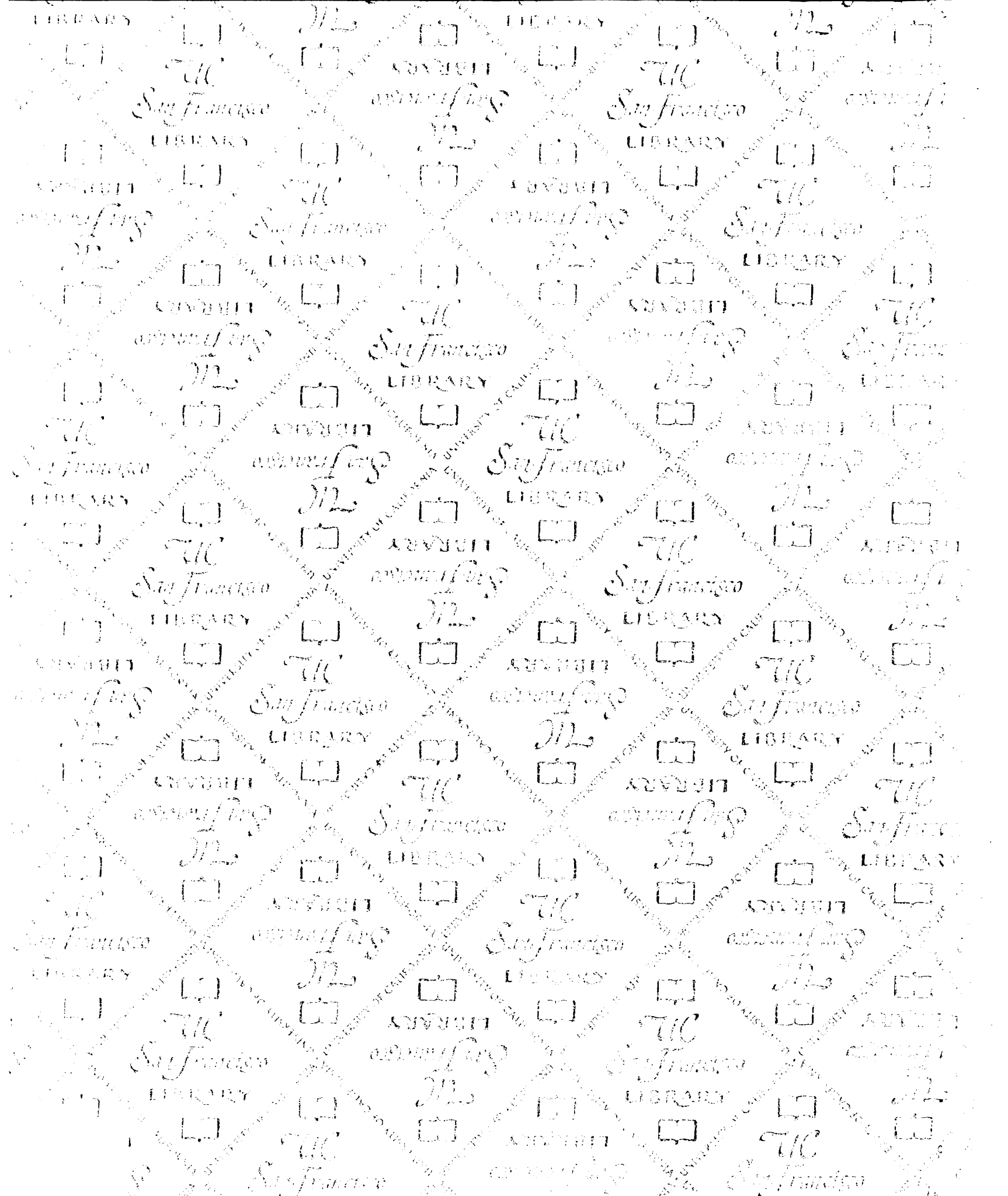
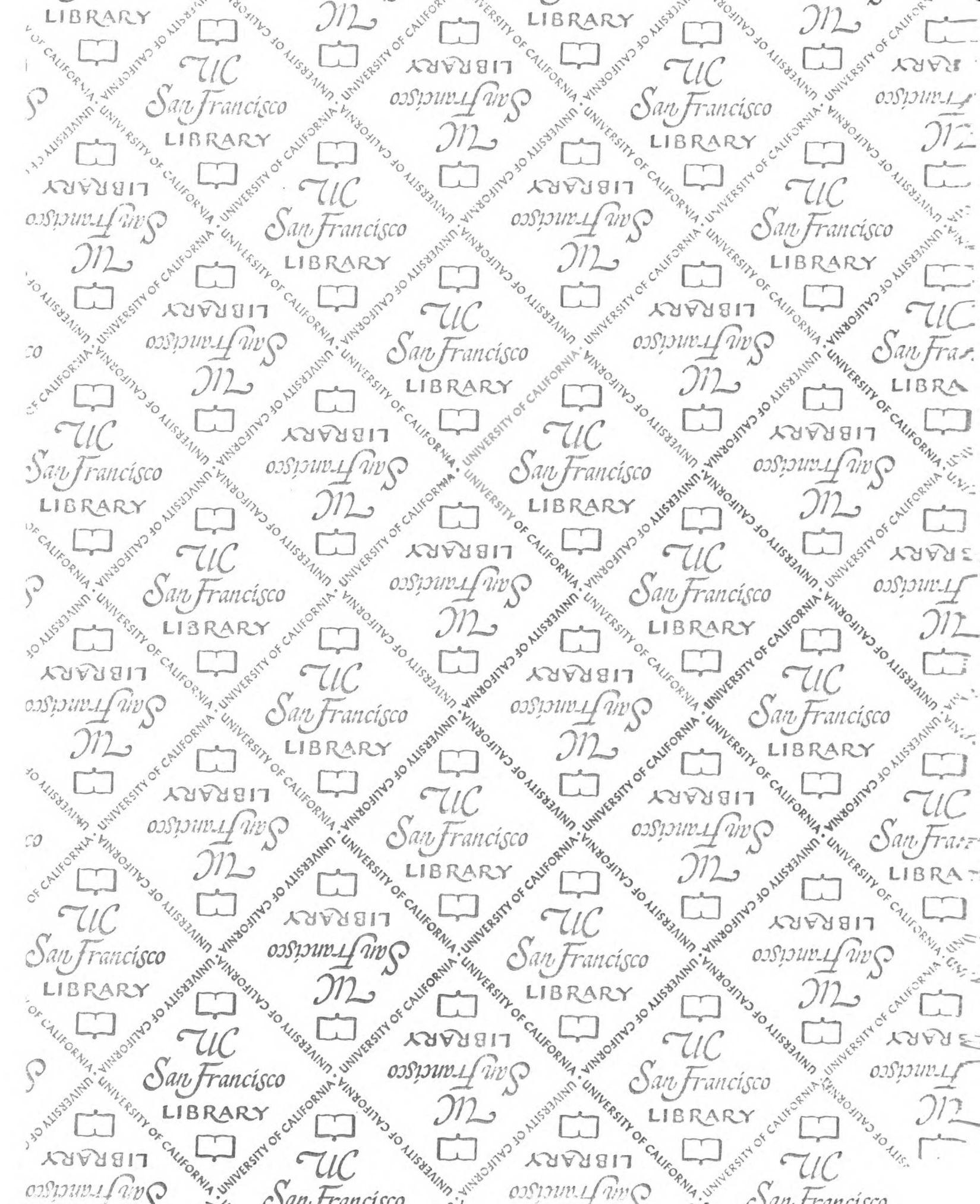
Title
Regulation of membrane protein integration at the translocon

