized avidin (33, 43). Similar to soluble proteins, membrane proteins fused to BAD result in biotinylated proteins (41, 141, 184). Lac permease constructs with the BAD of the oxaloacetate decarboxylase of *K. pneumoniae* either inserted in the middle cytoplasmic loop or fused at the C terminus catalyze active transport and become biotinylated in vivo (180). Due to incomplete in vivo biotinylation, only a fraction of the transport protein was recovered upon purification using immobilized avidin. Recently, a method for in vitro biotinylation of the LacY-BAD fusion protein, resulting in complete biotinylation, has been described (142). Similarly, fusions of the same BAD to the C terminus of the citrate transporter CitS resulted in active and in vivo biotinylated protein and could be purified in a functional state (141). Membrane topology studies based on the BAD make use of the compartment-specific in vivo biotinylation of the domain (87, 209) or the high sensitivity to detect the biotinylated proteins in combination with proteolysis (184).

BAD in topology studies. The membrane topology of the lactose transporter of *E. coli* LacY has been determined by a variety of biochemical techniques. The BAD of the oxaloacetate decarboxylase from *K. pneumoniae* was inserted into various loops to investigate the suitability of the BAD for topology studies. The rationale was that BAD inserted in a cytoplasmic domain would be biotinylated because it is accessible to the cytoplasmic biotin ligase enzyme and, consequently, BAD inserted in a periplasmic domain would not be biotinylated. The following observations led to the conclusion that the approach was hindered by unexpected effects. Insertion of the BAD in two out of four cytoplasmic loops of LacY did not result in biotinylation of the domain (209). Lack of biotinylation was also observed when the BAD was inserted at one specific position in the most C-terminal cytoplasmic domain of the

citrate transporter CitS protein, while efficient biotinylation was observed when the domain was inserted at a different position in the same loop (M. van Geest and J. S. Lolkema, submitted for publication). Possibly, the biotinylation site is buried in the structure of the folded membrane protein, rendering the domain inaccessible to the biotin ligase enzyme. Alternatively, interactions of the domain with other regions of the membrane protein may prevent the proper folding of the BAD necessary for recognition by the biotin ligase enzyme. Insertions of the BAD into periplasmic domains of LacY resulted in biotinylated proteins that were all inactive in lactose transport (209). It could be shown that the domain was not translocated to the periplasm. Insertion of the BAD into a periplasmic loop blocks the translocation of the loop and, in all likelihood, the insertion of the two adjacent helices, resulting in a misfolded, inactive protein (209). In contrast, fusion of the BAD to the periplasmic C terminus of CitS resulted in an active fusion protein, and it was shown that the domain was translocated to the periplasm (184). Apparently, translocation of the domain is blocked only when sandwiched in a hydrophilic domain and not when fused to a hydrophilic domain, suggesting that the BAD is useful as a topological marker when using C-terminal deletion fusions. Biotinylation of the BAD of the 1.3S subunit of *Propionibacterium shermanii* transcarboxylase fused behind a series of C-terminal deletion mutants of the inner membrane protein MalF was indeed in agreement with the known topology of MalF (87). In fact, MalF-BAD fusions did not suffer from the problems encountered with MalF-PhoA fusions, in which the fusion site was at the beginning of a cytoplasmic loop, which resulted erroneously in the export of the PhoA moiety (see also "Complementarity" above) (26). In the corresponding MalF-BAD fusion, the BAD was securely locked in the cytoplasm, suggesting that the domain requires a less strong signal for cytoplasmic localization

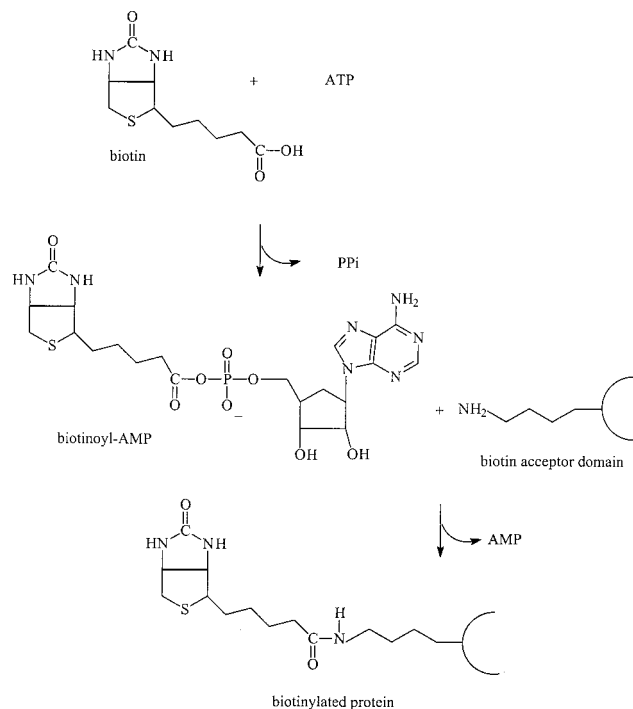


FIG. 6. In vivo biotinylation. Shown is the condensation reaction of biotin and a lysine residue, yielding a biotinylated protein. The reaction is catalyzed by the enzyme biotin ligase.

than does PhoA (87). This feature may be related to the fact that PhoA is a periplasmic protein while the BAD is part of a cytoplasmic protein.

In conclusion, C-terminal deletion fusions of the BAD may be a good alternative to LacZ fusions as positive indicators of cytoplasmic localization, with PhoA as the complementary positive indicator of periplasmic localization. Insertions of the BAD in putative cytoplasmic and periplasmic loops seem to result in many artifacts that make the domain less suitable as a topological marker. Only insertions that leave the protein active and that result in biotinylation report a trustworthy localization of the loop in the cytoplasm. Insertions of the BAD into putative cytoplasmic loops of CitS that did not compromise transport activity were used to verify a topology model of CitS that was based on C-terminal deletion PhoA fusions (van Geest and Lolkema, submitted) (see "Proteolysis" below).

Folding kinetics. Biotin acceptor domain fusions have been used to study the rate of translocation of secretory proteins or periplasmic domains of membrane proteins (87, 153). The degree of biotinylation is taken as a measure of the time spent in the cytoplasm before translocation. The degree of biotinylation is determined by the rate of biotinylation of the BAD by the biotin ligase in the cytoplasm and the rate of translocation to the periplasm catalyzed by the export machinery. When the BAD was fused to the periplasmic proteins PhoA or maltose binding protein, little or no biotinylation of the domain was observed in *E. coli*. Biotinylation of the domain was observed when the protein secretion process was slowed by mutations in components of the export machinery or by azide treatment, which specifically inhibits the ATP driven translocation. Thus, the biotinylation of the BAD functions as an indicator of impaired protein export. The rates of translocation of several periplasmic domains of MalF were studied upon sodium azide treatment or depletion of SecE (a component of the export machinery), demonstrating the involvement of the export machinery in the insertion mechanism of the MalF protein (87). Similar studies have been performed with the BAD fused to the periplasmic C terminus of the citrate transporter CitS. The level of biotinylation was much lower than observed for the BAD fused to the cytoplasmic N terminus, consistent with the periplasmic location of the domain. Upon addition of sodium azide to the growth medium, the level of biotinylation of the C-terminal fusion increased significantly (M. van Geest and J. S. Lolkema, unpublished results).

Proteolysis

Membrane proteins, like all proteins, contain cleavage sites for various proteolytic enzymes. Externally added proteolytic enzymes cannot cross the membrane, and therefore cytoplasmic sites are protected against cleavage by the membrane upon exposure of bacterial right-side-out membranes or spheroplasts whereas these sites are not protected against cleavage in inside-out membranes. Similarly, in ER microsomes, the lumenally exposed sites are protected against cleavage. Cleavage of the site(s) can be analyzed after sodium dodecyl sulfate-polyacrylamide gel electrophoresis and immunoblotting using antibodies directed against the protein. The technique has a clear advantage when it can be applied to native proteins without the need for modification. The limitations of the technique are that the naturally occurring cleavage sites are often not sensitive to the proteases and that more sites may be present, which complicates the analysis. To overcome the latter problem, proteolytic sites for specific proteases, for instance factor Xa, have been engineered into different loops of membrane proteins (154, 199). Similar to this are methods in which the accessibility

of hydrophilic loops to nonspecific proteases, such as proteinase K, is assayed after the insertion of a detectable proteic tag in the loop. Insertion into a cytoplasmic domain results in degradation of the tag upon exposure of inside-out membranes to proteinase K, and the protein or fragments thereof is no longer detectable. In contrast, treatment of spheroplasts or right-side-out membranes will result in the detection of a protected fragment. Insertion into a periplasmic loop gives the opposite result. Tags that have been used are the BAD of the oxaloacetate decarboxylase of *K. pneumoniae* (184, 209), a tag consisting of six histidine residues (His tag) (122, 185), and several antigenic epitopes, such as an epitope of the hemagglutinin of influenza virus (94–96), the 8-amino-acid FLAG epitope (135), the Myc epitope (31), and the secretory protein prolactin (38). Insertion of antigenic epitopes into membrane proteins which are expressed in eukaryotic cells allow the determination of the sidedness of the tag directly by immunofluorescence microscopy (94–96). The accessibility of an epitope tag or protease site to the antibody or protease has been improved by inserting multiple copies of the tag into the membrane protein (95, 96, 154). As with all these approaches, only functional insertion mutants are reliable for topology mapping, since their use guarantees that insertion of the tag did not alter the native membrane topology of the protein.

MEMBRANE INSERTION

Topology and Insertion

The previous part of this review presented an overview of biochemical approaches aimed at elucidating the membrane topology of polytopic membrane proteins. Some of the techniques work with intact and functional membrane proteins, but most techniques involve modified or truncated proteins. The changes to the proteins are made at the level of the coding sequences, while the results are interpreted in terms of the structure of the proteins as they sit in the membrane, thereby ignoring the steps in between that lead to the mature protein. Certain aspects of membrane protein structure may be properly understood only in the light of how these proteins are inserted into the lipid bilayer. Since membrane proteins contain domains that must be translocated across the bilayer, it is no surprise that the enzymatic machinery responsible for translocating proteins across the membrane also plays a role during the assembly of integral membrane proteins. An important notion is that the insertion machinery may not only insert the membrane protein but also determine its topology. Depending on the "communication" between the nascent polypeptide chain and the insertion machinery, the polypeptide is inserted into the membrane. Only when we are able to understand the "language" used by the polypeptide chain and the insertion machinery will it become possible to predict the complete structure of a membrane protein from the amino acid sequence. The success of the biochemical approaches that involve engineering of the proteins to determine membrane topology may improve significantly when we understand how the signals in the amino acid sequence interact with the insertion machinery. The next section briefly reviews the insertion machineries in the ER membrane and the bacterial cytoplasmic membrane and provides a discussion of several techniques used to detect topogenic signals in the polypeptide chain that trigger a certain response in the insertion machinery and thereby to determine the folding of the protein in the membrane.

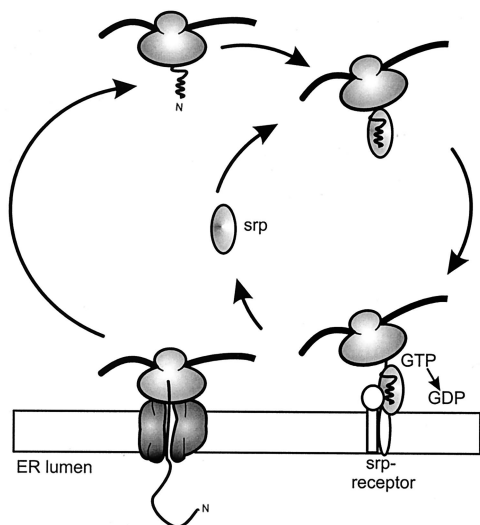


FIG. 7. Membrane protein insertion into the ER membrane. The SRP recognizes the first transmembrane segment emerging from the ribosome and targets the whole complex to the ER membrane. The ribosome-nascent chain complex is delivered at the translocon, where translation resumes and the protein is inserted into the membrane.

