

# Integral membrane protein biosynthesis: why topology is hard to predict

Carolyn M. Ott<sup>1</sup> and Vishwanath R. Lingappa<sup>2,3,\*</sup>

<sup>1</sup>Program in Biological Sciences, University of California, San Francisco, CA 94143-0444, USA

<sup>2</sup>Departments of Physiology and <sup>3</sup>Medicine, University of California, San Francisco, CA 94143-0444, USA

\*Author for correspondence (e-mail: vrl@itsa.ucsf.edu)

*Journal of Cell Science* 115, 2003-2009 (2002) © The Company of Biologists Ltd

## Summary

**Integral membrane protein biogenesis requires the coordination of several events: accurate targeting of the nascent chain to the membrane; recognition, orientation and integration of transmembrane (TM) domains; and proper formation of tertiary and quaternary structure. Initially unanticipated inter- and intra-protein interactions probably mediate each stage of biogenesis for single spanning, polytopic and C-terminally anchored membrane proteins. The importance of these regulated interactions is**

**illustrated by analysis of topology prediction algorithm failures. Misassigned or misoriented TM domains occur because the primary sequence and overall hydrophobicity of a single TM domain are not the only determinants of membrane integration.**

Key words: Translocon, Endoplasmic reticulum, Biogenesis, Signal transduction, Topogenesis

## Introduction

Understanding the details of integral membrane protein biogenesis is important for the study of any process or pathway that involves these proteins, including signaling cascades, vesicle trafficking and intercellular communication. Structural information is commonly used to predict protein function, and an important feature of the tertiary structure of an integral membrane protein is its topology or its distribution relative to the membrane. Very few integral membrane proteins have had their topology determined experimentally, however, and of those proteins examined, several exhibit topological heterogeneity. That is, polypeptides with identical sequences can span the membrane differently. Researchers therefore commonly rely on topology prediction algorithms, which we will discuss after reviewing the details of biosynthesis. Although these algorithms are helpful for providing a first approximation, they are often imprecise and sometimes predict incorrect topology (see below). An appreciation of the complexity of integral membrane protein biosynthesis empowers scientists to think more critically about a variety of problems: when the data does not exactly fit the model, an alternate topological form may be part of the explanation.

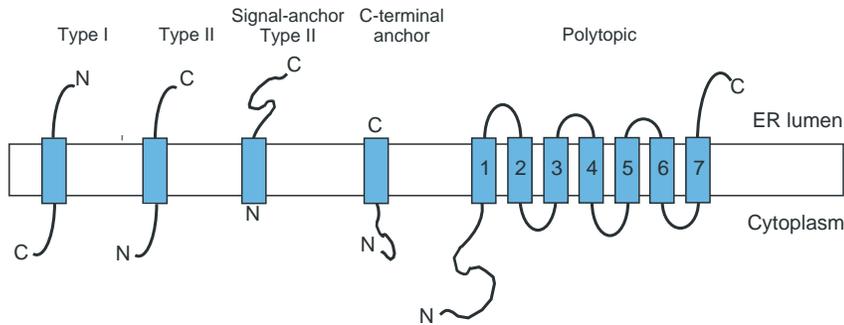
Here we focus on the biosynthesis of mammalian integral membrane proteins that use one or more  $\alpha$ -helical membrane-spanning domains to integrate into the lipid bilayer. Some integral membrane proteins have a single membrane-spanning domain (bitopic), others have several (polytopic). Bitopic membrane proteins are categorized according to the properties of their transmembrane (TM) domains (Fig. 1). During biogenesis, the N-terminus of a type I integral membrane protein is in the ER lumen, whereas in a type II integral membrane protein the N-terminus is in the cytoplasm. Integral membrane proteins that use their first transmembrane domain as both a signal sequence and a stop transfer sequence are classified as signal-anchored proteins. C-terminally anchored

proteins have a signal-anchored domain at the extreme C-terminus.

## Overview of integral membrane protein biogenesis

Biosynthesis of integral membrane proteins involves several interrelated events: targeting of the nascent chain to the ER, translocation of all necessary domains into the ER lumen, recognition and proper orientation of TM domains, integration of TM domains into the lipid bilayer and, in some cases, formation of multimeric complexes. Nucleus-encoded proteins begin translation in the cytosol. Secretory and integral membrane proteins have a signal sequence that is recognized by the signal recognition particle (SRP) shortly after emerging from the ribosome (Walter and Johnson, 1994). Through interactions with its receptor on the surface of the ER, SRP transfers the ribosome-nascent-chain complex to the translocon, an aqueous pore in the ER membrane responsible for translocation and integration (Corsi and Schekman, 1996; Matlack et al., 1998; Fulga et al., 2001). At the ER, upon entering the translocon, integral membrane proteins differ from secretory proteins in that translocation stops and TM domains are oriented and integrated into the bilayer. In vivo the orientation and integration of membrane proteins determines protein topology and is coupled to protein folding (Booth and Curran, 1999; Sanders and Nagy, 2000).

Synthesis of polytopic membrane proteins is more complex than that of bitopic membrane proteins. For example, instead of synthesizing the cytosolic domain of a type I membrane protein and then terminating translocation, the translocation machinery has to be switched on again and begin to translocate another TM domain, another luminal domain, etc. How are these switches controlled? They are regulated by several factors that can act independently or in concert. The hydrophobicity of the TM domain plays an important role.