Author
Ott, Carolyn M

Publication Date
2004

Peer reviewed|Thesis/dissertation


The San Francisco LIBRARY
SOCIETY OF CALIFORNIA


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA ·

Regulation of membrane protein integration at the translocon

by

Carolyn M. Ott

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgments

I could not have reached the end of this stage of my journey without the help of many people along the way. I would like to thank, first and foremost, my family. I especially want to thank my husband Jason who picked up everything and came with me to San Francisco (without ever having visited) so I could go to graduate school. He has put up with me almost every day of my graduate career. You are truly a blessing to me and I thank you. My parents have *always* supported me in every endeavor and this was no exception. They didn't always understand what I was doing or what I was going through, but when I was happy, they were happy for me, and when I was frustrated or discouraged, they supported me with phone calls and emails. My sisters are two of my closest friends and they, too, have encouraged me in little ways, like telling me how proud they are of me.

My time at UCSF has been shaped by my advisor, Vishu, and the members of his laboratory. Vishu always provided an environment of intellectual freedom. For much of my career he was as available as I needed him to be, working closely with me when I wanted his guidance and giving me space when I wanted to plot my own course. From Vishu I learned how to think big; to consider the implications of my research for cell biology in general. He is also a fantastic storyteller.

Manu and Tom, both during their time in the laboratory and since their departure, have been my primary sounding boards. Manu generously provided materials, protocols, and most importantly ideas and knowledge that helped me shape my experiments. I have also learned about the practice and politics of science as he has launched his own laboratory. Tom, for better or for worse, taught me to be skeptical, both of my data and

that of others. He also taught me how to sort through Vishu's creative ideas to find those that were both feasible and significant. In addition to his scientific influence, Tom has been a friend.

I have especially enjoyed working with two technicians in the laboratory, May and Fred. They must receive credit for generating most of the 400 plus clones I used during my studies. They also provided invaluable help, such as culturing cells, running gels, and performing countless experiments with me. I also want to thank Jason and Michael pouring gels and preparing bacteria culture media, and Benny, Nona, Susan, and Angela for cleaning glassware and preparing other materials.