Membrane Protein Biogenesis

Insertion machinery in the ER membrane. In the ER, insertion of membrane proteins into the membrane and preprotein translocation into the lumen proceed by similar mechanisms (for reviews, see references 147 and 195). The process of targeting of membrane and secretory proteins begins in the cytosol when a hydrophobic segment of a nascent polypeptide chain, either a signal sequence or the first transmembrane segment, emerges from the ribosome and is recognized by the signal recognition particle (SRP) (Fig. 7). Binding of the SRP to the ribosome results in translational arrest, and the whole complex is targeted to translocation sites in the membrane via interactions with the SRP receptor. Translation resumes after dissociation of the SRP from the ribosome, and the nascent chain is transferred cotranslationally into the translocon. The ordered binding and recognition steps along this pathway are controlled by GTP binding and hydrolysis and GDP-GTP exchange (Fig. 7) (12, 40, 125, 146). The cotranslational integration of the membrane proteins is mediated by the translocon, consisting of the heterotrimeric Sec61 complex (120, 138, 159) and a number of other membrane proteins, including the translocating chain-associated membrane protein (TRAM) in mammalian cells (81, 82) or the heterotetrameric Sec62/63p complex in yeast (138). Since the Sec61 complex both translocates secretory proteins and integrates membrane proteins into the lipid bilayer, it must coordinate a number of different functions. For membrane proteins, cytoplasmic domains have to be left in the cytoplasm, luminal domains have to be passed through the membrane, and transmembrane domains have to be properly oriented and inserted into the membrane.

Insertion into the bacterial cytoplasmic membrane. (i) **Bacterial translocation machinery.** It is a matter of debate whether the same pathway is used in *E. coli* for export of preproteins and insertion of membrane proteins. In contrast to membrane protein insertion, translocation of preproteins across the inner membrane of *E. coli* has been studied extensively (50–52). Rather than occurring cotranslationally, preprotein translocation across the bacterial membrane occurs in a posttranslational manner and is carried out by a system involv-

ing chaperones like SecB and GroEL-GroES and the inner membrane components of the secretory (Sec) machinery, SecA, SecY, SecE, SecG, SecD, and SecF. Secretory proteins are synthesized as preproteins with an amino-terminal signal sequence. The signal sequence and/or mature domain of the preprotein is recognized in the cytoplasm by chaperones which ensure that the nascent secretory proteins do not fold or aggregate prematurely and which help to deliver the preproteins to specific receptors on the cytoplasmic membrane (for a review, see reference 60). The integral membrane components SecY, SecE, and SecG, which are homologous to subunits of the eukaryotic Sec61 complex, are believed to form a channel through which the nascent chain is extruded. SecA, which is peripherally bound to the SecYEG complex in the membrane and which does not have an eukaryotic counterpart, carries out an ATP-driven cycle of insertion and deinsertion into the translocase, during which 20 to 30 residues of the nascent chain are moved through the membrane (54, 179). In the bacterial system, SecA is the engine that drives translocation, while in the ER, translocation is driven by synthesis of the nascent chain on the ribosomes. The membrane-bound SecD and SecF components that are not essential for protein translocation per se stabilize SecA in its active conformation (53).

(ii) **Sec-dependent and Sec-independent insertion of membrane proteins.** Recent evidence has changed our views about membrane protein insertion in bacteria. Previously, it was believed that insertion occurred by two different mechanisms, a Sec-dependent mechanism and a Sec-independent one. Proteins could use one mechanism exclusively, or domains within a protein could use one or the other mechanism. A Sec-dependent mechanism was demonstrated for inner membrane proteins containing large periplasmic domains, such as the maltose transporter subunit MalF and leader peptidase Lep (176, 201). In contrast, insertion of proteins without large periplasmic domains was proposed to be Sec independent (14, 201). Sec dependency would be related to the length of periplasmic loops (9) and to the number of charged residues in the translocated loops (10). Sec-dependent insertion was believed to occur in a sequential, N- to C-terminal fashion, similar to that in the ER membrane (see “Insertion mechanism” below). Sec-independent insertion was, for thermodynamic reasons, thought to occur via “helical hairpins” composed of interacting antiparallel α -helices formed from two consecutive hydrophobic segments and their short intervening polar loops (58). Formation of the helical hairpin would sufficiently increase the net hydrophobicity of the sequence to drive insertion into the lipid bilayer. In summary, the Sec machinery would assist only in situations when insertion could not follow a simple spontaneous mechanism based on the physicochemical properties of the polypeptide chain and the membrane.

More recently, evidence has been accumulating that the *E. coli* proteins Ffh and FtsY (19, 150) play an important role in the assembly of inner membrane proteins. The Ffh and FtsY proteins are homologous to subunits of the eukaryotic SRP and SRP receptor proteins, suggesting an insertion pathway similar to the one in the ER (46, 162, 181, 182). Moreover, structural similarities between the *E. coli* SecYEG complex and the eukaryotic Sec61 complex (see also reference 113) make it likely that in *E. coli*, after targeting, membrane insertion proceeds via the Sec machinery. It was realized that studies that had indicated Sec-independent insertion were based on negative results obtained in experiments with conditional SecA and SecY strains that were selected primarily for secretion defects rather than for defects in the assembly of membrane proteins. Sec-independent insertion had been observed after sodium azide treatment, which is known to inhibit the ATPase

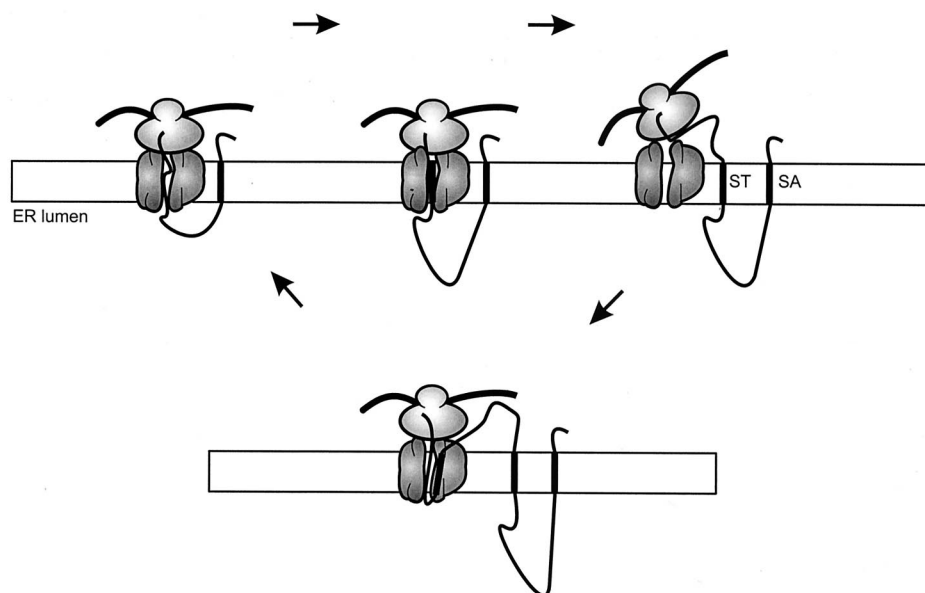


FIG. 8. Sequential insertion of hydrophobic sequences. Hydrophobic segments insert into the membrane as they emerge from the ribosome. The orientation of the segment is opposite to the orientation of the previous segment. A segment with an $N_{\text{cyt}}\text{-}C_{\text{lumen}}$ orientation is termed signal anchor (SA), and a segment with an $N_{\text{lumen}}\text{-}C_{\text{cyt}}$ orientation is termed stop transfer (ST).

activity of the SecA subunit only partially. Furthermore, it cannot be assumed a priori that translocon mutations have equal effects on the export of preproteins and insertion of membrane proteins (176). Very recently, Sec-dependent insertion in *E. coli* was demonstrated for three small bitopic membrane proteins with and without large periplasmic domains by using an *E. coli* strain with the SecE subunit depleted. Also, these proteins required the *E. coli* SRP for correct assembly into the inner membrane (46). These observations make it less likely that insertion of more complicated polytopic proteins bypasses the translocon. Recent reports actually showed that the SecY protein plays a role in determining the topology and insertion of membrane proteins in the bacterial membrane (128, 145).

Membrane protein insertion into the bacterial cytoplasmic membrane may be very similar to the process in the ER membrane. This view is supported by functional, heterologous expression of bacterial proteins in eukaryotic membranes and vice versa (73, 80).

Insertion mechanism. For reasons discussed above, most data on the actual insertion mechanism by the Sec machinery come from studies involving the eukaryotic ER system. The translocon forms an aqueous channel that permits the translocation of hydrophilic polypeptides (44, 69, 165) and opens laterally to allow the exit of transmembrane segments into the lipid bilayer (119, 127). Membrane integration is regulated by topogenic signals in the amino acid sequence of the nascent chain emerging from the ribosome (20). In a simple view, a series of alternating signal anchor (SA) and stop transfer (ST) sequences directs the integration of polytopic membrane proteins (22, 61, 112). A SA sequence is defined as a hydrophobic segment that has the ability to insert into the membrane with the N terminus in the cytoplasm and the C terminus in the lumen ($N_{\text{cyt}}\text{-}C_{\text{lum}}$ orientation). The sequence following the SA will be translocated to the lumen until translation of the next hydrophobic region, the ST sequence, which is retained in the membrane as an $N_{\text{lum}}\text{-}C_{\text{cyt}}$ transmembrane segment. The following sequences will have a cytoplasmic location until the next

SA sequence emerges from the ribosome and the cycle starts all over again (Fig. 8). Thus, according to this view, insertion of the polypeptide into the membrane occurs sequentially in an N- to C-terminal fashion, with the hydrophobic segments inserting into the membrane with alternating orientations. Consequently, the N-terminal segment determines the orientation of the protein in the membrane.

Although the assembly of some membrane proteins in the ER appears to occur via this mechanism, observations such as the exclusion of hydrophobic segments from the membrane (184, 187), the insertion of less hydrophobic sequences (13, 16, 17, 133), and observations made with artificial proteins (65, 66) and with C-terminally truncated membrane proteins (37, 199) have suggested that the insertion process may be more complicated, depending on the coordinate action of two or more topogenic sequences in the nascent chain. Recently, it has been documented that the ribosome plays a role in the insertion process (49, 111, 127). The ribosome would detect a newly synthesized transmembrane segment and use this information to influence the activity of the translocon (111). Structural changes at the ribosome may allow the ribosome and/or cytosolic proteins to facilitate the folding and processing of the nascent chain and/or determine the orientation of the nascent chain in the membrane. The transmembrane segments of the multidrug transporter P-glycoprotein did not integrate into the lipid bilayer until termination of synthesis, suggesting that the transmembrane segments remain associated with the insertion machinery until the polypeptide is released from the ribosome (23). Such observations suggest that the nascent chain contains, in addition to SA and ST segments, novel topogenic sequences that signal to and control the activity of the insertion machinery.