**Fig. 1.** Types of integral membrane protein. Integral membrane proteins can be synthesized in many different forms. Shown here are a type I integral membrane protein, a type II integral membrane protein, a C-terminally anchored integral membrane protein, a type II signal-anchored protein and a multi-spanning membrane protein.

However, some proteins also have a stop transfer effector (STE) sequence, a domain flanking the hydrophobic membrane-spanning domain, which appears to instruct the translocon not to translocate the domain intended for the cytosol (Yost et al., 1990). In addition, some TM domains facilitate the integration of other TM domains into the same protein.

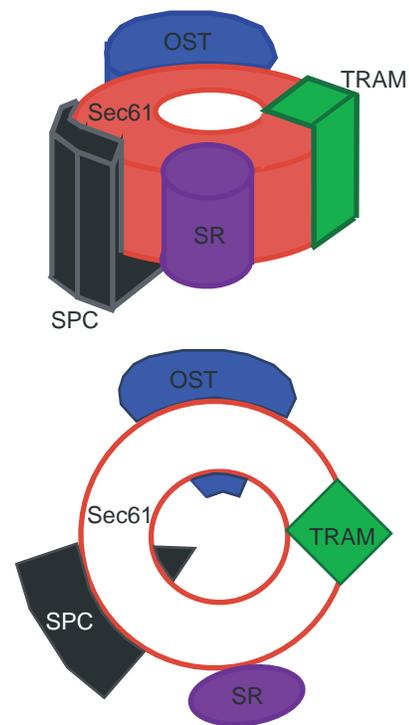
### Co-translational membrane protein biosynthesis

The translocon is a dynamic aqueous pore made up of several different proteins (Fig. 2). Sec61 (an  $\alpha\beta\gamma$  heterotrimer) forms the protein-conducting channel of the translocon (Hanein et al., 1996). The translocating-chain-associated membrane protein (TRAM) is required for translocation of some, but not all, substrates across the ER membrane (Gorlich et al., 1992; Gorlich and Rapoport, 1993). TRAM was first identified through its interaction with the nascent chain early in translocation (Gorlich et al., 1992). Signal sequence structure and the length of the charged N-terminal region determine whether or not a signal sequence requires TRAM for translocation (Voigt et al., 1996). TRAM might also have a regulatory role during protein biogenesis (Hegde et al., 1998c; Hegde and Lingappa, 1999) and has been shown to function in membrane protein integration (see below). Other proteins associated with the translocon include signal peptidase, which cleaves the signal sequence, and oligosaccharyl transferase (OST), which adds N-linked sugars to the nascent chain (Evans et al., 1986; Kelleher et al., 1992). The luminal protein BiP helps maintain the permeability barrier of the membrane early in translocation and during integration (Hamman et al., 1998; Haigh and Johnson, 2002).

Similar to signal sequences, TM domains have differing requirements for TRAM during integration. Attempts to determine exactly how a TM domain passes from the translocon into the lipid bilayer have produced seemingly conflicting results. First it was reported that the TM domain of a type I membrane protein remains associated with translocon components until translation termination (Thrift et al., 1991). The TM domain transits from an environment in which it contacts Sec61 $\alpha$  to an environment in which it contacts TRAM; this suggests lateral movement and lipid integration (Do et al., 1996). More recent studies of a signal-anchored protein led to the alternative model that integration of the TM domain into the lipid bilayer occurs shortly after synthesis and is not dependent on TRAM or translation termination (Mothes

et al., 1997). Changing the properties of the TM domain decreases its ability to partition into the lipid bilayer co-translationally and enables the nascent chain to crosslink to TRAM (Heinrich et al., 2000). It is highly likely that both models are correct and that only some TM domains interact with TRAM during integration, probably those that linger in the translocon.

The translocon must be dynamic. Unlike many other pores, substrates can move through it in two dimensions: into the ER lumen or into the ER membrane. To accommodate the needs of different substrates; it must also be capable of expanding. Fluorescence quenching experiments in the absence of a



**Fig. 2.** The translocon. Two views of the translocon are shown. Sec61 (shown in red) is a heterotrimer and makes up the core of the translocon. TRAM (shown in green) is tightly associated with the translocon and is required for the translocation of many substrates. Other complexes associated with the translocon are signal peptidase (SP; shown in black) and oligosaccharyl transferase (OST; shown in blue).

ribosome indicate that the pore has a diameter of between 9 and 15 Å (Hamman et al., 1997); however, recent electron microscopy data suggest that the pore is closed but dimpled (Beckmann et al., 2001). Sec61 complexes visualized by electron microscopy had a pore size of ~20Å, which is large enough for a single  $\alpha$  helix (Hanein et al., 1996). Other experimental evidence, both direct and indirect, indicates that the channel has a diameter of 40-60Å, which could accommodate up to six TM domains (Borel and Simon, 1996; Hamman et al., 1997).