It never occurred to me to become a scientist until I arrived at the University of Nebraska, Lincoln. There, Dr. Herman Knoche encouraged me to apply for a Howard Hughes program, which funded undergraduate research. I know of few other universities where freshmen are given the opportunity to do independent research. At UNL, much of the credit for my scientific formation goes to Dr. Spreitzer, who introduced me to real scientific research. After first year chemistry laboratory classes, I don't think I would have ever wanted to be a scientist had it not been for the experiences I had in his laboratory. Hong, then a graduate student in Dr. Spreitzer's laboratory was my first and best teacher in the laboratory. I watched and tried to support him as he finished graduate school, but I didn't understand then what he was going through...now I do.

I am grateful to the Howard Hughes Medical Institute, which funded most of my graduate career, along with a significant amount of my undergraduate research. The National Science Foundation funded my first year of graduate school.

This section would not be complete if I did not recognize the source of all the blessings in my life, the Lord. He has provided me with the interest and ability to do the work described here, as well as a supportive family and community in which to grow and learn. Several years ago I scribbled some words on a piece of paper and posted them on my desk, "To strive to be the very best is to seek glory for myself. To strive to do my very best is to glorify the Lord." I know this small body of research reveals a miniscule part of the wonder of Your creation, but I thank You for the opportunity You have given me and I hope that this work contributes to the larger body of scientific knowledge for Your glory and the improvement of the human condition.

Much of the material presented in Chapter I is reprinted by permission of Ott and Lingappa (2002). Journal of Cell Science, vol. 115, pp. 2003-2009, © 2002 The Company of Biologists Ltd. The reprinted work presented herein is substantially that of the author of this thesis.

Regulation of membrane protein integration at the translocon

Carolyn M. Ott

Understanding the details of integral membrane protein biogenesis is important for the study of processes that involve these proteins, such as signaling cascades, vesicle trafficking, and intercellular communication. Biosynthesis of integral membrane proteins involves several interrelated events including recognition, orientation and integration of transmembrane (TM) domains. Any alteration in this process can affect the final folded state of the protein and thus, the protein's function. The goal of the research presented here is to gain a better understanding of the features of the nascent chain that influence transmembrane domain orientation and integration. We used the prion protein (PrP) as a model protein for studying integration because it can be made as an integral membrane protein (in either of two orientations) or as a secretory protein.

The N terminal region of the PrP transmembrane domain and the stop transfer effector (STE; the region immediately preceding the transmembrane domain) are known to affect integration. Here we present the first evidence that signal sequences can also promote integration. Interestingly, increasing the amount of time nascent chains with signal sequences that are poor integration effectors associate with the translocon increases the ability of the transmembrane domain to integrate. These results have important implications for discriminating between two current models of membrane protein integration.

Little is known about how STEs function, or how TM domain orientation is determined. Upon substituting the STE and TM domains from other proteins into PrP we found that STEs from different proteins are not functionally interchangeable and do not affect orientation of TM domains. We learned that the hydrophobic region of the TM domain mediates the requirement for an STE. We also found that the C terminal region of the TM domain influences topology.

The Met/Val polymorphism at position 129 in the C terminal region of the TM domain has been associated with familial, sporadic and infectious prion diseases. We compared ^{Ctm}PrP synthesis from Met129 and Val129 PrP constructs and found that nascent chains with a Met at 129 make more ^{Ctm}PrP. This is significant because ^{Ctm}PrP is thought to cause the neurodegeneration associated with prion diseases.

5/25/04

Vishwanath R. Lingappa
Thesis Committee Chairman

Table of Contents

List of Figures

Chapter I

An Introduction to Translocation and Membrane Protein Integration

Introduction

Membrane proteins, proteins that pass through or are attached to the lipid bilayer, serve many functions in eukaryotic cells. Some examples of the varied functions of membrane proteins in the plasma membrane include binding of ligands to initiate signal transduction cascades, formation of channels for uptake or release of nutrients, and mediating cell-cell interactions. Membrane proteins in intracellular organelles are important for tasks such as turning on the unfolded protein response, and targeting vesicles to and from organelles and the plasma membrane. Despite their varied functions and cellular destinations most integral membrane proteins are synthesized and integrated into the lipid bilayer at the endoplasmic reticulum (ER). Understanding how integral membrane proteins are made can have important implications for studying how they function.

Here we focus on the biosynthesis of mammalian integral membrane proteins that use one or more α-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 or multi-spanning). Bitopic membrane proteins are categorized according to the properties of their transmembrane (TM) domains (see figure 1). During biogenesis, the N-terminus of a type I, or Ntm integral membrane protein is in the ER lumen, whereas in a type II, or Ctm 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-terminal anchored proteins have a signal-anchored at the extreme C-terminus.