Studies aiming at unravelling the complicated insertion mechanism follow two lines of research: (i) mechanistic studies of the insertion machinery and (ii) identification of insertion and topogenic signals in the polypeptide sequence. The remainder of this review discusses the latter studies. Two steps are discriminated. The first is identification of segments in the

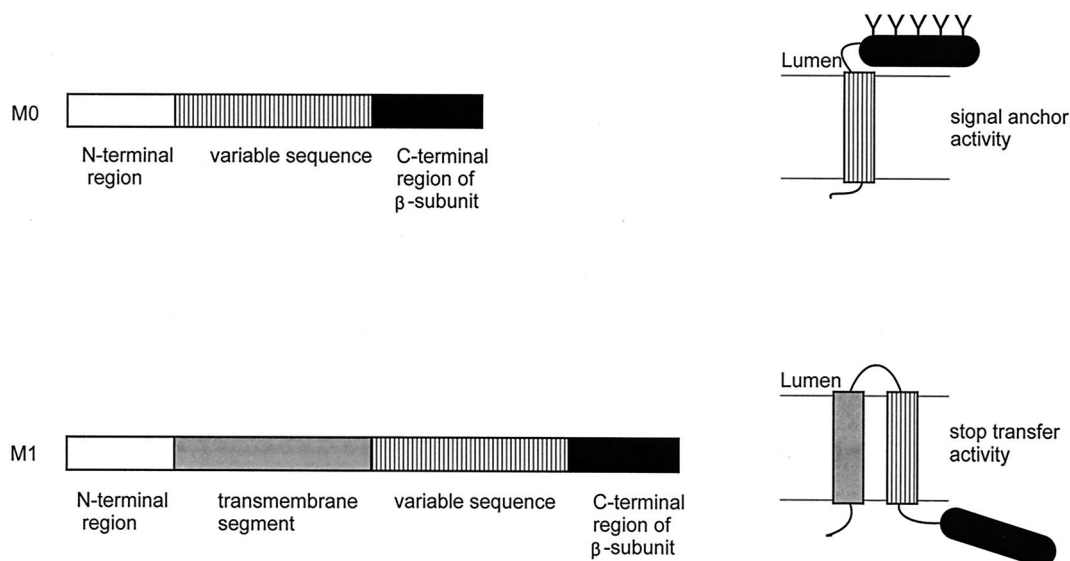


FIG. 9. M0 and M1 system. The variable sequence is preceded by the N-terminal region excluding (M0) or including (M1) the first transmembrane segment of the α -subunit of the gastric H^+, K^+ P-type ATPase and followed by the C-terminal region of the β -subunit of the same enzyme to assay for ST and SA activity, respectively (left). Insertion of the variable sequence into the membrane results in glycosylation of the C-terminal region of the β -subunit in the M0 vector and in the lack of glycosylation in the M1 vector (right).

amino acid sequence that have the potential to insert into the membrane in one or both orientations in the absence of other putative transmembrane segments or other topogenic signals. The second step is reevaluation of the insertion potencies of the putative transmembrane segments when combined with adjacent regions on the polypeptide chain that may or may not contain known insertion signals like charged residues. The second step reveals the relative strength of the signals and the presence of unknown signals. A putative transmembrane segment may be excluded from the membrane because its insertion is overruled by another stronger signal; alternatively, a less hydrophobic segment may be inserted because of strong orientation signals in adjacent transmembrane segments. Unexpected behavior may be helpful in decoding unknown signals in the amino acid sequence.

Insertion of Isolated Hydrophobic Segments

Principles. Sequences throughout the amino acid sequence of a polytopic membrane protein are identified that are able to insert independently of each other into the ER membrane, with a preferred orientation. Different algorithms (56, 99) that predict putative membrane-spanning sequences are used to guide the initial selection of potential candidates. The glycosylation fusion approach (see "Glycosylation fusions" above) has been used mainly to assay the sequences for their membrane insertion properties. Briefly, a reporter molecule containing one or more glycosylation sites is fused behind the sequence encoding the putative transmembrane segment and in vitro transcription-translation is carried out in the presence of microsomes. Glycosylation of the reporter corresponds to insertion of the segment into the ER membrane, indicating that the segment exhibits SA activity, whereas a lack of glycosylation corresponds to a lack of SA activity. Fusion of a sequence encoding a transmembrane segment in front of the putative segment indicates insertion of the segment in the opposite orientation, i.e., ST activity. Different types of systems have been developed for these studies. Insertion vehicles based on the H^+/K^+ -ATPase (M0 and M1 vectors), the *E. coli* inner

membrane protein leader peptidase (Lep), and the prolactin molecule have been used to analyze both eukaryotic (13, 133) and bacterial (123, 187) membrane proteins in the ER. Possible methods to perform similar studies in the bacterial membrane are discussed below (see "Similar studies in *E. coli*").

Insertion vehicles. (i) M0 and M1 system. The H^+/K^+ -ATPase is a P-type ATPase that pumps H^+ and K^+ in opposite directions across the gastric membrane, driven by ATP hydrolysis. The enzyme consists of two subunits, α and β , that both are integral membrane proteins. The C-terminal region of the β -subunit contains five N-linked glycosylation sites. A vector was developed containing the DNA sequences encoding the cytoplasmic N-terminal region of the α -subunit excluding (M0 vector) or including (M1 vector) the first transmembrane segment, followed by a variable region representing the amino acid sequence under investigation, which is followed by the C-terminal region of the β -subunit containing the glycosylation sites (Fig. 9). The M0 translation product, without an insertion in the variable region, was not glycosylated when synthesized in the presence of microsomes, demonstrating that the β -subunit fragment does not insert into the membrane (13). The M0 vector assays for the SA capacity of the inserted sequences, while the M1 vector identifies segments with ST activity. Membrane insertion sequences of the H^+/K^+ -ATPase (13), the ER and sarcoplasmic Ca^{2+} -ATPases (16), a P-type ATPase of *Helicobacter pylori* (123), and the rabbit gastric cholecystokinin A receptor (17) have been identified using the M0 and M1 system (15). Topogenic properties of putative transmembrane segments of the glycine transporter GLYT1 were analyzed in a similar system (132).

A drawback of the M0 fusion vector is the lack of an ER-targeting sequence preceding the variable region. As a result, the vector assays for both targeting and translocation and the intrinsic signal anchor capacity of a single segment cannot be measured. Insertion of a segment into the M0 vector, which results in a nonglycosylated product, does not necessarily reflect a segment incapable of insertion, as was concluded by various authors (13, 16, 17, 123).

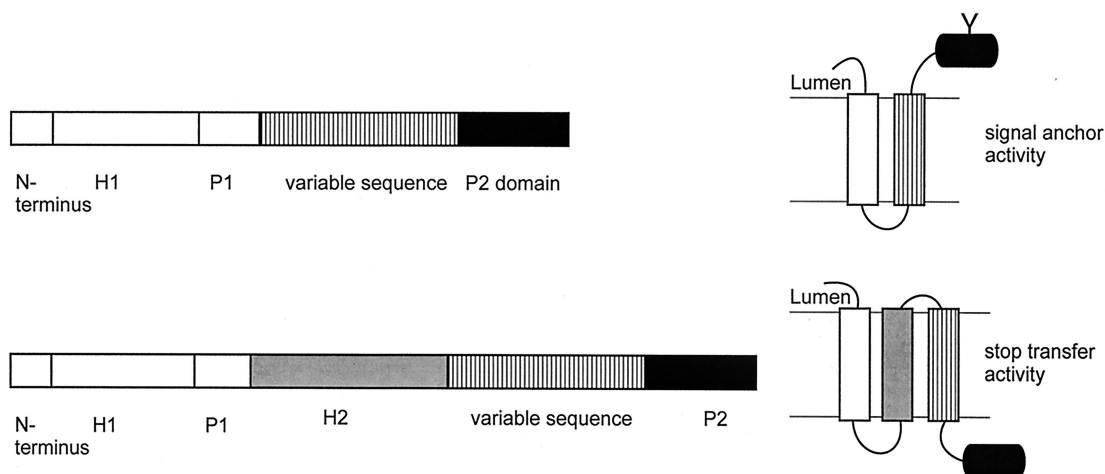


FIG. 10. Lep system. Vectors based on Leader peptidase of *E. coli*. The variable sequence replaces the second transmembrane segment, H2, of Lep (top) to measure SA activity or is inserted in the P2 domain to assay for ST activity (bottom). Membrane insertion of the variable sequence results in glycosylation and lack of glycosylation, respectively (right).

(ii) **Lep system.** The system based on the *E. coli* membrane protein leader peptidase (Lep) does not suffer from the above problem. Leader peptidase is involved in the removal of targeting sequences from exported proteins. Lep is anchored to the cytoplasmic membrane by two N-terminal transmembrane segments, H1 and H2, connected by a cytoplasmic P1 domain, and it also contains the periplasmic catalytic C-terminal P2 domain. The insertion of Lep into the microsomal membrane has been studied extensively (65, 89, 192) using *in vitro* transcription-translation. In the presence of microsomes, the Lep molecule inserts into the membrane with both the N and C termini facing the lumen and with the P1 and P2 domains in the cytoplasm and lumen, respectively. The H1 domain has an intrinsic ability to target to the ER membrane and to insert with the N terminus in the lumen. The orientation of the H1 segment is determined largely by the highly positively charged P1 loop. An engineered glycosylation site C-terminal of H2 is efficiently glycosylated upon correct insertion into the microsomal membrane (89, 130). In the insertion vehicle, the H2 domain is replaced by segments encoding individual putative transmembrane segments of membrane proteins and glycosylation of the P2 domain is used as an indicator of the SA activity of the segment (Fig. 10). Since all constructs contain the H1 domain and the P1 loop, targeting of the molecules to the membrane is mediated by H1 and the N-terminal end of the variable sequence is locked in the cytoplasm. Thus, the construct measures the “intrinsic” SA function of various fragments that have already been targeted to the microsomal membrane. A similar Lep construct, in which the variable sequence was inserted in between the H2 and P2 domain, was used to measure the ST efficiency of a segment. The constructs have been used to study ST activities of artificial peptides with different hydrophobicities (153) and SA and ST activities of the putative transmembrane segments of the bacterial citrate transporter CitS (187).

(iii) **Prolactin system.** Three types of insertion vehicles for isolated putative segments have been developed based on prolactin (Fig. 11). Prolactin is a secreted protein, and preprolactin consists of a signal sequence followed by the mature prolactin domain. A vector described by Kuroiwa et al. (100) measures ST activity. The signal peptide of prolactin, mature prolactin, an inserted transmembrane segment, and a second

mature prolactin molecule are arranged sequentially. Upon *in vitro* transcription-translation in the presence of microsomes, the whole construct is targeted to the membrane by the signal sequence and the first prolactin moiety is exported to the lumen. ST activity of the inserted transmembrane segment results in a cytoplasmic location of the second prolactin moiety, where it is sensitive to externally added proteinase K. Lack of ST activity results in translocation of the molecule to the lumen, where it is fully resistant to the protease.

An insertion vehicle designed to measure SA activity consists of the H1 domain of leader peptidase followed by the variable region, followed by the prolactin domain. An engineered glycosylation site in front of the H1 domain was used in an independent assay to ensure that the H1 domain inserted correctly. SA activity of the inserted transmembrane segment results in translocation of the prolactin domain to the lumen, where it is resistant to externally added proteinase K. When the inserted segment does not function as an SA, the molecule remains outside the membrane and is degraded by the protease (133).

A third vector was designed to examine the preferred orientation of the segment under investigation (133). The segment is preceded by a hydrophilic domain containing a glycosylation site and followed by mature prolactin. If the segment prefers to insert in an ST orientation, a glycosylated molecule is obtained in which the prolactin moiety is sensitive to proteases, while a preferred SA orientation results in the absence of glycosylation and a prolactin moiety that is protected against proteolytic attack.

The prolactin vectors have been used to study the topogenic functions of transmembrane segments of the human anion exchanger Band 3 (133), the multidrug transporter P-glycoprotein (78, 126, 210), the γ -aminobutyrate transporter GAT-1 (37), and the cystic fibrosis transmembrane conductance regulator, CFTR (116); they have also been used to study SA and ST functions of artificial sequences (101, 102). In some studies, the insertion vehicles contained different reporter molecules, such as interleukin-2 (101).