### A role for intraprotein interactions

A common assumption is that every TM domain is recognized, oriented and integrated independently. This stems from the idea that the simplest way to achieve the correct topology of a polytopic membrane protein is to orient the first TM domain and then alternate between 'start transfer' and 'stop transfer' signals to thread the peptide chain through the membrane. This appears to be one viable mode of membrane protein biogenesis (Rothman et al., 1988); however, some proteins use more complex processes. Stop transfer effectors (STEs) were found in studies of the prion protein (PrP) (Yost et al., 1990) and have also been identified in IgM (Falcone et al., 1999). The characteristics of STEs are not well defined because few have been examined experimentally. In general STEs encompass 10 to 20 residues directly upstream of the TM domain. Mutagenesis studies of the IgM STE found that negatively charged residues are important for stop transfer function (Falcone et al., 1999). The PrP STE, however, contains no negatively charged residues but several positive ones, which may mean that it interacts with different STE receptors. (Receptor-mediated recognition is discussed more in the following section.)

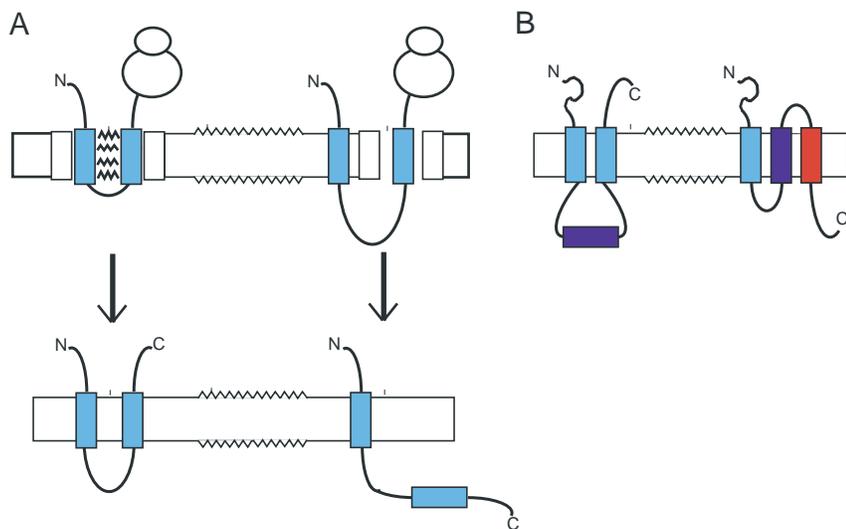
Intraprotein interactions that affect membrane protein biosynthesis can be classified as weak integrators or strong orientation effectors. TM domains that require interaction with adjacent TM domains for proper integration (weak integrators) are present in the multidrug resistance protein MDR1, the *Neurospora* proton transporter H<sup>+</sup>-ATPase and the erythrocyte protein band 3 (Skach and Lingappa, 1993; Lin and Addison,

1995; Ota et al., 2000). In these proteins, specific TM domains can target and properly orient independently, but integration efficiency is poor if the TM domain is unable to interact with adjacent TM domains. Increasing the distance between TM domains reduces the cis interactions and results in translocation of the weak TM domain (Fig. 3a). Orientation of TM domains can also be affected by cis interactions. In the case of the erythrocyte protein band 3, the eighth TM domain (TM8) – a strong orientation effector – is required for both proper orientation and integration of TM7 (Fig. 3b). TM8 is such a strong orientation effector that it can cause the integration of both hydrophobic and hydrophilic domains (Ota et al., 1998b).

### The role of other protein factors

Interactions between TM domains cannot explain how two proteins that have identical primary structures and use the same basic translocation machinery can be synthesized in two different orientations. Several proteins, including the prion protein (PrP), ductin, myelin proteolipid protein (PLP) and the cystic fibrosis transmembrane conductance regulator (CFTR) exist in multiple topological forms (Lopez et al., 1990; Dunlop et al., 1995; Hegde et al., 1998b; Wahle and Stoffel, 1998). Although a nascent chain may access one of the many available folding funnels, studies of PrP have demonstrated that this distribution can be altered both in cis and in trans.

Interprotein interactions can play a role in both TM domain integration and STE recognition. PrP can be synthesized in three different topological forms: <sup>N</sup>tmPrP, a type I membrane protein in which the N-terminus is in the lumen; <sup>C</sup>tmPrP, a type II membrane protein in which the C-terminus is in the lumen; and a secretory form called <sup>sec</sup>PrP. In vitro, in the absence of translocation accessory factor (TrAF) activity, PrP is made exclusively as the <sup>C</sup>tmPrP form (Hegde et al., 1998b), which causes neurodegeneration in mice and humans when synthesized in vivo (Hegde et al., 1998a). Little more is known about TrAF, but perhaps it regulates how or when other factors, such as TRAM, interact with PrP and probably with many other proteins. Early studies suggested that receptor-mediated recognition events occur during translocation starting and stopping (Mize et al., 1986), which is consistent with the



**Fig. 3.** Examples of intra-protein interactions necessary for proper biosynthesis. (A) Weak integrators are TM domains that require association with an adjacent TM domain to integrate into the lipid bilayer. Increasing the length of the loop between the two TM domains (as shown on the right) prevents the necessary interactions (shown as zigzag lines) from occurring, possibly because the first TM domain integrates before the necessary interactions can take place (Ota et al., 2000). (B) A strong orientation effector (shown here as a red region) forces TM orientation and integration of adjacent domains that would not integrate independently (purple region).

subsequent identification of STEs (Yost et al., 1990). Recently, crosslinking studies of an IgM STE sequence identified two membrane proteins involved in STE recognition or function (Falcone et al., 1999). Characterization of these STE receptors will be one of the next steps toward understanding how integration is regulated.