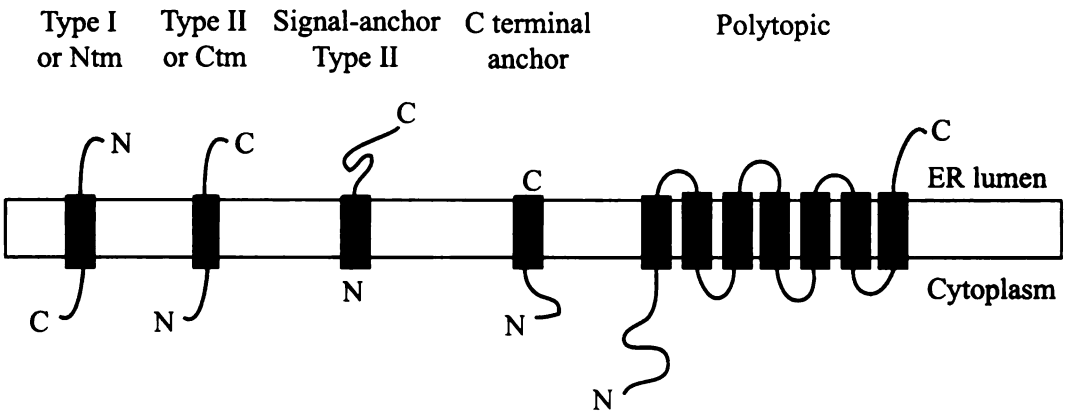


Figure 1: Types of integral membrane proteins

Integral membrane proteins can be synthesized in many different forms. Shown here are a type I integral membrane protein with the N terminus in the lumen, a type II integral membrane protein with the C terminus in the lumen, a type II signal anchored protein, a C terminal anchored integral membrane protein and a multi-spanning membrane protein.

In order to incorporate into the lipid bilayer, TM domains are recognized and integrated co-translationally. This happens at the entrance to secretory pathway, a specialized channel in the membrane of the ER called the translocon. (One exception is tail-anchored proteins whose transmembrane domain is not released from the ribosome until translation has terminated. Biosynthesis of these proteins will not be discussed here.) Both secretory and membrane proteins are targeted to the translocon shortly after the nascent chain has emerged from the ribosome. The translocon facilitates the transfer of proteins both vectorially into the ER lumen and laterally into the ER membrane. Regulated changes in the biosynthesis of integral membrane proteins can alter the final folded state of the protein and thus its function.

This chapter summarizes our current understanding of the molecular details involved in integral membrane protein biosynthesis. Specifically, it reviews how nascent chains target to the translocon, the role of translocon proteins in regulating translocation and integration, the properties of the nascent chain that affect integration, and the role of other intra- and inter- protein interactions in regulating biosynthesis. In addition, the use of topology prediction programs, programs that predict the orientation of the TM domain relative to the membrane, is addressed in light of our current understanding of the complexity of integral membrane protein biosynthesis.

Protein targeting and translocation initiation

Nuclear-encoded proteins begin translation in the cytosol. The first amino acids to emerge from the ribosome upon translation of secretory and integral membrane proteins are the signal sequence. Signal sequences vary in length and composition but

are typically 15 to 50 amino acids in length and have a central hydrophobic domain and hydrophilic N and C-terminal domains (Martoglio and Dobberstein 1998). In some membrane proteins, called signal anchor proteins, the same sequence serves as both the signal sequence and a TM domain. The signal recognition particle (SRP) binds to the hydrophobic domain of the signal sequence and translation arrests. The SRP-ribosome-nascent chain complex targets to the ER membrane, where SRP binds to the SRP receptor (SR) (Gilmore et al. 1982a; Gilmore et al. 1982b; Meyer and Dobberstein 1980a; Meyer and Dobberstein 1980b). Upon hydrolysis of GTP, SRP transfers the ribosome-nascent chain complex to the translocon (Fulga et al. 2001). SRP then acts as a GTPase activating protein to stimulate SR hydrolysis of GTP. This releases SRP from the ER membrane and makes it available for targeting subsequent rounds of targeting (for review see(Keenan et al. 2001)).

The translocon is a dynamic aqueous pore made up of several different proteins. Sec61 (an $\alpha\beta\gamma$ heterotrimer) forms the core of the protein-conducting channel of the translocon (Gorlich and Rapoport 1993). In eukaryotes, the channel is thought to be composed of multiple copies of the Sec61 complex (Hanein et al. 1996; Snapp et al. 2003). Interestingly, recent x-ray crystal structure data of SecYEβ, the archea homolog of Sec 61 indicated that a single heterotrimer can serve as a functional channel (Van den Berg et al. 2004). 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). TRAM is thought to have a regulatory role during protein biogenesis (Hegde and Lingappa 1999;

Hegde et al. 1998c) and has been shown to function in membrane protein integration (see below). The translocon associated protein (TRAP) complex is necessary for translocation initiation of select signal sequences (Fons et al. 2003). 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 lumenal protein BiP helps maintain the permeability barrier of the membrane early in translocation and during integration (Haigh and Johnson 2002; Hamman et al. 1998).