SA and ST activities of isolated transmembrane segments. Many hydrophobic segments of membrane proteins function equally efficient as SA and ST sequences when assayed in the M0 and M1 system. These include the four N-terminal seg-

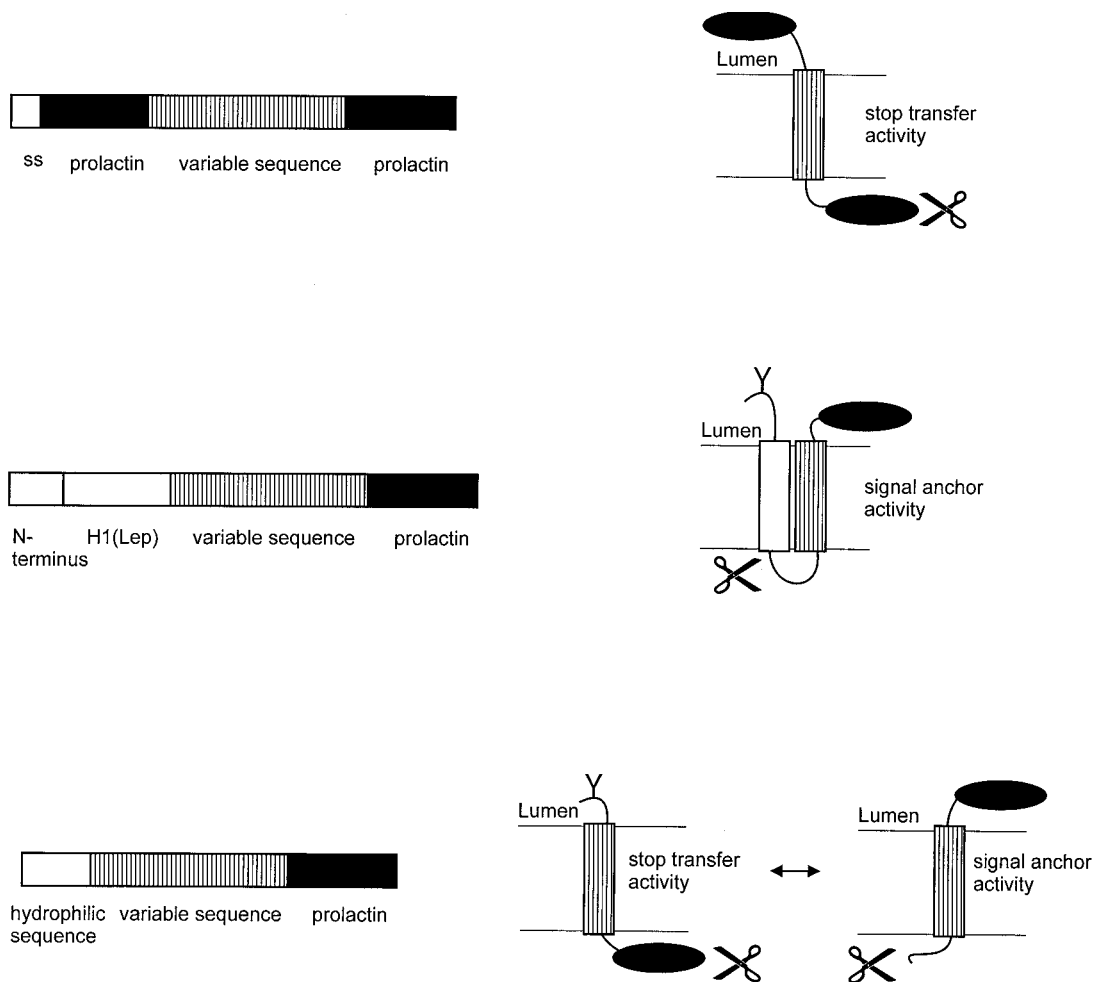


FIG. 11. Prolactin system. Shown are vectors based on the secretory protein prolactin. The variable sequence is inserted between preprolactin and prolactin (top), the H1 domain of leader peptidase and prolactin (middle), and a hydrophilic domain and prolactin (bottom). The assay of the membrane insertion activity of the variable sequence is based on protection of the prolactin domain against proteolytic degradation (right). The scissors represents the protease.

ments of the H^+/K^+ -ATPase (13) and the Ca^{2+} -ATPase (16) and most of the transmembrane segments of the CCK-A receptor (17). For the P-type ATPase of *H. pylori* (123), the ability to insert in both orientations seemed to correlate with the orientation of the segments in the native protein. All four segments expected to have an SA orientation in the native protein showed both SA and ST activity when assayed as individual segments. In contrast, three of the four segments expected to have an ST orientation in the native protein were not able to act as SA sequences but did show ST activity. Similar observations were made with ST segments of the H^+/K^+ -ATPase (13), the Ca^{2+} -ATPase (16), and the CCK-A receptor (17), suggesting that SA activity is more critically dependent on a particular segment than is ST activity. It should be noted, though, that the M0 and M1 system does not discriminate between SA activity and targeting (see above). Since the average hydrophobicity of the ST segments was not lower than the average hydrophobicity of the SA segments, factors other than hydrophobicity must play a role in membrane insertion or targeting.

The 12 hydrophobic segments of the citrate transporter CitS all demonstrated SA activity when assayed in the Lep insertion vehicle. Thus, in contrast to what was observed above, all ST segments in the native protein inserted equally well as SA

sequences in the membrane as isolated segments. No correlation between the hydrophobicity of the segment and insertion efficiency was seen. Interestingly, one hydrophobic segment of CitS that is located in the periplasm in the native protein (184, 187) was fully functional as SA and exhibited significant but incomplete ST activity. Apparently, signals present in the remaining part of the CitS polypeptide mediate the exclusion of this hydrophobic segment from the membrane during the insertion of CitS (see below).

Various sequences in the region of the fifth and sixth transmembrane domains of the H^+/K^+ -ATPase (13) and the Ca^{2+} -ATPase (16) did not show any membrane-inserting properties when assayed as individual segments in the M0 and M1 system. The average hydrophobicity of these segments was lower than that of the other segments. Similarly, one transmembrane segment of the CCK-A receptor did not show any membrane insertion properties and 4 out of 14 segments of the anion exchanger Band 3 that were supposed to be SA sequences showed insufficient SA activity. Since these segments are found to be transmembrane segments in the native proteins, their insertion in the membrane must be dependent on other signals in the polypeptide chain.

Similar studies in *E. coli*. The topogenic functions of isolated transmembrane segments in the *E. coli* membrane have

not been studied as extensively as those in the ER membrane, and approaches that combine insertion vehicles and *in vitro* translation-insertion have not been explored to their full potential. The efficiency of ST activity of a set of pseudo-random amino acid segments was measured both in the ER membrane and in *E. coli* (153). The *E. coli* study was done *in vivo* using a modified Lep molecule as the insertion vehicle in which the P2 domain was replaced by the mature part of alkaline phosphatase (PhoA). Most sequences behaved similarly in both systems, indicating that the criteria for functional ST sequences are largely conserved between prokaryotes and eukaryotes, but subtle differences were detected.

The insertion properties of various isolated sequences of the ArsB protein, a subunit of the bacterial arsenical resistance pump, were assayed *in vivo* in *E. coli* by fusing the segments in front of the β -lactamase reporter protein (202). Similarly, the first transmembrane segment of the *E. coli* serine chemoreceptor Tsr was fused in front of both β -galactosidase and the mature part of alkaline phosphatase (see "The enzymes: alkaline phosphatase, β -galactosidase, and β -lactamase" above) to assay the tolerance of its signal anchor function to mutations. It was concluded that a threshold hydrophobicity of the segment is required for efficient insertion into the membrane (107). The C-terminal transmembrane segment of the citrate transporter CitS, which has an $N_{\text{cyt}}\text{-}C_{\text{per}}$ orientation, was fused with its C terminus to the mature part of PhoA. Expression of the construct resulted in efficient secretion of the PhoA molecule to the periplasm, demonstrating independent insertion of the segment (van Geest and Lolkema, unpublished). Extension of these methods to other segments and other membrane proteins would allow a detailed comparison between interactions of the eukaryotic and bacterial insertion systems with nascent polypeptide chains and may provide important information about the insertion mechanisms in both systems.

Interaction between Multiple Insertion Signals

Charge distribution and hydrophobic segments. Besides hydrophobic segments, the distribution of positively charged amino acids appears to be an important determinant in membrane topology, in both eukaryotic and bacterial membranes. Statistical studies of bacterial cytoplasmic membrane proteins have demonstrated that cytoplasmic domains contain more positive charged amino acids than do periplasmic domains, which resulted in the so-called positive inside rule (190, 191). Statistical analyses indicated that a similar positive inside rule is applicable to eukaryotic membrane proteins (79, 90, 105, 166, 193), although the tendency for positive charges to be absent from extracellular domains is somewhat weaker (65). The influence of positive charges on the topology of single and double membrane spanning proteins (7, 8, 129, 190, 193) as well as of polytopic membrane proteins (65, 66) has been experimentally demonstrated. The positive charges prevent the translocation of N- and C-terminal segments as well as the connecting loops of polytopic membrane proteins and therefore determine to some extent the orientation of a transmembrane segment. Thus, hydrophobic segments and positive charges in the connecting loops represent two types of topogenic signals.

In marked contrast to the above results, introduction of a positively charged lysine residue into the cytoplasmic domain just downstream of the second transmembrane segment of the serine chemoreceptor protein Tsr of *E. coli* resulted in enhanced export of the domain (161). The lysine residue was introduced into a stretch of 11 residues that is strongly amphipathic when folded as an α -helix. A total of 10 different muta-

tions in the amphipathic stretch resulted in a correlation between enhanced export of the cytoplasmic domain and the amphipathicity of the stretch. It was suggested that an amphipathic sequence located immediately downstream of a transmembrane sequence represents a determinant of membrane topology.

Model proteins, consisting of four transmembrane segments and a distribution of positively charged residues in the loops that does not conform to the positive inside rule in either of the two orientations in the membrane ("frustrated proteins"), have been characterized in the *E. coli* (66) and the ER (65) membrane. In both membrane types, some proteins adopted a "leave-one-out" topology, in which only three segments inserted into the membrane but in which the highly positively charged loops were in the cytoplasm. Leave-one-out topologies were also observed for P-glycoprotein mutants (211) and GLUT1 glucose transporter mutants (158) as a result of a change in the distribution of positively charged residues in domains flanking transmembrane segments. The leave-one-out topologies demonstrate that the topogenic signal imposed by the distribution of positively charged residues may overrule the topogenic signal imposed by the hydrophobic segment; the distribution of positively charged residues is a stronger signal than a hydrophobic segment. These studies demonstrate that mechanisms exist by which a hydrophobic segment can be excluded from the membrane.

Interaction between different hydrophobic segments. (i) Insertion depending on upstream segments. In the topology model of the anion exchanger Band 3, which is based on various experimental approaches (for a review, see reference 173), the first transmembrane segment, an SA segment, is connected by a short loop to the second transmembrane segment, an ST segment. A C-terminal deletion mutant containing the first two segments folded as in the complete molecule. However, when the second segment was separated from the first transmembrane segment by a loop containing 200 residues, the second segment exhibited insufficient ST activity (133). ST activity of the second segment improved when it was located closer to the first segment, as in the original Band 3 molecule. Thus, the distance between the two segments was critical for the topogenic process. The authors proposed that with the short connecting loop the preceding sequence might still be in the translocon, allowing a direct interaction between the two hydrophobic segments which would make it easier for the second segment to stop in the translocon. In this way, a segment with a lower hydrophobicity may become integrated into the membrane. With the long connecting loop, the first transmembrane segment may leave the translocation channel prematurely and cannot "help" with the insertion of the second segment (133).

The citrate transporter CitS contains a hydrophobic segment that in the complete molecule is translocated to the periplasm (187). The segment exhibited both SA and ST activities when assayed as an individual segment. It appeared that the presence of the preceding transmembrane segment in CitS was largely responsible for the exclusion of the segment from the membrane. Possibly, interactions between the two segments in the translocon occur which direct the second hydrophobic segment from the translocon to the periplasm. The nature of these interactions is unknown.

(ii) Insertion depending on downstream segments. Some transmembrane segments of Band 3 which are inserted in the membrane in an SA orientation in the complete protein exhibited either insufficient or no insertion activity when assayed as individual segments (133). The segments are followed by highly hydrophobic segments with a strong preference for the

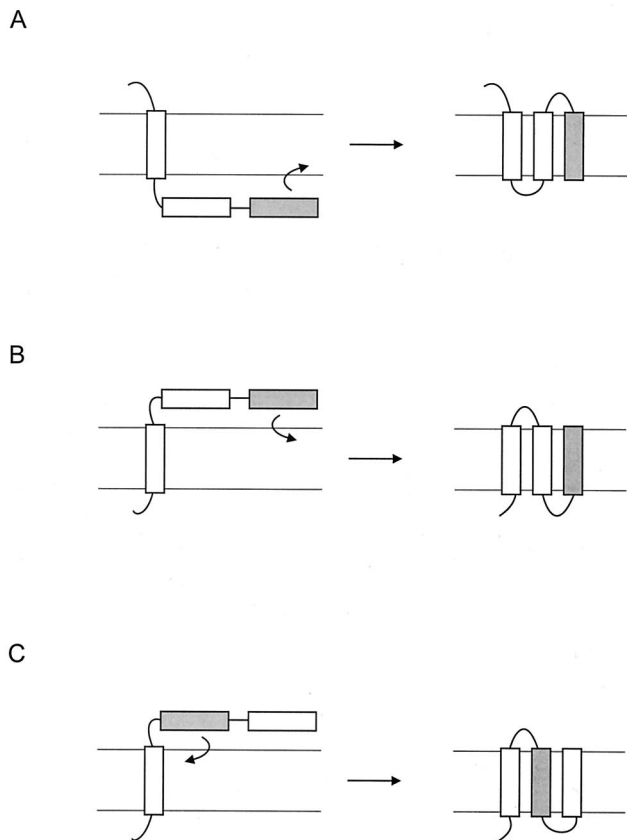


FIG. 12. Hypothetical folding intermediates. In the folding intermediates (left), two successive transmembrane segments are not inserted in the membrane. The hydrophobic segment, indicated as a gray box, drives the insertion of the preceding (A and B) or following (C) segment to yield the membrane topologies indicated on the right. The intermediates are based on observations made with the anion exchanger Band 3 (A), the protein translocation complex subunit Sec61 (A and B), and the citrate transporter CitS (C).