Chaperone activity also appears to have a role in integration. At least one protein factor in the ER membrane is proposed to be responsible for proper biosynthesis of the gap junction component connexin. In vitro synthesis or in vivo overexpression of connexin results in the production of aberrantly cleaved molecules because signal peptidase mistakes the first TM domain for a signal peptide. In vivo cleavage of the TM domain is believed to be prevented by an unidentified chaperone in the membrane, which recognizes the nascent chain and blocks the access of signal peptidase. In vitro this chaperone may be absent or non-functional (Falk and Gilula, 1998).

Co-translocational modification of nascent chains can also affect biosynthesis. Oligosaccharyl transferase (OST) associates with the translocon and glycosylates nascent chains as they emerge in the ER. To look at possible effects of glycosylation on TM domain orientation, Goder et al. (Goder et al., 1999) created a chimeric protein that can be synthesized in either of two topological forms. When they engineered glycosylation sites, they found that reorientation of a transmembrane domain in the translocon was prevented by glycosylation of the luminal TM loop. These results suggest that regulation of glycosylation of native proteins can control folding and orientation of proteins according to the needs of the cell.

The interprotein interactions described above probably affect biosynthesis of many different membrane proteins. Substrate-specific interprotein interactions also affect biosynthesis. In the membrane, as in the cytosol, proteins associate to form functional complexes. Studies of the P-type Na<sup>+</sup>/K<sup>+</sup>-ATPase revealed that the correct insertion of the polytopic  $\alpha$  subunit seventh and eighth TM domains requires association of the bitopic  $\beta$  subunit with the extra-cytosolic loop between the two TM domains (Beguín et al., 1998). When the  $\beta$  subunit encounters the proper region of the  $\alpha$  subunit, it appears to induce a conformational change that promotes proper folding and integration of the TM domains. Specific trans interactions that facilitate proper formation of membrane protein complexes might prevent the nascent chain from making undesirable or deleterious associations with itself or other proteins.

We are beginning to understand more about the proteins that influence membrane protein biosynthesis, but there is much left to learn. Characterization of both TrAFs and the STE receptors will improve our understanding of the mechanism of membrane domain integration, as will additional examples of substrate-specific interactions. Identification of the chaperone involved in connexin biosynthesis will enable us to learn how membrane chaperones function. Finally, discovery of proteins that use glycosylation to control orientation in vivo will clarify other ways in which biosynthesis can be regulated.

### A role for signal sequences in orientation and integration

Signal sequences are vital for targeting proteins to the

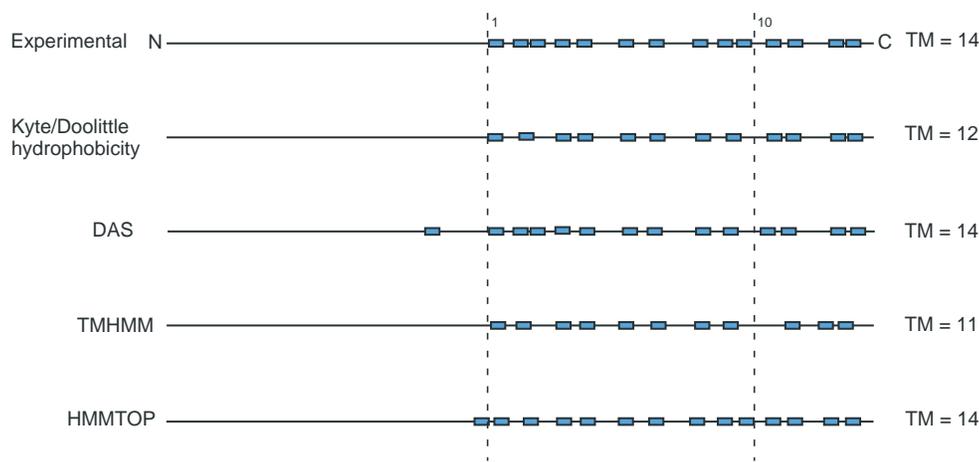
translocon, but they also affect the orientation of subsequent TM domains. Recent research has highlighted an unexpected role for the signal sequence in biosynthesis of secretory proteins. When engineered onto an identical protein, different signal sequences can alter the interactions between the ribosome and the translocon (Rutkowski et al., 2001) or affect glycosylation (D. T. Rutkowski, C.M.O. and V.R.L., unpublished). PrP is one example in which N-terminally cleaved signal sequences affect TM domain orientation and integration. Mutations in the signal sequence, the STE or the TM domain of PrP can dramatically change the fraction of chains synthesized in each of the three topological forms, N<sup>tm</sup>PrP, C<sup>tm</sup>PrP and secPrP (Kim et al., 2001). The mechanism by which the signal sequence influences membrane protein orientation and integration is unclear, but studies of the effect of signal sequences on glycosylation demonstrate that translocon-signal sequence interactions regulate co-translocational modification (D. T. Rutkowski, C.M.O. and V.R.L., unpublished).