After the ribosome-nascent chain complex is transferred to the translocon, translation resumes. The signal sequence is then recognized by Sec61 complex (Jungnickel and Rapoport 1995; Mothes et al. 1998). Most, but not all, proteins also require the presence of TRAM for proper translocation. 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). In contrast, the requirement of a signal sequence for the TRAP complex is not dependant on hydrophobicity, but on the post-targeting translocation initiation efficiency of the signal sequence (Fons et al. 2003). The signal sequence is reoriented with the N terminus on the cytosolic side of the translocon shortly after translocation begins to generate secretory proteins and transmembrane proteins which have the C terminus localized in the ER lumen (Goder and Spiess 2003).

The translocon must be very 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 ribosome indicate the pore has a diameter between 9 and 15 Å (Hamman et al. 1997); however, recent electron microscopy data suggests 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 a 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).

Integration of membrane proteins

At the ER, upon entering the translocon, integral membrane proteins differ from secretory proteins in that at some point 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 switch on again and begin to translocate another TM domain, another lumenal domain, etc. How are these switches controlled? They are the result of several factors that can act independently or in concert. The hydrophobicity of the TM domain plays an important role. 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 to not translocate the domain intended for the cytosol (Lopez et al. 1990).

In addition, some TM domains facilitate integration of other TM domains in the same protein.

Similar to signal sequences, TM domains appear to 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. The simplest model of integration is that upon entering the translocon, a hydrophobic TM domain very quickly has access to lipids, and it "slips" into the lipid bilayer (Singer et al. 1987). The demonstration that the hydrophobic region of the TM domain of the type II signal anchor protein invariant chain crosslinks to both lipids and proteins supports this idea (Martoglio et al. 1995). Both single and double spanning constructs derived from the bacterial membrane protein leader peptidase crosslink early to the translocon channel protein Sec61 alpha, but as the chain gets longer and the TM domain crosslinks more strongly to lipids, the crosslink to Sec61 alpha is lost (Mothes et al. 1997). The TM domain of leader peptidase was found to partition into lipid-detergent micelles at the same chain length as in intact membranes, but without the need of the Sec61 complex, suggesting that Sec61 just functions to remove the barrier presented by the charged phospholipids head groups (Heinrich et al. 2000). These results also lead to the conclusion that the equilibrium constant for passive partitioning of the TM domain into the lipid bilayer is influenced only by the hydrophobicity of the TM domain.

The alternate view that integration of the TM domain is protein mediated is based on experiments that use crosslinking to demonstrate that the nascent TM domain stays in close proximity to translocon components until termination. Before the components of the translocon were identified it was demonstrated that lysine residues just outside the

TM domain of chimeric type II (N-terminus in the cytosol) signal anchor and type I (with the N-terminus in the lumen) membrane proteins crosslink to ER proteins until the chain is terminated (Thrift et al. 1991). In addition, evidence was found for up to four TM domains to co-reside in the translocon (Borel and Simon 1996). Later experiments showed that the hydrophobic region of the TM domain of progressively longer type I chains crosslink to TRAM in addition to Sec61 alpha (Do et al. 1996). From these results the authors conclude that the TM domain moves from the pore and association with Sec61 stepwise towards the outside of the translocon where it associates with TRAM until release of the nascent chain from the ribosome. The idea of an active role for proteins in integration was further supported by the work of (McCormick et al. 2003) who found position specific crosslinks to Sec61 alpha and TRAM from consecutive residues within the TM domain. As the nascent chains grew longer the distribution of the crosslinks to each face of the helix changed slightly. Although the results were a varied slightly for different TM domains, all TM domains tested still crosslinked with translocon proteins until late in translocation, many until termination. From these results the authors concluded that the TM domain remains bound to the translocon because the nascent chain is not free to rotate. These results indicate that the translocon binds to the TM domains, but it does not directly demonstrate that integration is protein mediated. 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.

Regulation of protein integration in *cis*

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 ten to twenty 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.

Intra-protein 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) have been found in the multidrug resistance protein (MDR1), the *Neurospora* proton transporter, H+-ATPase, and the erythrocyte protein band 3 (Lin and Addison 1995; Ota et al. 2000; Skach and Lingappa 1993). 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 (see figure 2a). 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 (see figure 2b). TM8 is such a strong orientation effector that it can cause the integration of hydrophobic or hydrophilic domains (Ota et al. 1998).

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 (Rutkowski et al. 2003). PrP is one example in which N-terminal 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 ^{Ntm}PrP, ^{Ctm}PrP, and ^{sec}PrP (Kim et al. 2001). In addition, the ability of signal sequences to open the lumenal gate of the translocon can affect the localization of the N terminus, and thus the topologic distribution of nascent chains (Kim et al. 2002).