ST orientation over the SA orientation. The authors suggested that segments with weak insertion signals might be integrated into the membrane by the immediate downstream segments (134) (Fig. 12A).

Recently, it was demonstrated that instead of inserting into the membrane in an $N_{\text{per}}-C_{\text{cyt}}$ orientation, transmembrane segment VIII of the citrate transporter CitS is translocated into the periplasm in the absence of C-terminal sequences of the protein. The presence of downstream transmembrane segment IX appeared to be sufficient to mediate the insertion of both segments (186). In contrast to the Band 3 case above, segment VIII is a highly hydrophobic segment, which makes it difficult to understand why it would require the presence of the less hydrophobic segment IX for insertion. Instead, it could be segment IX that requires the presence of segment VIII for insertion. First, segment VIII would be translocated to the periplasm. Only when segment IX arrives in the periplasm would segment VIII drive the insertion of the helical hairpin consisting of segments VIII and IX (Fig. 12C). This sequence of events requires a mechanism that prevents segment VIII from inserting before segment IX is translated and translocated to the periplasm. Surprisingly, the same phenomenon was not observed upon insertion of CitS in the ER membrane (187).

A requirement for downstream sequences was demonstrated

for two transmembrane segments, with limited hydrophobicity, of the Sec61 protein from yeast (199). Segment TM2 required the presence of its downstream segment to stop translocation, while segment TM5 required C-terminal sequences to function as an SA sequence. Interestingly, in a construct consisting of the N-terminal five segments, TM5 was a transmembrane segment when the C-terminal sequences were coexpressed, suggesting that the insertion took place after the protein had left the translocon (Fig. 12B).

A requirement for downstream sequences was also observed for some of the transmembrane segments of the K^+, H^+ -ATPase (13) and for the amino-terminal segment TM1 of the CFTR (116).

(iii) **Folding intermediates.** The apparent interdependence between transmembrane segments for insertion may be a manifestation of the presence of folding intermediates during the insertion process. The intermediates may serve to insert segments that are relevant for catalysis and hence less hydrophobic. Figure 12 shows a number of hypothetical folding intermediates based on the examples above. Each intermediate requires very specific properties of the involved segments, resulting in specific actions of the insertion machinery to prevent a sequential insertion mechanism of the hydrophobic segments. A moderately hydrophobic segment at the cytoplasmic side of the membrane may be inserted into the membrane by the following hydrophobic segment (the "driving" segment) when the latter has a strong ST activity (Fig. 12A [based on Band 3 and Sec61 topologies]). In the folding intermediate, the hydrophobic driving segment does not enter the translocon as an SA sequence but is retained in the cytoplasm. Only when both segments are completed do the two insert as a helical hairpin. The latter action may take place in the translocon or after lateral movement into the lipid bilayer, either spontaneously or assisted by chaperone-like proteins. Similarly, a moderately hydrophobic periplasmic or luminal segment may be inserted by a downstream hydrophobic segment (Fig. 12B [based upon the Sec61 topology]). In the folding intermediate, the driving segment is exported to the periplasm. The loop between the two segments is initially translocated across the membrane to end up in the cytoplasm after the coordinate insertion of the two segments. Finally, the driving segment may also precede the less hydrophobic segment (Fig. 12C [based upon the CitS topology]). The insertion machinery translocates the driving segment across the membrane, where it waits for the less hydrophobic segment to be exported before both insert in concert.

The presence of folding intermediates as described above would imply that the folding of the polypeptide into the transmembrane segments cannot be understood solely based upon the energetics of transmembrane helix partitioning between the aqueous and membrane phases (139). The folding would be determined by the insertion machinery, and not all transmembrane domains may form autonomous folding domains. Importantly, such domains are likely to play a pivotal role in the function of the protein. For such membrane proteins, it is unlikely that they will refold from the denatured state into the native structure *in vitro*. The catalytically active structure of the protein is trapped in a local free energy minimum that is higher than the overall free energy minimum of the polypeptide and that cannot be reached in the absence of the biosynthetic machinery.

PROSPECTS

High-resolution three-dimensional structures are indispensable to our understanding of protein function. Given the enor-

mous number of sequences that are produced in genome-sequencing projects, it is not realistic to assume that the structures of all the encoded proteins will be generated by crystallographic approaches, especially for membrane proteins. The structures will have to be generated by computational techniques. This review emphasizes that biochemical approaches provide, besides topological information, important information about the insertion process. Only when the molecular mechanism behind the insertion of a membrane protein is understood will it become possible to predict the three-dimensional structures from the amino acid sequence by a simple successive decoding of all the topogenic signals present in the amino acid sequence. Knowledge about new topogenic sequences present in the nascent chain of membrane proteins and about the interactions between different topogenic signals will help to develop better prediction algorithms and biochemical techniques to predict and determine membrane protein structure.

REFERENCES

1. Akabas, M. H., and A. Karlin. 1995. Identification of acetylcholine receptor channel-lining residues in the M1 segment of the alpha-subunit. *Biochemistry* **34**:12496–12500.
2. Akabas, M. H., C. Kaufmann, P. Archdeacon, and A. Karlin. 1994. Identification of acetylcholine receptor channel-lining residues in the entire M2 segment of the alpha subunit. *Neuron* **13**:919–927.
3. Akabas, M. H., C. Kaufmann, T. A. Cook, and P. Archdeacon. 1994. Amino acid residues lining the chloride channel of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **269**:14865–14868.
4. Akabas, M. H., D. A. Stauffer, M. Xu, and A. Karlin. 1992. Acetylcholine receptor channel structure probed in cysteine-substitution mutants. *Science* **258**:307–310.
5. Akiyama, Y., and K. Ito. 1993. Folding and assembly of bacterial alkaline phosphatase in vitro and in vivo. *J. Biol. Chem.* **268**:8146–8150.
6. Allard, J. D., and K. P. Bertrand. 1992. Membrane topology of the pBR322 tetracycline resistance protein. TetA-PhoA gene fusions and implications for the mechanism of TetA membrane insertion. *J. Biol. Chem.* **267**:17809–17819.
7. Andersson, H., E. Bakker, and G. von Heijne. 1992. Different positively charged amino acids have similar effects on the topology of a polytopic transmembrane protein in *Escherichia coli*. *J. Biol. Chem.* **267**:1491–1495.
8. Andersson, H., and G. von Heijne. 1993. Position-specific Asp-Lys pairing can affect signal sequence function and membrane protein topology. *J. Biol. Chem.* **268**:21389–21393.
9. Andersson, H., and G. von Heijne. 1993. Sec dependent and sec independent assembly of *E. coli* inner membrane proteins: the topological rules depend on chain length. *EMBO J.* **12**:683–691.
10. Andersson, H., and G. von Heijne. 1994. Positively charged residues influence the degree of SecA dependence in protein translocation across the *E. coli* inner membrane. *FEBS Lett.* **347**:169–172.
11. Athappilly, F. K., and W. A. Hendrickson. 1995. Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**:1407–1419.
12. Bacher, G., H. Lutcke, B. Jungnickel, T. A. Rapoport, and B. Dobberstein. 1996. Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. *Nature* **381**:248–251.
13. Bamberg, K., and G. Sachs. 1994. Topological analysis of H⁺, K⁺-ATPase using in vitro translation. *J. Biol. Chem.* **269**:16909–16919.
14. Bassilana, M., and C. Gwizdek. 1996. In vivo membrane assembly of the *E. coli* polytopic protein, melibiose permease, occurs via a Sec-independent process which requires the protonmotive force. *EMBO J.* **15**:5202–5208.
15. Bayle, D., D. Weeks, S. Hallen, K. Melchers, K. Bamberg, and G. Sachs. 1997. In vitro translation analysis of integral membrane proteins. *J. Recept. Signal Transduct. Res.* **17**:29–56.
16. Bayle, D., D. Weeks, and G. Sachs. 1995. The membrane topology of the rat sarcoplasmic and endoplasmic reticulum calcium ATPases by in vitro translation scanning. *J. Biol. Chem.* **270**:25678–25684.
17. Bayle, D., D. Weeks, and G. Sachs. 1997. Identification of membrane insertion sequences of the rabbit gastric cholecystokinin-A receptor by in vitro translation. *J. Biol. Chem.* **272**:19697–19707.
18. Bennett, E. R., and B. I. Kanner. 1997. The membrane topology of GAT-1, a (Na⁺ + Cl⁻)-coupled gamma-aminobutyric acid transporter from rat brain. *J. Biol. Chem.* **272**:1203–1210.
19. Bernstein, H. D., M. A. Poritz, K. Strub, P. J. Hoben, S. Brenner, and P. Walter. 1989. Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. *Nature* **340**:482–486.
20. Bibi, E. 1998. The role of the ribosome-translocon complex in translation and assembly of polytopic membrane proteins. *Trends Biochem. Sci.* **23**:51–55.
21. Bibi, E., and O. Beja. 1994. Membrane topology of multidrug resistance protein expressed in *Escherichia coli* N-terminal domain. *J. Biol. Chem.* **269**:19910–19915.
22. Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA* **77**:1496–1500.
23. Borel, A. C., and S. M. Simon. 1996. Biogenesis of polytopic membrane proteins: membrane segments assemble within translocation channels prior to membrane integration. *Cell* **85**:379–389.
24. Boyd, D. 1994. Use of gene fusions to determine membrane protein topology, p. 144–163. In S. H. White (ed.), *Membrane protein structure. Experimental approaches*. Oxford University Press, Oxford, United Kingdom.
25. Boyd, D., and J. Beckwith. 1990. The role of charged amino acids in the localization of secreted and membrane proteins. *Cell* **62**:1031–1033.
26. Boyd, D., B. Traxler, and J. Beckwith. 1993. Analysis of the topology of a membrane protein by using a minimum number of alkaline phosphatase fusions. *J. Bacteriol.* **175**:553–556.
27. Broome-Smith, J. K., and B. G. Spratt. 1986. A vector for the construction of translational fusions to TEM beta-lactamase and the analysis of protein export signals and membrane protein topology. *Gene* **49**:341–349.
28. Broome-Smith, J. K., M. Tadayyon, and Y. Zhang. 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol. Microbiol.* **4**:1637–1644.
29. Calamia, J., and C. Manoil. 1990. lac permease of *Escherichia coli*: topology and sequence elements promoting membrane insertion. *Proc. Natl. Acad. Sci. USA* **87**:4937–4941.
30. Chang, C. N., W. J. Kuang, and E. Y. Chen. 1986. Nucleotide sequence of the alkaline phosphatase gene of *Escherichia coli*. *Gene* **44**:121–125.
31. Chang, H. C., and D. R. Bush. 1997. Topology of NAT2, a prototypical example of a new family of amino acid transporters. *J. Biol. Chem.* **272**:30552–30557.
32. Chang, X. B., Y. X. Hou, T. J. Jensen, and J. R. Riordan. 1994. Mapping of cystic fibrosis transmembrane conductance regulator membrane topology by glycosylation site insertion. *J. Biol. Chem.* **269**:18572–18575.
33. Chapman-Smith, A., D. L. Turner, J. E. Cronan, Jr., T. W. Morris, and J. C. Wallace. 1994. Expression, biotinylation and purification of a biotin-domain peptide from the biotin carboxy carrier protein of *Escherichia coli* acetyl-CoA carboxylase. *Biochem. J.* **302**:881–887.
34. Chavez, R. A., and Z. W. Hall. 1991. The transmembrane topology of the amino terminus of the alpha subunit of the nicotinic acetylcholine receptor. *J. Biol. Chem.* **266**:15532–15538.
35. Chen, J. G., S. Liu-Chen, and G. Rudnick. 1998. Determination of external loop topology in the serotonin transporter by site-directed chemical labeling. *J. Biol. Chem.* **273**:12675–12681.
36. Chepur, V., and R. B. Gennis. 1990. The use of gene fusions to determine the topology of all of the subunits of the cytochrome c terminal oxidase complex of *Escherichia coli*. *J. Biol. Chem.* **265**:12978–12986.
37. Clark, J. A. 1997. Analysis of the transmembrane topology and membrane assembly of the GAT-1 gamma-aminobutyric acid transporter. *J. Biol. Chem.* **272**:14695–14704.
38. Clark, J. A. 1998. Analysis of transporter topology using deletion and epitope tagging. *Methods Enzymol.* **296**:293–307.
39. Claros, M. G., and G. von Heijne. 1994. TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* **10**:685–686.
40. Connolly, T., and R. Gilmore. 1993. GTP hydrolysis by complexes of the signal recognition particle and the signal recognition particle receptor. *J. Cell Biol.* **123**:799–807.
41. Consler, T. G., B. L. Persson, H. Jung, K. H. Zen, K. Jung, G. G. Prive, G. E. Verner, and H. R. Kaback. 1993. Properties and purification of an active biotinylated lactose permease from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **90**:6934–6938.
42. Cosgriff, A. J., and A. J. Pittard. 1997. A topological model for the general aromatic amino acid permease, AroP, of *Escherichia coli*. *J. Bacteriol.* **179**:3317–3323.
43. Cronan, J. E., Jr. 1990. Biotinylation of proteins in vivo. A post-translational modification to label, purify, and study proteins. *J. Biol. Chem.* **265**:10327–10333.
44. Crowley, K. S., S. Liao, V. E. Worrell, G. D. Reinhart, and A. E. Johnson. 1994. Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell* **78**:461–471.
45. Danielsen, S., D. Boyd, and J. Neuhard. 1995. Membrane topology analysis of the *Escherichia coli* cytosine permease. *Microbiology* **141**:2905–2913.
46. De Gier, J. W., P. A. Scotti, A. Saaf, Q. A. Valent, A. Kuhn, J. Luijck, and G. von Heijne. 1998. Differential use of the signal recognition particle translocase targeting pathway for inner membrane protein assembly in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **95**:14646–14651.
47. Derman, A. I., and J. Beckwith. 1991. *Escherichia coli* alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. *J. Bacteriol.* **173**:7719–7722.
48. Derman, A. I., W. A. Prinz, D. Belin, and J. Beckwith. 1993. Mutations that