### Post-translational targeting and integration

In common with co-translational membrane protein biosynthesis, post-translational targeting and integration is regulated and receptor mediated. C-terminally anchored membrane proteins must target and integrate post-translationally because their signal anchors are not synthesized until translation is almost complete. Initially, the hydrophobicity of the anchor sequence alone was thought to drive both targeting and membrane insertion (Palade, 1975). If this were the case, however, tail-anchored proteins would integrate into any lipid bilayer; proteins destined for the Golgi apparatus or synaptic vesicles could target and integrate directly, bypassing the ER, which is generally not the case. The details of C-terminal integration that are beginning to emerge indicate that post-translational targeting is translocon independent. Studies of synaptobrevin, a vesicle-associated membrane protein (VAMP), revealed that it inserts post-translationally into the ER membrane in an ATP-dependent manner. Integration of synaptobrevin was not affected by membrane depletion of the SRP receptor, Sec61, or other co-translational translocation machinery (Kutay et al., 1995). Studies of another VAMP, Vamp1, found that membrane binding is saturable and involves a trypsin-sensitive factor in the membrane (Kim et al., 1997; Kim et al., 1999). It is unclear whether all the VAMPs use the same targeting and integration pathways, but identification of their receptors will help answer this question.

### Predictive algorithms for integral membrane protein topology

An understanding of the complexity of integral membrane protein biosynthesis allows us to view predictive algorithms in a new light. Researchers commonly rely on algorithms that predict the topology of a protein. These algorithms are available on the Internet and simply require input of the protein sequences. They are especially useful for genome-wide analysis of predicted open reading frames and for identifying relationships between protein families, because they can provide a rough approximation of membrane topology (von Heijne, 1999). Assumptions are often made about the accuracy



**Fig. 4.** Comparison of the experimentally determined and predicted topology of band 3. Band 3 is a polytopic membrane protein that has an N-terminal cytosolic domain. In the diagram, the TM domain is represented as a rectangle and the number of predicted TM domains is indicated for each. The topology of band 3 has been extensively experimentally characterized (Popov et al., 1997; Tanner, 1997; Ota et al., 1998a). Three types of prediction methods are represented: the hydrophathy index (Kyte and Doolittle, 1982); the Dense Alignment Surface (DAS) method (Cserzo et al., 1997); and

two hidden Markov model (HMM) methods, TMHMM (Tusnady and Simon, 1998; Tusnady and Simon, 2001) and HMMTOP (Sonnhammer et al., 1998). For reference, the location of the first and tenth TM domains of the experimentally determined topology are indicated by vertical dotted lines.

and universality of these programs, which can cause problems for users. Integral membrane proteins that can be synthesized in multiple topological forms will elude predictive algorithms. However, predictive algorithms can also incorrectly assign the topology of proteins currently believed to be made in only one topological form. Fig. 4 compares the experimentally determined topology of band 3 with that predicted by four common prediction algorithms. The number, location and boundaries of the TM domains predicted depend on the algorithm used. Below we explain the information prediction algorithms use and their limitations.

Integral membrane proteins have several common features. First, the membrane-spanning domain is generally a hydrophobic  $\alpha$  helix. Interestingly, several residues considered to be helix breakers in aqueous environments, such as glycine, isoleucine and valine, do not disrupt helix formation in the lipid environment of the membrane (Deber et al., 2001). Another trend is the 'positive-inside' rule: the cytoplasmic portion of the integral membrane protein tends to be enriched in positively charged residues (von Heijne, 1992). The problem for topology prediction is that these 'rules' are far from absolute. For example, the positive-inside rule, although largely true in prokaryotes, for which it was formulated, appears to be less true in eukaryotes (Andrews et al., 1992).

Many prediction algorithms have been developed during the past twenty years. The first prediction methods simply evaluated the hydrophobicity of individual residues; regions with several hydrophobic residues were predicted to be TM domains (Kyte and Doolittle, 1982). The dense alignment surface (DAS) method analyzes the frequency with which groups of amino acids are found in the TM domains of proteins in the test set (Cserzo et al., 1997). The latest generation of topology-prediction programs use machine-learning algorithms called hidden Markov models (HMM), which are trained by analyzing the residues that tend to occupy defined regions in the integral membrane proteins. Two such algorithms, transmembrane HMM (TMHMM) and HMMTOP, assess five or seven (respectively) defined regions of an integral membrane protein, such as the helix core, the TM domain

boundaries and cytosolic and luminal domains. Instead of looking at the probability of individual or groups of amino acids to populate each region as in TMHMM, HMMTOP assigns topology by comparing the residues found in one region with those found in other regions (Sonnhammer et al., 1998; Tusnady and Simon, 1998). To evaluate a protein, the programs look for distribution of amino acids in patterns similar to those defined in the training set.