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

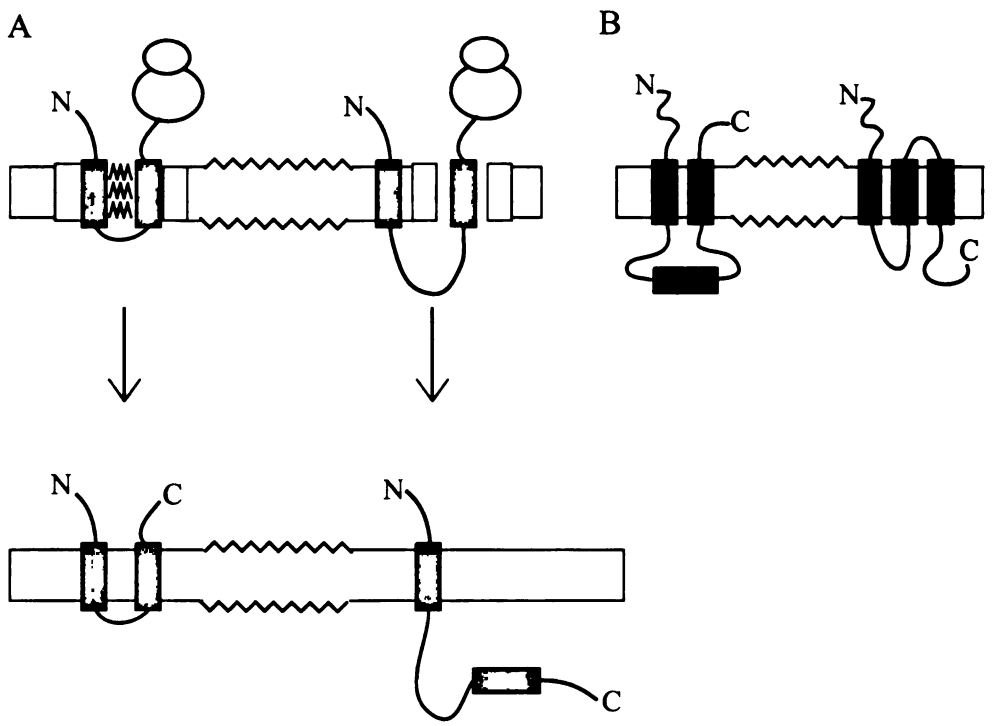


Figure 2: Examples of intraprotein interactions necessary for proper biosynthesis

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).

relationship between folding, orientation, integration and assembly. New data suggest that some TM domains may never exist independently. In some cases (such as the P-type Na^+/K^+ -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.

Regulation of protein integration in *trans*

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 (Dunlop et al. 1995; Hegde et al. 1998c; Lopez et al. 1990; Wahle and Stoffel 1998). Although a nascent chain may access one of multiple available folding funnels, studies of PrP have demonstrated that this distribution can be altered both in *cis* and in *trans*.

Inter-protein interactions can play a role in both TM domain integration and STE recognition. The prion protein (PrP) can be synthesized in three different topological forms: ^{Ntm}PrP, a type I membrane protein in which the N-terminus is in the lumen; ^{Ctm}PrP, a type II membrane protein in which the C-terminus is in the lumen; and a secretory form called ^{sec}PrP. *In vitro*, in the absence of the TRAP complex, PrP is made almost completely as ^{Ctm}PrP (Fons et al. 2003; Hegde et al. 1998c), which causes

neurodegeneration in mice and humans when synthesized *in vivo* (Hegde et al. 1998b). Early studies suggested that receptor-mediated recognition events occur during translocation starting and stopping (Mize et al. 1986), which is consistent with the 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 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 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 affects of glycosylation on TM domain orientation, (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 lumenal TM loop. These results suggests that regulation of

glycosylation of native proteins can control folding and orientation of proteins according to the needs of the cell.

The inter-protein interactions described above probably affect biosynthesis of many different membrane proteins. Substrate-specific inter-protein interactions also affect biosynthesis. In the membrane, as in the cytosol, proteins associate to form functional complexes. Studies of the P-type Na^+/K^+-ATPase revealed that the correct insertion of the polytopic α subunit seventh and eighth TM domains requires association of the bitopic β subunit with the extra-cytosolic loop between the two TM domains (Beguin et al. 1998). When the β subunit encounters the proper region of the α 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 learn more about proteins that influence membrane protein biosynthesis, but there is much left to learn. Characterization of 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.

Predictive algorithms - why topology is hard to predict

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; therefore researchers commonly rely on topology prediction algorithms. However, of those proteins examined, several exhibit topological heterogeneity. That is, polypeptides with identical sequences can span the membrane differently. Although these algorithms are helpful for providing a first approximation, they are often imprecise and sometimes predict incorrect topologies (see below). An understanding of the complexity of integral membrane protein biosynthesis allows us to view predictive algorithms in a new light. These algorithms are available on the Internet and simply require input of 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 and universality of these programs, which can cause problems for users. Integral membrane proteins that can be synthesized in multiple topologies 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. Figure 3 compares the experimentally determined topology of band 3 to that predicted by four common prediction algorithms. The number, location and boundaries of the TM domains predicted