- allow disulfide bond formation in the cytoplasm of *Escherichia coli*. *Science* **262**:1744–1747.
49. Do, H., D. Falcone, J. Lin, D. W. Andrews, and A. E. Johnson. 1996. The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* **85**:369–378.
 50. Driessen, A. J., P. Fekkes, and J. P. van der Wolk. 1998. The Sec system. *Curr. Opin. Microbiol.* **1**:216–222.
 51. Driessen, A. J. M. 1996. Translocation of proteins across the bacterial cytoplasmic membrane, p. 759–790. *In* W. N. Konings, H. R. Kaback, and J. S. Lolkema (ed.), *Transport processes in membranes*. Elsevier Biomedical Press, Amsterdam, The Netherlands.
 52. Duong, F., J. Eichler, A. Price, M. R. Leonard, and W. Wickner. 1997. Biogenesis of the gram-negative bacterial envelope. *Cell* **91**:567–573.
 53. Duong, F., and W. Wickner. 1997. The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. *EMBO J.* **16**:4871–4879.
 54. Economou, A., and W. Wickner. 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* **78**:835–843.
 55. Ehrmann, M., D. Boyd, and J. Beckwith. 1990. Genetic analysis of membrane protein topology by a sandwich gene fusion approach. *Proc. Natl. Acad. Sci. USA* **87**:7574–7578.
 56. Eisenberg, D., E. Schwarz, M. Komaromy, and R. Wall. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**:125–142.
 57. Eitinger, T., and B. Friedrich. 1994. A topological model for the high-affinity nickel transporter of *Alcaligenes eutrophus*. *Mol. Microbiol.* **12**:1025–1032.
 58. Engelman, D. M., and T. A. Steitz. 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell* **23**:411–422.
 59. Fall, R. R. 1979. Analysis of microbial biotin proteins. *Methods Enzymol.* **62**:390–398.
 60. Fekkes, P., and A. J. Driessen. 1999. Protein targeting to the bacterial cytoplasmic membrane. *Microbiol. Mol. Biol. Rev.* **63**:161–173.
 61. Friedlander, M., and G. Blobel. 1985. Bovine opsin has more than one signal sequence. *Nature* **318**:338–343.
 62. Frillingos, S., M. Sahin-toth, J. Wu, and H. R. Kaback. 1998. Cys-scanning mutagenesis: a novel approach to structure function relationships in polytopic membrane proteins. *FASEB J.* **12**:1281–1299.
 63. Froshauer, S., G. N. Green, D. Boyd, K. McGovern, and J. Beckwith. 1988. Genetic analysis of the membrane insertion and topology of MalF, a cytoplasmic membrane protein of *Escherichia coli*. *J. Mol. Biol.* **200**:501–511.
 64. Fu, D., and P. C. Maloney. 1998. Structure-function relationships in OxlT, the oxalate/formate transporter of *Oxalobacter formigenes*. Topological features of transmembrane helix 11 as visualized by site-directed fluorescent labeling. *J. Biol. Chem.* **273**:17962–17967.
 65. Gafvelin, G., M. Sakaguchi, H. Andersson, and G. von Heijne. 1997. Topological rules for membrane protein assembly in eukaryotic cells. *J. Biol. Chem.* **272**:6119–6127.
 66. Gafvelin, G., and G. von Heijne. 1994. Topological “frustration” in multi-spanning *E. coli* inner membrane proteins. *Cell* **77**:401–412.
 67. Geller, D., D. Taglicht, R. Edgar, A. Tam, O. Pines, S. Michaelis, and E. Bibi. 1996. Comparative topology studies in *Saccharomyces cerevisiae* and in *Escherichia coli*. The N-terminal half of the yeast ABC protein Ste6. *J. Biol. Chem.* **271**:13746–13753.
 68. Georgiou, C. D., T. J. Dueweke, and R. B. Gennis. 1988. Beta-galactosidase gene fusions as probes for the cytoplasmic regions of subunits I and II of the membrane-bound cytochrome d terminal oxidase from *Escherichia coli*. *J. Biol. Chem.* **263**:13130–13137.
 69. Gilmore, R., and G. Blobel. 1985. Translocation of secretory proteins across the microsomal membrane occurs through an environment accessible to aqueous perturbants. *Cell* **42**:497–505.
 70. Gott, P., and W. Boos. 1988. The transmembrane topology of the sn-glycerol-3-phosphate permease of *Escherichia coli* analysed by phoA and lacZ protein fusions. *Mol. Microbiol.* **2**:655–663.
 71. Gowrishankar, J. 1989. Nucleotide sequence of the osmoregulatory *proU* operon of *Escherichia coli*. *J. Bacteriol.* **171**:1923–1931.
 72. Green, N. M. 1975. Avidin. *Adv. Protein Chem.* **29**:85–133.
 73. Grishammer, R., and C. G. Tate. 1995. Overexpression of integral membrane proteins for structural studies. *Q. Rev. Biophys.* **28**:315–422.
 74. Groves, J. D., and M. J. Tanner. 1994. Role of N-glycosylation in the expression of human band 3-mediated anion transport. *Mol. Membr. Biol.* **11**:31–38.
 75. Grunewald, M., A. Bendahan, and B. I. Kanner. 1998. Biotinylation of single cysteine mutants of the glutamate transporter GLT-1 from rat brain reveals its unusual topology. *Neuron* **21**:623–632.
 76. Haardt, M., and E. Bremer. 1996. Use of phoA and lacZ fusions to study the membrane topology of ProW, a component of the osmoregulated ProU transport system of *Escherichia coli*. *J. Bacteriol.* **178**:5370–5381.
 77. Hagting, A., J. van der Velde, B. Poolman, and W. N. Konings. 1997. Membrane topology of the di- and tripeptide transport protein of *Lactococcus lactis*. *Biochemistry* **36**:6777–6785.
 78. Han, E. S., and J. T. Zhang. 1998. Mechanism involved in generating the carboxyl-terminal half topology of P-glycoprotein. *Biochemistry* **37**:11996–12004.
 79. Hartmann, E., T. A. Rapoport, and H. F. Lodish. 1989. Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci. USA* **86**:5786–5790.
 80. Hennessey, E. S., L. Hashemzadeh-Bonehi, L. A. Hunt, and J. K. Broome-Smith. 1993. Assembly of eukaryotic class III (N-out, C-in) membrane proteins into the *Escherichia coli* cytoplasmic membrane. *FEBS Lett.* **331**:159–161.
 81. High, S., S. S. Andersen, D. Gorlich, E. Hartmann, S. Prehn, T. A. Rapoport, and B. Dobberstein. 1993. Sec61p is adjacent to nascent type I and type II signal-anchor proteins during their membrane insertion. *J. Cell Biol.* **121**:743–750.
 82. High, S., B. Martoglio, D. Gorlich, S. S. Andersen, A. J. Ashford, A. Giner, E. Hartmann, S. Prehn, T. A. Rapoport, and B. Dobberstein. 1993. Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence. *J. Biol. Chem.* **268**:26745–26751.
 83. Hipfner, D. R., K. C. Almquist, E. M. Leslie, J. H. Gerlach, C. E. Grant, R. G. Deeley, and S. P. Cole. 1997. Membrane topology of the multidrug resistance protein (MRP). A study of glycosylation-site mutants reveals an extracytosolic NH2 terminus. *J. Biol. Chem.* **272**:23623–23630.
 84. Hresko, R. C., M. Kruse, M. Strube, and M. Mueckler. 1994. Topology of the Glut 1 glucose transporter deduced from glycosylation scanning mutagenesis. *J. Biol. Chem.* **269**:20482–20488.
 85. Iwata, S., J. W. Lee, K. Okada, J. K. Lee, M. Iwata, B. Rasmussen, T. A. Link, S. Ramaswamy, and B. K. Jap. 1998. Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex. *Science* **281**:64–71.
 86. Iwata, S., C. Ostermeier, B. Ludwig, and H. Michel. 1995. Structure at 2.8 Å resolution of cytochrome c oxidase from *Paracoccus denitrificans*. *Nature* **376**:660–669.
 87. Jander, G., J. E. Cronan, Jr., and J. Beckwith. 1996. Biotinylation in vivo as a sensitive indicator of protein secretion and membrane protein insertion. *J. Bacteriol.* **178**:3049–3058.
 88. Javitch, J. A. 1998. Probing structure of neurotransmitter transporters by substituted-cysteine accessibility method. *Methods Enzymol.* **296**:331–346.
 89. Johansson, M., I. Nilsson, and G. von Heijne. 1993. Positively charged amino acids placed next to a signal sequence block protein translocation more efficiently in *Escherichia coli* than in mammalian microsomes. *Mol. Gen. Genet.* **239**:251–256.
 90. Jones, D. T., W. R. Taylor, and J. M. Thornton. 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**:3038–3049.
 91. Jung, H. 1998. Topology and function of the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute cotransporter family. *Biochim. Biophys. Acta* **1365**:60–64.
 92. Jung, H., R. Rubenhagen, S. Tebbe, K. Leifker, N. Tholema, M. Quick, and R. Schmid. 1998. Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* **273**:26400–26407.
 93. Kaback, H. R. 1996. The lactose permease of *Escherichia coli*: past, present and future, p. 203–227. *In* W. N. Konings, H. R. Kaback, and J. S. Lolkema (ed.), *Handbook of biological physics. Transport processes in eukaryotic and prokaryotic organisms*. Elsevier Science, Amsterdam, The Netherlands.
 94. Kast, C., V. Canfield, R. Levenson, and P. Gros. 1995. Membrane topology of P-glycoprotein as determined by epitope insertion: transmembrane organization of the N-terminal domain of mdr3. *Biochemistry* **34**:4402–4411.
 95. Kast, C., V. Canfield, R. Levenson, and P. Gros. 1996. Transmembrane organization of mouse P-glycoprotein determined by epitope insertion and immunofluorescence. *J. Biol. Chem.* **271**:9240–9248.
 96. Kast, C., and P. Gros. 1998. Epitope insertion favors a six transmembrane domain model for the carboxy-terminal portion of the multidrug resistance-associated protein. *Biochemistry* **37**:2305–2313.
 97. Kimura, T., M. Ohnuma, T. Sawai, and A. Yamaguchi. 1997. Membrane topology of the transposon 10-encoded metal-tetracycline/H⁺ antiporter as studied by site-directed chemical labeling. *J. Biol. Chem.* **272**:580–585.
 98. Kimura-Someya, T., S. Iwaki, and A. Yamaguchi. 1998. Site-directed chemical modification of cysteine-scanning mutants as to transmembrane segment II and its flanking regions of the Tn10-encoded metal-tetracycline/H⁺ antiporter reveals a transmembrane water-filled channel. *J. Biol. Chem.* **273**:32806–32811.
 99. Klein, P., M. Kanehisa, and C. DeLisi. 1985. The detection and classification of membrane-spanning proteins. *Biochim. Biophys. Acta* **815**:468–476.
 100. Kuroiwa, T., M. Sakaguchi, K. Mihara, and T. Omura. 1990. Structural requirements for interruption of protein translocation across rough endoplasmic reticulum membrane. *J. Biochem.* **108**:829–834.
 101. Kuroiwa, T., M. Sakaguchi, K. Mihara, and T. Omura. 1991. Systematic analysis of stop-transfer sequence for microsomal membrane. *J. Biol. Chem.* **266**:9251–9255.
 102. Kuroiwa, T., M. Sakaguchi, T. Omura, and K. Mihara. 1996. Reinitiation