Integral membrane protein topology prediction programs generally attempt to provide four different kinds of information: (1) whether or not the protein is likely to be an integral membrane protein; (2) how many membrane-spanning domains the protein has; (3) the orientations of the transmembrane domains; and (4) the boundaries of the membrane and non-membrane domains. Incorrect predictions can come from several different sources. The hydrophobic core of a soluble protein can be misidentified as a TM domain. Short TM domains or TM domains containing charged residues can be overlooked, as can regions adjacent to strong orientation effector sequences. In Fig. 4 the number of TM domains predicted for band 3 by each program is variable, and even the program that predicts the correct number of TM domains fails to identify the location of the first TM domain correctly. The transmembrane hidden Markov model (TMHMM) predicts an odd number of transmembrane domains and consequent localization of the band 3 C-terminus to the lumen. Prediction errors in the topology assignment of an early TM domain in a multi-spanning membrane protein can result in an incorrectly predicted orientation of the subsequent TM domains.

The training set used by a program can limit its predictive power. Current test sets contain limited information about eukaryotic membrane proteins, because the topologies of relatively few eukaryotic integral membrane proteins have been experimentally determined. Much of the information we do have has come from biochemical analysis. Relatively few crystal structures are available, because membrane proteins are generally hard to crystallize. Bias in the training set comes from both the small sample size available and the fact that certain membrane proteins are more amenable to structural

analysis (Rosenbusch et al., 2001). It is very difficult to determine the exact boundaries of a TM domain by biochemical and structural approaches and so the accuracy of boundaries assigned by prediction programs are difficult to assess (Deber et al., 2001).

Prediction algorithms will continue to develop and take advantage of new technology. Significant improvement, however, will probably require a better understanding of integral membrane protein biosynthesis. As the properties that mediate cis and trans protein interactions are defined, they can be included in the algorithms, perhaps identifying those proteins whose topologies are most difficult to predict.

## Conclusions

Historically, biological events are first defined in the simplest cases; subsequent efforts focus on more complex exceptions. Eventually the exceptions help redefine the rules. This certainly appears true in the case of transmembrane integration in which polytopic membrane proteins have revealed aspects of biogenesis that were not apparent from the studies of simple secretory proteins or bitopic integral membrane proteins.

The new information about cis interactions during biosynthesis should affect how we think about membrane protein folding. The current model of membrane folding involves two stages: (1) folding of independent TM domains; and (2) assembly of those separate domains into a functional protein through lateral helix-helix interactions (Popot and Engelman, 1990; Popot and Engelman, 2000). This model may not fully consider the relationship between folding, orientation, integration and assembly. New data suggest that some TM domains may never exist as independent TM domains. In some cases (such as the P-type Na<sup>+</sup>/K<sup>+</sup>-ATPase described above), multiprotein complex formation is linked to TM domain recognition, orientation and integration. During membrane protein folding, generation of a final folded state is not the result of a linear progression from primary to quaternary structure. Instead, secondary and tertiary structure can be formed simultaneously.

A growing body of evidence that many factors regulate the recognition, orientation and integration of TM domains indicates a level of complexity, and perhaps topological heterogeneity, not apparent from the amino-acid sequence alone. The molecular bases of the cis and trans interactions that affect integral membrane protein biogenesis are not yet well enough understood for us to assess whether they can be incorporated into prediction algorithms. Until such time, although prediction programs have improved significantly, they should still be used cautiously.

Regulation at the levels of transcription, splicing and translation is universally acknowledged, but regulation also occurs during translocation, integration and perhaps even folding. It is highly probable that the decision to make a specific form of a multi-topogenic protein occurs during biogenesis, prior to integration. Cell signaling cascades can regulate integral membrane protein biosynthesis by utilizing chaperones and other accessory factors. Controlling the translocon environment affects cis and trans interactions. Understanding the intricate regulation of integral membrane protein biosynthesis will enable researchers in many fields to understand how these proteins function.

We would like to thank D. T. Rutkowski, S. Saghafi, V. Denick, F. Moore, C. Adler, and P. Walter for useful discussion and critical reading of the manuscript. This work was supported by grants to V.R.L. from the U.S. National Institutes of Health, American Heart Association and the Sandler Foundation.