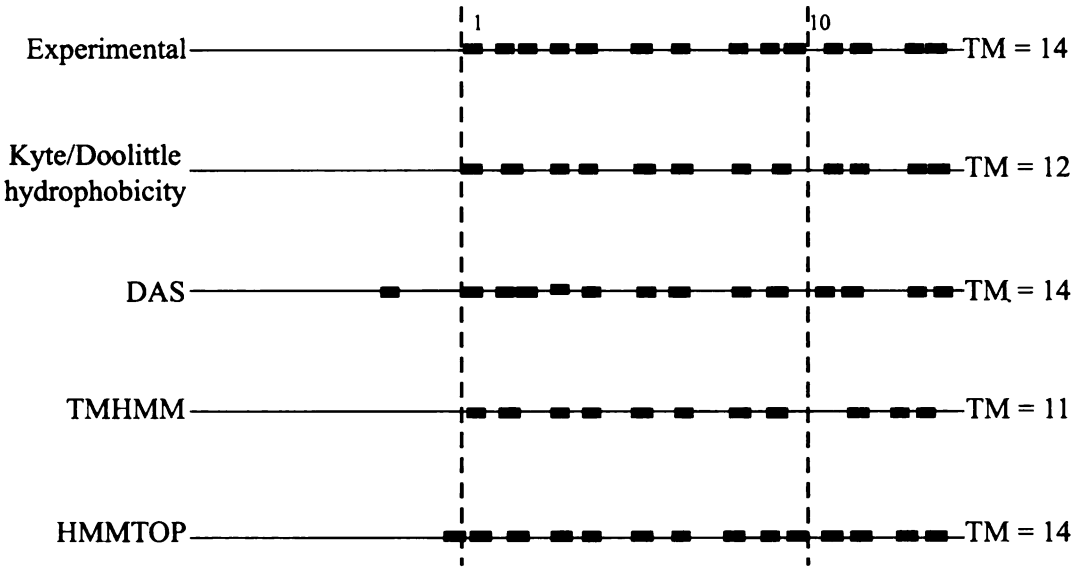


Figure 3: Comparison of the experimentally determined and predicted topology of band 3

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 hydropathy 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.

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 in the last 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 uses machine-learning algorithms called hidden Markov models (HMM) that 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 lumenal domains and compare the results to those found in training sets. 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 to 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 figure 3 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 incorrectly predicted orientation of 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 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.

Summary

Integral membrane protein biogenesis requires 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 likely mediate each stage of biogenesis for single spanning and polytopic membrane proteins. The importance of these regulated interactions is illustrated by looking at 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 determinant of membrane integration.

Many of the mechanisms that regulate integral membrane protein biosynthesis remain debated or elusive. The following body of work elucidates a previously unappreciated role for the signal sequence in regulating TM domain integration. These results have significant implications for understanding both the mechanism and timing of integration. In addition, studies of the properties of the PrP TM domain that affect integration revealed that a disease associated polymorphism alters the topologic distribution of nascent chains which have a significant impact on generation of a diseased state.

Chapter II

Signal Sequences Can Influence Integration of Single-Spanning Membrane Proteins

Summary

Biosynthesis of the prion protein at the endoplasmic reticulum generates multiple topological forms. The topology of an individual chain is determined first by the localization of the N terminus and then by potential integration of the transmembrane domain into the lipid bilayer. Here we provide the first evidence that signal sequences affect the latter of these events by demonstrating that some, but not other, signal sequences and signal sequence mutations result in significant increases in the fraction of prion protein nascent chains that integrate into the lipid bilayer. Through analysis of the prolactin signal sequence, an especially poor integration effector, we find that the N terminal and hydrophobic regions of the signal sequence affect integration most significantly. Mutations in either region result in a considerable increase in the number of chains that integrate. The affect of the signal sequence cannot be attributed to timing of signal cleavage or the state of the ribosome membrane junction, parameters previously found to affect protein biogenesis. Interestingly, we found that signal sequences that are poor integration effectors can promote integration under experimental conditions that allow the nascent chain more time to integrate. We show that the role of the signal sequence in promoting integration is separate from its role in early translocation initiation events. These results have important implications for discriminating between two current models of single-spanning membrane protein biogenesis. We conclude that the transmembrane domain remains functionally competent to integrate until translation termination and that integration of the PrP transmembrane domain is not passive, but mediated by protein-protein interactions.

Introduction

Signal sequences emerge from the ribosome shortly after translation initiation and are bound by the signal recognition particle (SRP), which targets nascent secretory and membrane proteins to the endoplasmic reticulum (ER) (Keenan et al. 2001). After interacting with the SRP receptor, SRP transfers the signal sequence to the translocon, an aqueous channel in the ER membrane. After targeting, the signal sequence is recognized by the core component of the translocon, the Sec61 complex (Gorlich and Rapoport 1993; Jungnickel and Rapoport 1995; Morgan et al. 2002). The signal sequence is reoriented with the N terminus on the cytosolic side of the translocon shortly after translocation begins to generate secretory proteins and transmembrane proteins which have the C terminus localized in the ER lumen (Goder and Spiess 2003). The targeting and translocation initiation functions of the signal sequence are thought to be independent of the nascent passenger protein. Signal sequences have no strict consensus sequence. They vary in length from 15 to 50 amino acids (aa) but in general have hydrophilic N and C-terminal domains, and a central hydrophobic domain (Martoglio and Dobberstein 1998). Because of these loose structural requirements, signal sequences were long considered interchangeable.