- of protein translocation across the endoplasmic reticulum membrane for the topogenesis of multispinning membrane proteins. *J. Biol. Chem.* **271**:6423–6428.
103. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105–132.
 104. Landolt-Marticorena, C., and R. A. Reithmeier. 1994. Asparagine-linked oligosaccharides are localized to single extracytosolic segments in multi-span membrane glycoproteins. *Biochem. J.* **302**:253–260.
 105. Landolt-Marticorena, C., K. A. Williams, C. M. Deber, and R. A. Reithmeier. 1993. Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. *J. Mol. Biol.* **229**:602–608.
 106. LeBlanc, H. N., and J. T. Beatty. 1996. Topological analysis of the *Rhodospirillum rubrum* PucC protein and effects of C-terminal deletions on light-harvesting complex II. *J. Bacteriol.* **178**:4801–4806.
 107. Lee, E., and C. Manoel. 1994. Mutations eliminating the protein export function of a membrane-spanning sequence. *J. Biol. Chem.* **269**:28822–28828.
 108. Lee, M., and C. Manoel. 1996. Molecular genetic analysis of membrane protein topology, p. 189–201. In W. N. Konings, H. R. Kaback, and J. S. Lolkema (ed.), *Handbook of biological physics. Transport processes of eukaryotic and prokaryotic organisms*. Elsevier Science, Amsterdam, The Netherlands.
 109. Lewis, M. J., J. A. Chang, and R. D. Simoni. 1990. A topological analysis of subunit alpha from *Escherichia coli* F1F0-ATP synthase predicts eight transmembrane segments. *J. Biol. Chem.* **265**:10541–10550.
 110. Li, S. J., and J. E. Cronan. 1992. The gene encoding the biotin carboxylase subunit of *Escherichia coli* acetyl-CoA carboxylase. *J. Biol. Chem.* **267**:855–863.
 111. Liao, S., J. Lin, H. Do, and A. E. Johnson. 1997. Both luminal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* **90**:31–41.
 112. Lingappa, V. R., J. R. Lingappa, and G. Blobel. 1979. Chicken ovalbumin contains an internal signal sequence. *Nature* **281**:117–121.
 113. Lolkema, J. S., and D. J. Slotboom. 1998. Estimation of structural similarity of membrane proteins by hydrophobicity profile alignment. *Mol. Membr. Biol.* **15**:33–42.
 114. Long, J. C., S. Wang, and S. B. Vik. 1998. Membrane topology of subunit a of the F1F0 ATP synthase as determined by labeling of unique cysteine residues. *J. Biol. Chem.* **273**:16235–16240.
 115. Loo, T. W., and D. M. Clarke. 1995. Membrane topology of a cysteine-less mutant of human P-glycoprotein. *J. Biol. Chem.* **270**:843–848.
 116. Lu, Y., X. Xiong, A. Helm, K. Kimani, A. Bragin, and W. R. Skach. 1998. Co- and posttranslational translocation mechanisms direct cystic fibrosis transmembrane conductance regulator N terminus transmembrane assembly. *J. Biol. Chem.* **273**:568–576.
 117. Manoel, C. 1990. Analysis of protein localization by use of gene fusions with complementary properties. *J. Bacteriol.* **172**:1035–1042.
 118. Manoel, C., and J. Beckwith. 1986. A genetic approach to analyzing membrane protein topology. *Science* **233**:1403–1408.
 119. Martoglio, B., M. W. Hofmann, J. Brunner, and B. Dobberstein. 1995. The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. *Cell* **81**:207–214.
 120. Matlack, K. E., W. Mothes, and T. A. Rapoport. 1998. Protein translocation: tunnel vision. *Cell* **92**:381–390.
 121. McAllister, H. C., and M. J. Coon. 1966. Further studies on the properties of liver propionyl coenzyme A holocarboxylase synthetase and the specificity of holocarboxylase formation. *J. Biol. Chem.* **241**:2855–2861.
 122. McKenna, E., D. Hardy, and H. R. Kaback. 1992. Insertional mutagenesis of hydrophilic domains in the lactose permease of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **89**:11954–11958.
 123. Melchers, K., T. Weitzenegger, A. Buhmann, W. Steinhilber, G. Sachs, and K. P. Schafer. 1996. Cloning and membrane topology of a P type ATPase from *Helicobacter pylori*. *J. Biol. Chem.* **271**:446–457.
 124. Michaelis, S., H. Inouye, D. Oliver, and J. Beckwith. 1983. Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. *J. Bacteriol.* **154**:366–374.
 125. Miller, J. D., H. Wilhelm, L. Gierasch, R. Gilmore, and P. Walter. 1993. GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. *Nature* **366**:351–354.
 126. Moss, K., A. Helm, Y. Lu, A. Bragin, and W. R. Skach. 1998. Coupled translocation events generate topological heterogeneity at the endoplasmic reticulum membrane. *Mol. Biol. Cell* **9**:2681–2697.
 127. Mothes, W., S. U. Heinrich, R. Graf, I. Nilsson, G. von Heijne, J. Brunner, and T. A. Rapoport. 1997. Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell* **89**:523–533.
 128. Newitt, J. A., and H. D. Bernstein. 1998. A mutation in the *Escherichia coli* secY gene that produces distinct effects on inner membrane protein insertion and protein export. *J. Biol. Chem.* **273**:12451–12456.
 129. Nilsson, I., and G. von Heijne. 1990. Fine-tuning the topology of a polytopic membrane protein: role of positively and negatively charged amino acids. *Cell* **62**:1135–1141.
 130. Nilsson, I. M., and G. von Heijne. 1993. Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* **268**:5798–5801.
 131. Olender, E. H., and R. D. Simon. 1992. The intracellular targeting and membrane topology of 3-hydroxy-3-methylglutaryl-CoA reductase. *J. Biol. Chem.* **267**:4223–4235.
 132. Olivares, L., C. Aragon, C. Gimenez, and F. Zafra. 1997. Analysis of the transmembrane topology of the glycine transporter GLYT1. *J. Biol. Chem.* **272**:1211–1217.
 133. Ota, K., M. Sakaguchi, N. Hamasaki, and K. Mihara. 1998. Assessment of topogenic functions of anticipated transmembrane segments of human band 3. *J. Biol. Chem.* **273**:28286–28291.
 134. Ota, K., M. Sakaguchi, G. von Heijne, N. Hamasaki, and K. Mihara. 1998. Forced transmembrane orientation of hydrophilic polypeptide segments in multispinning membrane proteins. *Mol. Cell* **2**:495–503.
 135. Pan, C. J., K. J. Lei, B. Annabi, W. Hemrika, and J. Y. Chou. 1998. Transmembrane topology of glucose-6-phosphatase. *J. Biol. Chem.* **273**:6144–6148.
 136. Pascual, J. M., C. C. Shieh, G. E. Kirsch, and A. M. Brown. 1995. Multiple residues specify external tetraethylammonium blockade in voltage-gated potassium channels. *Biophys. J.* **69**:428–434.
 137. Pi, J., and A. J. Pittard. 1996. Topology of the phenylalanine-specific permease of *Escherichia coli*. *J. Bacteriol.* **178**:2650–2655.
 138. Pohlshroder, M., W. A. Prinz, E. Hartmann, and J. Beckwith. 1997. Protein translocation in the three domains of life: variations on a theme. *Cell* **91**:563–566.
 139. Popot, J. L., and D. M. Engelman. 1990. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* **29**:4031–4037.
 140. Popov, M., L. Y. Tam, J. Li, and R. A. Reithmeier. 1997. Mapping the ends of transmembrane segments in a polytopic membrane protein. Scanning N-glycosylation mutagenesis of extracytosolic loops in the anion exchanger, band 3. *J. Biol. Chem.* **272**:18325–18332.
 141. Pos, K. M., M. Bott, and P. Dimroth. 1994. Purification of two active fusion proteins of the Na(+)-dependent citrate carrier of *Klebsiella pneumoniae*. *FEBS Lett.* **347**:37–41.
 142. Pouny, Y., C. Weitzman, and H. R. Kaback. 1998. In vitro biotinylation provides quantitative recovery of highly purified active lactose permease in a single step. *Biochemistry* **37**:15713–15719.
 143. Pourcher, T., E. Bibi, H. R. Kaback, and G. Leblanc. 1996. Membrane topology of the melibiose permease of *Escherichia coli* studied by melB-phoA fusion analysis. *Biochemistry* **35**:4161–4168.
 144. Prinz, W. A., and J. Beckwith. 1994. Gene fusion analysis of membrane protein topology: a direct comparison of alkaline phosphatase and beta-lactamase fusions. *J. Bacteriol.* **176**:6410–6413.
 145. Prinz, W. A., D. H. Boyd, M. Ehrmann, and J. Beckwith. 1998. The protein translocation apparatus contributes to determining the topology of an integral membrane protein in *Escherichia coli*. *J. Biol. Chem.* **273**:8419–8424.
 146. Rapiejko, P. J., and R. Gilmore. 1994. Signal sequence recognition and targeting of ribosomes to the endoplasmic reticulum by the signal recognition particle do not require GTP. *Mol. Biol. Cell* **5**:887–897.
 147. Rapoport, T. A., B. Jungnickel, and U. Kutay. 1996. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu. Rev. Biochem.* **65**:271–303.
 148. Rees, D. C., L. DeAntonio, and D. Eisenberg. 1989. Hydrophobic organization of membrane proteins. *Science* **245**:510–513.
 149. Rees, D. C., H. Komiya, T. O. Yeates, J. P. Allen, and G. Feher. 1989. The bacterial photosynthetic reaction center as a model for membrane proteins. *Annu. Rev. Biochem.* **58**:607–633.
 150. Romisch, K., J. Webb, J. Herz, S. Prehn, R. Frank, M. Vingron, and B. Dobberstein. 1989. Homology of 54K protein of signal-recognition particle, docking protein and two *E. coli* proteins with putative GTP-binding domains. *Nature* **340**:478–482.
 151. Rothman, A., E. Padan, and S. Schuldiner. 1996. Topological analysis of NhaA, a Na⁺/H⁺ antiporter from *Escherichia coli*. *J. Biol. Chem.* **271**:32288–32292.
 152. Saaf, A., M. Monne, J. W. De Gier, and G. von Heijne. 1998. Membrane topology of the 60-kDa oxa1p homologue from *Escherichia coli*. *J. Biol. Chem.* **273**:30415–30418.
 153. Saaf, A., E. Wallin, and G. von Heijne. 1998. Stop-transfer function of pseudo-random amino acid segments during translocation across prokaryotic and eukaryotic membranes. *Eur. J. Biochem.* **251**:821–829.
 154. Sahin-Toth, M., R. L. Dunten, and H. R. Kaback. 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.
 155. Samols, D., C. G. Thornton, V. L. Murtif, G. K. Kumar, F. C. Haase, and H. G. Wood. 1988. Evolutionary conservation among biotin enzymes. *J. Biol. Chem.* **263**:6461–6464.
 156. San Millan, J. L., D. Boyd, R. Dalbey, W. Wickner, and J. Beckwith. 1989. Use of phoA fusions to study the topology of the *Escherichia coli* inner membrane protein leader peptidase. *J. Bacteriol.* **171**:5536–5541.
 157. Sarsero, J. P., and A. J. Pittard. 1995. Membrane topology analysis of