## References

- Andrews, D. W., Young, J. C., Mirels, L. F. and Czarnota, G. J. (1992). The role of the N region in signal sequence and signal-anchor function. *J. Biol. Chem.* **267**, 7761-7769.
- Beckmann, R., Spahn, C. M. T., Eswar, N., Helters, J., Penczek, P. A., Sali, A., Frank, J. and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* **107**, 361-372.
- Beguín, P., Hasler, U., Beggah, A., Horisberger, J. D. and Geering, K. (1998). Membrane integration of Na,K-ATPase alpha-subunits and beta-subunit assembly. *J. Biol. Chem.* **273**, 24921-24931.
- Booth, P. J. and Curran, A. R. (1999). Membrane protein folding. *Curr. Opin. Struct. Biol.* **9**, 115-121.
- Borel, A. C. and Simon, S. M. (1996). Biogenesis of polytopic membrane proteins – membrane segments assemble within translocation channels prior to membrane integration. *Cell* **85**, 379-389.
- Corsi, A. K. and Schekman, R. (1996). Mechanism of polypeptide translocation into the endoplasmic reticulum. *J. Biol. Chem.* **271**, 30299-30302.
- Cserzo, M., Wallin, E., Simon, I., vonHeijne, G. and Elofsson, A. (1997). Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng.* **10**, 673-676.
- Deber, C. M., Wang, C., Liu, L. P., Prior, A. S., Agrawal, S., Muskat, B. L. and Cuticchia, A. J. (2001). TM Finder: A prediction program for transmembrane protein segments using a combination of hydrophobicity and nonpolar phase helicity scales. *Protein Sci.* **10**, 212-219.
- Do, H., Falcone, D., Lin, J., Andrews, D. W. and Johnson, A. E. (1996). The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* **85**, 369-378.
- Dunlop, J., Jones, P. C. and Finbow, M. E. (1995). Membrane insertion and assembly of Ductin – a polytopic channel with dual orientations. *EMBO J.* **14**, 3609-3616.
- Evans, E. A., Gilmore, R. and Blobel, G. (1986). Purification of microsomal signal peptidase as a complex. *Proc. Natl. Acad. Sci. USA* **83**, 581-585.
- Falcone, D., Do, H., Johnson, A. E. and Andrews, D. W. (1999). Negatively charged residues in the IgM stop-transfer effector sequence regulate transmembrane polypeptide integration. *J. Biol. Chem.* **274**, 33661-33670.
- Falk, M. M. and Gilula, N. B. (1998). Connexin membrane protein biosynthesis is influenced by polypeptide positioning within the translocon and signal peptidase access. *J. Biol. Chem.* **273**, 7856-7864.
- Fulga, T. A., Sinning, I., Dobberstein, B. and Pool, M. R. (2001). SR beta coordinates signal sequence release from SRP with ribosome binding to the translocon. *EMBO J.* **20**, 2338-2347.
- Goder, V., Bieri, C. and Spiess, M. (1999). Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. *J. Cell Biol.* **147**, 257-265.
- Gorlich, D., Hartmann, E., Prehn, S. and Rapoport, T. A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature* **357**, 47-52.
- Gorlich, D. and Rapoport, T. A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* **75**, 615-630.
- Haigh, N. and Johnson, A. (2002). A new role for BiP: closing the aqueous translocon pore during protein integration into the ER membrane. *J. Cell Biol.* **156**, 261-270.
- Hamman, B. D., Chen, J. C., Johnson, E. E. and Johnson, A. E. (1997). The aqueous pore through the translocon has a diameter of 40-60 Å during cotranslational protein translocation at the ER membrane. *Cell* **89**, 535-544.
- Hamman, B. D., Hendershot, L. M. and Johnson, A. E. (1998). BiP maintains the permeability barrier of the ER membrane by sealing the luminal end of the translocon pore before and early in translocation. *Cell* **92**, 747-758.
- Hanein, D., Matlack, K. E., Jungnickel, B., Plath, K., Kalies, K. U., Miller, K. R., Rapoport, T. A. and Akey, C. W. (1996). Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* **87**, 721-732.

- Hegde, R. S. and Lingappa, V. R.** (1999). Regulation of protein biogenesis at the endoplasmic reticulum membrane. *Trends Cell Biol.* **9**, 132-137.
- Hegde, R. S., Mastroianni, J. A., Scott, M. R., DeFea, K. A., Tremblay, P., Torchia, M., DeArmond, S. J., Prusiner, S. B. and Lingappa, V. R.** (1998a). A transmembrane form of the prion protein in neurodegenerative disease. *Science* **279**, 827-834.
- Hegde, R. S., Voigt, S. and Lingappa, V. R.** (1998b). Regulation of protein topology by trans-acting factors at the endoplasmic reticulum. *Mol. Cell* **2**, 85-91.
- Hegde, R. S., Voigt, S., Rapoport, T. A. and Lingappa, V. R.** (1998c). TRAM regulates the exposure of nascent secretory proteins to the cytosol during translocation into the endoplasmic reticulum. *Cell* **92**, 621-631.
- Heinrich, S. U., Mothes, W., Brunner, J. and Rapoport, T. A.** (2000). The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* **102**, 233-244.
- Kelleher, D. J., Kreibich, G. and Gilmore, R.** (1992). Oligosaccharyltransferase activity is associated with a protein complex composed of ribophorins I and II and a 48 kd protein. *Cell* **69**, 55-65.
- Kim, P. K., JaniakSpens, F., Trimble, W. S., Leber, B. and Andrews, D. W.** (1997). Evidence for multiple mechanisms for membrane binding and integration via carboxyl-terminal insertion sequences. *Biochemistry* **36**, 8873-8882.
- Kim, P. K., Hollerbach, C., Trimble, W. S., Leber, B. and Andrews, D. W.** (1999). Identification of the endoplasmic reticulum targeting signal in vesicle-associated membrane proteins. *J. Biol. Chem.* **274**, 36876-36882.
- Kim, S. J., Rahbar, R. and Hegde, R. S.** (2001). Combinatorial control of prion protein biogenesis by the signal sequence and transmembrane domain. *J. Biol. Chem.* **276**, 26132-26140.
- Kutay, U., Ahnert-Hilger, G., Hartmann, E., Wiedenmann, B. and Rapoport, T. A.** (1995). Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. *EMBO J.* **14**, 217-223.
- Kyte, J. and Doolittle, R. F.** (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**, 105-132.
- Lin, J. L. and Addison, R.** (1995). A novel integration signal that is composed of two transmembrane segments is required to integrate the *Neurospora* plasma membrane H<sup>+</sup>-ATPase into microsomes. *J. Biol. Chem.* **270**, 6935-6941.
- Lopez, C. D., Yost, C. S., Prusiner, S. B., Myers, R. M. and Lingappa, V. R.** (1990). Unusual topogenic sequence directs prion protein biogenesis. *Science* **248**, 226-229.
- Matlack, K. E., Mothes, W. and Rapoport, T. A.** (1998). Protein translocation: tunnel vision. *Cell* **92**, 381-390.
- Mize, N. K., Andrews, D. W. and Lingappa, V. R.** (1986). A stop transfer sequence recognizes receptors for nascent chain translocation across the endoplasmic reticulum membrane. *Cell* **47**, 711-719.
- Mothes, W., Heinrich, S. U., Graf, R., Nilsson, I., von Heijne, G., Brunner, J. and Rapoport, T. A.** (1997). Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell* **89**, 523-533.
- Ota, K., Sakaguchi, M., Hamasaki, N. and Mihara, K.** (1998a). Assessment of topogenic functions of anticipated transmembrane segments of human band 3. *J. Biol. Chem.* **273**, 28286-28291.
- Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N. and Mihara, K.** (1998b). Forced transmembrane orientation of hydrophilic polypeptide segments in multispanning membrane proteins. *Mol. Cell* **2**, 495-503.
- Ota, K., Sakaguchi, M., Hamasaki, N. and Mihara, K.** (2000). Membrane integration of the second transmembrane segment of band 3 requires a closely apposed preceding signal-anchor sequence. *J. Biol. Chem.* **275**, 29743-29748.
- Palade, G.** (1975). Intracellular aspects of the process of protein synthesis. *Science* **189**, 347-358.
- Popot, J. L. and Engelman, D. M.** (1990). Membrane protein folding and oligomerization – the 2-stage model. *Biochemistry* **29**, 4031-4037.
- Popot, J. L. and Engelman, D. M.** (2000). Helical membrane protein folding, stability, and evolution. *Annu. Review Biochem.* **69**, 881-922.
- Popov, M., Tam, L. Y., Li, J. and Reithmeier, R. A.** (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.
- Rosenbusch, J. P., Lustig, A., Grabo, M., Zulauf, M. and Regenass, M.** (2001). Approaches to determining membrane protein structures to high resolution: do selections of subpopulations occur? *Micron* **32**, 75-90.
- Rothman, R. E., Andrews, D. W., Calayag, M. C. and Lingappa, V. R.** (1988). Construction of defined polytopic integral transmembrane proteins. The role of signal and stop transfer sequence permutations. *J. Biol. Chem.* **263**, 10470-10480.
- Rutkowski, D. T., Lingappa, V. R. and Hegde, R. S.** (2001). Substrate-specific regulation of the ribosome-translocon junction by N-terminal signal sequences. *Proc. Natl Acad. Sci. USA* **98**, 7823-7828.
- Sanders, C. R. and Nagy, J. K.** (2000). Misfolding of membrane proteins in health and disease: the lady or the tiger? *Curr. Opin. Struct. Biol.* **10**, 438-442.
- Skach, W. R. and Lingappa, V. R.** (1993). Amino-terminal assembly of human P-glycoprotein at the endoplasmic reticulum is directed by cooperative actions of two internal sequences. *J. Biol. Chem.* **268**, 23552-23561.
- Sonnhammer, E. L. L., von Heijne, G. and Krogh, A.** (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 175-182.
- Tanner, M. J.** (1997). The structure and function of band 3 (AE1): recent developments. *Mol. Membr. Biol.* **14**, 155-165.
- Thriff, R. N., Andrews, D. W., Walter, P. and Johnson, A. E.** (1991). A nascent membrane protein is located adjacent to ER membrane proteins throughout its integration and translation. *J. Cell Biol.* **112**, 809-821.
- Tusnady, G. E. and Simon, I.** (1998). Principles governing amino acid composition of integral membrane proteins: Application to topology prediction. *J. Mol. Biol.* **283**, 489-506.
- Tusnady, G. E. and Simon, I.** (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849-850.
- Voigt, S., Jungnickel, B., Hartmann, E. and Rapoport, T. A.** (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J. Cell Biol.* **134**, 25-35.
- von Heijne, G.** (1992). Membrane protein structure prediction – Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225**, 487-494.
- von Heijne, G.** (1999). Recent advances in the understanding of membrane protein assembly and structure. *Q. Rev. Biophys.* **32**, 285-307.
- Wahle, S. and Stoffel, W.** (1998). Cotranslational integration of myelin proteolipid protein (PLP) into the membrane of endoplasmic reticulum: Analysis of topology by glycosylation scanning and protease domain protection assay. *Glia* **24**, 226-235.
- Walter, P. and Johnson, A. E.** (1994). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **10**, 87-119.
- Yost, C. S., Lopez, C. D., Prusiner, S. B., Myers, R. M. and Lingappa, V. R.** (1990). Non-hydrophobic extracytoplasmic determinant of stop transfer in the prion protein. *Nature* **343**, 669-672.