- Escherichia coli* K-12 Mtr permease by alkaline phosphatase and beta-galactosidase fusions. *J. Bacteriol.* **177**:297–306.
158. Sato, M., R. Hresko, and M. Mueckler. 1998. Testing the charge difference hypothesis for the assembly of a eucaryotic multispanning membrane protein. *J. Biol. Chem.* **273**:25203–25208.
 159. Schatz, G., and B. Dobberstein. 1996. Common principles of protein translocation across membranes. *Science* **271**:1519–1526.
 160. Seal, R. P., and S. G. Amara. 1998. A reentrant loop domain in the glutamate carrier EAAT1 participates in substrate binding and translocation. *Neuron* **21**:1487–1498.
 161. Seligman, L., and C. Manoil. 1994. An amphipathic sequence determinant of membrane protein topology. *J. Biol. Chem.* **269**:19888–19896.
 162. Seluanov, A., and E. Bibi. 1997. FtsY, the prokaryotic signal recognition particle receptor homologue, is essential for biogenesis of membrane proteins. *J. Biol. Chem.* **272**:2053–2055.
 163. Shenoy, B. C., and H. G. Wood. 1988. Purification and properties of the synthetase catalyzing the biotinylation of the apousubunit of transcarboxylase from *Propionibacterium shermanii*. *FASEB J.* **2**:2396–2401.
 164. Silhavy, T. J., H. A. Shuman, J. Beckwith, and M. Schwartz. 1977. Use of gene fusions to study outer membrane protein localization in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **74**:5411–5415.
 165. Simon, S. M., and G. Blobel. 1991. A protein-conducting channel in the endoplasmic reticulum. *Cell* **65**:371–380.
 166. Sipos, L., and G. von Heijne. 1993. Predicting the topology of eukaryotic membrane proteins. *Eur. J. Biochem.* **213**:1333–1340.
 167. Skach, W. R., L. B. Shi, M. C. Calayag, A. Frigeri, V. R. Lingappa, and A. S. Verkman. 1994. Biogenesis and transmembrane topology of the CHIP28 water channel at the endoplasmic reticulum. *J. Cell Biol.* **125**:803–815.
 168. Skvirsky, R. C., S. Reginald, and X. Shen. 1995. Topology analysis of the colicin V export protein CvaA in *Escherichia coli*. *J. Bacteriol.* **177**:6153–6159.
 169. Slotboom, D. J., J. S. Lolkema, and W. N. Konings. 1996. Membrane topology of the C-terminal half of the neuronal, glial, and bacterial glutamate transporter family. *J. Biol. Chem.* **271**:31317–31321.
 170. Song, H. Y., and W. A. Cramer. 1991. Membrane topography of ColE1 gene products: the immunity protein. *J. Bacteriol.* **173**:2935–2943.
 171. Tam, L. Y., C. Landolt-Marticorena, and R. A. Reithmeier. 1996. Glycosylation of multiple extracytosolic loops in Band 3, a model polytopic membrane protein. *Biochem. J.* **318**:645–648.
 172. Tang, X. B., J. Fujinaga, R. Kopito, and J. R. Casey. 1998. Topology of the region surrounding Glu681 of human AE1 protein, the erythrocyte anion exchanger. *J. Biol. Chem.* **273**:22545–22553.
 173. Tanner, M. J. 1997. The structure and function of band 3 (AE1): recent developments. *Mol. Membr. Biol.* **14**:155–165.
 174. Tate, C. G., and P. J. Henderson. 1993. Membrane topology of the L-rhamnose-H⁺ transport protein (RhaT) from enterobacteria. *J. Biol. Chem.* **268**:26850–26857.
 175. Traxler, B., D. Boyd, and J. Beckwith. 1993. The topological analysis of integral cytoplasmic membrane proteins. *J. Membr. Biol.* **132**:1–11.
 176. Traxler, B., and C. Murphy. 1996. Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery. *J. Biol. Chem.* **271**:12394–12400.
 177. Tsukihara, T., H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. Shinzawa-Itoh, R. Nakashima, R. Yaono, and S. Yoshikawa. 1996. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. *Science* **272**:1136–1144.
 178. Turk, E., C. J. Kerner, M. P. Lostao, and E. M. Wright. 1996. Membrane topology of the human Na⁺/glucose cotransporter SGLT1. *J. Biol. Chem.* **271**:1925–1934.
 179. Uchida, K., H. Mori, and S. Mizushima. 1995. Stepwise movement of preproteins in the process of translocation across the cytoplasmic membrane of *Escherichia coli*. *J. Biol. Chem.* **270**:30862–30868.
 180. Ujwal, M. L., H. Jung, E. Bibi, C. Manoil, C. Altenbach, W. L. Hubbell, and H. R. Kaback. 1995. Membrane topology of helices VII and XI in the lactose permease of *Escherichia coli* studied by lacY-phoA fusion analysis and site-directed spectroscopy. *Biochemistry* **34**:14909–14917.
 181. Ulbrandt, N. D., J. A. Newitt, and H. D. Bernstein. 1997. The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* **88**:187–196.
 182. Valent, Q. A., P. A. Scotti, S. High, J. W. De Gier, G. von Heijne, G. Lentzen, W. Wintermeyer, B. Oudega, and J. Luirink. 1998. The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J.* **17**:2504–2512.
 183. Valiyaveetil, F. I., and R. H. Fillingame. 1998. Transmembrane topography of subunit a in the *Escherichia coli* F1F0 ATP synthase. *J. Biol. Chem.* **273**:16241–16247.
 184. van Geest, M., and J. S. Lolkema. 1996. Membrane topology of the sodium ion-dependent citrate carrier of *Klebsiella pneumoniae*. Evidence for a new structural class of secondary transporters. *J. Biol. Chem.* **271**:25582–25589.
 185. Reference deleted.
 186. van Geest, M., and J. S. Lolkema. 1999. Transmembrane segment VIII of the Na⁺/citrate transporter CitS requires downstream TMS IX for insertion in the *Escherichia coli* membrane. *J. Biol. Chem.* **274**:29705–29711.
 187. van Geest, M., I. Nilsson, G. von Heijne, and J. S. Lolkema. 1999. Insertion of a bacterial secondary transport protein in the endoplasmic reticulum membrane. *J. Biol. Chem.* **274**:2816–2823.
 188. van Iwaarden, P. R., J. C. Pastore, W. N. Konings, and H. R. Kaback. 1991. Construction of a functional lactose permease devoid of cysteine residues. *Biochemistry* **30**:9595–9600.
 189. Vannier, B., X. Zhu, D. Brown, and L. Birnbaumer. 1998. The membrane topology of human transient receptor potential 3 as inferred from glycosylation-scanning mutagenesis and epitope immunocytochemistry. *J. Biol. Chem.* **273**:8675–8679.
 190. von Heijne, G. 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* **341**:456–458.
 191. von Heijne, G. 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225**:487–494.
 192. von Heijne, G. 1994. Membrane proteins: from sequence to structure. *Annu. Rev. Biophys. Biomol. Struct.* **23**:167–192.
 193. von Heijne, G., and Y. Gavel. 1988. Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* **174**:671–678.
 194. Wahle, S., and W. Stoffel. 1996. Membrane topology of the high-affinity L-glutamate transporter (GLAST-1) of the central nervous system. *J. Cell Biol.* **135**:1867–1877.
 195. Walter, P., and A. E. Johnson. 1994. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **10**:87–119.
 196. Weljly, J. K., P. Shenbagamurthi, W. J. Lennarz, and F. Naider. 1983. Substrate recognition by oligosaccharyltransferase. Studies on glycosylation of modified Asn-X-Thr/Ser tripeptides. *J. Biol. Chem.* **258**:11856–11863.
 197. Wessels, H. P., and M. Spiess. 1988. Insertion of a multispanning membrane protein occurs sequentially and requires only one signal sequence. *Cell* **55**:61–70.
 198. Whitley, P., T. Zander, M. Ehrmann, M. Haardt, E. Bremer, and G. von Heijne. 1994. Sec-independent translocation of a 100-residue periplasmic N-terminal tail in the *E. coli* inner membrane protein proW. *EMBO J.* **13**:4653–4661.
 199. Wilkinson, B. M., A. J. Critchley, and C. J. Stirling. 1996. Determination of the transmembrane topology of yeast Sec61p, an essential component of the endoplasmic reticulum translocation complex. *J. Biol. Chem.* **271**:25590–25597.
 200. Wilmes-Riesenberg, M. R., and B. L. Wanner. 1992. *TnphoA* and *TnphoA'* elements for making and switching fusions for study of transcription, translation, and cell surface localization. *J. Bacteriol.* **174**:4558–4575.
 201. Wolfe, P. B., M. Rice, and W. Wickner. 1985. Effects of two sec genes on protein assembly into the plasma membrane of *Escherichia coli*. *J. Biol. Chem.* **260**:1836–1841.
 202. Wu, J., L. S. Tisa, and B. P. Rosen. 1992. Membrane topology of the ArsB protein, the membrane subunit of an anion-translocating ATPase. *J. Biol. Chem.* **267**:12570–12576.
 203. Xu, M., and M. H. Akabas. 1993. Amino acids lining the channel of the gamma-aminobutyric acid type A receptor identified by cysteine substitution. *J. Biol. Chem.* **268**:21505–21508.
 204. Xu, M., D. F. Covey, and M. H. Akabas. 1995. Interaction of picrotoxin with GABAA receptor channel-lining residues probed in cysteine mutants. *Biophys. J.* **69**:1858–1867.
 205. Yan, R. T., and P. C. Maloney. 1995. Residues in the pathway through a membrane transporter. *Proc. Natl. Acad. Sci. USA* **92**:5973–5976.
 206. Young, C. S., and J. T. Beatty. 1998. Topological model of the *Rhodobacter capsulatus* light-harvesting complex I assembly protein LhaA (previously known as ORF1696). *J. Bacteriol.* **180**:4742–4745.
 207. Yu, N., Y. Cao, S. Mager, and H. A. Lester. 1998. Topological localization of cysteine 74 in the GABA transporter, GAT1, and its importance in ion binding and permeation. *FEBS Lett.* **426**:174–178.
 208. Yun, C. H., S. R. Van Doren, A. R. Crofts, and R. B. Gennis. 1991. The use of gene fusions to examine the membrane topology of the L subunit of the photosynthetic reaction center and of the cytochrome b subunit of the bc1 complex from *Rhodobacter sphaeroides*. *J. Biol. Chem.* **266**:10967–10973.
 209. Zen, K. H., T. G. Conslar, and H. R. Kaback. 1995. Insertion of the polytopic membrane protein lactose permease occurs by multiple mechanisms. *Biochemistry* **34**:3430–3437.
 210. Zhang, J. T., M. Chen, E. Han, and C. Wang. 1998. Dissection of de novo membrane insertion activities of internal transmembrane segments of ATP-binding-cassette transporters: toward understanding topological rules for membrane assembly of polytopic membrane proteins. *Mol. Biol. Cell* **9**:853–863.
 211. Zhang, J. T., C. H. Lee, M. Duthie, and V. Ling. 1995. Topological determinants of internal transmembrane segments in P-glycoprotein sequences. *J. Biol. Chem.* **270**:1742–1746.
 212. Zhou, J., R. T. Fazio, and D. F. Blair. 1995. Membrane topology of the MotA protein of *Escherichia coli*. *J. Mol. Biol.* **251**:237–242.