In the last few years, new roles for signal sequences have come to light by studying the effects of their substitution and mutation on secretory protein biosynthesis (Kim et al. 2002; Kim et al. 2001; Li et al. 1996; Rutkowski et al. 2001; Rutkowski et al. 2003). If signal sequences were interchangeable, then substituting one signal sequence for another would be expected to have little effect. However, examination of the ribosome-membrane junction shortly after translocation initiation reveals that while some

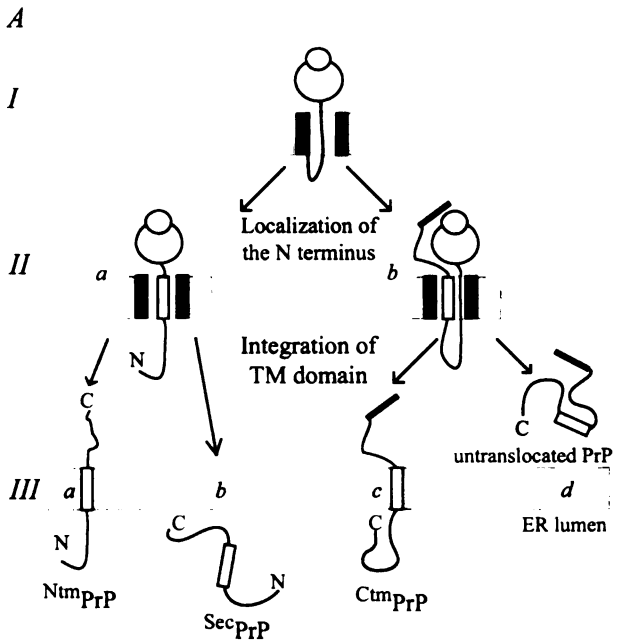
signal sequences close the ribosome-membrane junction early, shielding the nascent chain from the cytosol, other signal sequences close the junction later, allowing longer exposure of the chain to the cytoplasm (Rutkowski et al. 2001). Another step in translocation is opening of the lumenal gate of the translocon to allow passage of the nascent chain into the ER. Kim et al. (Kim et al. 2002) demonstrated that the ability of a signal sequence to open the lumenal gate is matched with the gating requirement of the mature domain. From this they concluded that signal sequence diversity is functional—not simply random variation.

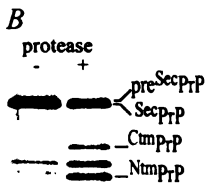
A recent study revealed that mutation of signal sequences, including that of the well-studied secretory protein, prolactin (Prl), could alter maturation of the attached nascent chain. Specifically, subtle changes in the hydrophobic region of the signal sequence affect the interaction of the nascent chain with the translocon and, as a result, the timing of signal cleavage and the extent of mature domain glycosylation. These point mutations did not alter the timing of ribosome-membrane junction closure, suggesting another, separate, substrate-specific role for signal sequences in regulating biogenesis (Rutkowski et al. 2003).

The prion protein (PrP) is a glycoprotein that has come to prominence because of its involvement in an unusual set of neurodegenerative disorders, termed prion diseases (Prusiner 1998). Studies of PrP biogenesis have revealed some remarkable features: most notably, it can be synthesized in three different topological forms. Two different steps in biogenesis, the localization of the N terminus and membrane integration, determine the topology of an individual PrP nascent chain (Kim et al. 2001). ^{Ntm}PrP results from lumenal localization of the N terminus and integration of the transmembrane (TM)

domain (see figure 4A). If integration fails to occur, ^{Sec}PrP is generated, which is entirely translocated into the ER lumen. Localization of the N terminus to the cytosol, followed by integration of the TM domain generates ^{Ctm}PrP. These forms can be distinguished after proteolysis because they are protected from protease to different extents (see figure 4B). Several studies have demonstrated that mutation of the PrP signal sequence, or substitution of signal sequences from other proteins, can alter topologic distribution. Specifically, signal sequence mediated regulation of both ribosome-membrane junction closure and translocon gating affect localization of the PrP N terminus and result in increased ^{Ctm}PrP (Kim et al. 2002; Rutkowski et al. 2001). Regulation of the second stage in PrP topology determination, integration, has been considered the task of the TM domain (Kim et al. 2001).

An exception to this paradigm that has not yet been explored is the observation that substitution of the PrP signal sequence with that of Prl (generating the chimera Prl-PrP) alters the topological distribution of PrP yielding almost entirely ^{Sec}PrP (Rutkowski et al. 2001). In this case the consequence of swapping the signal sequences includes *both* shifting the localization of the N terminus to the ER lumen and reducing TM domain integration. In this study we compare the ratio of ^{Ntm}PrP to ^{Sec}PrP and we observe that several other signal sequences affect ^{Ntm}PrP integration as well. Analysis of the effect of chimeric constructs and mutations on integration reveals that the N terminal and hydrophobic domains of the signal sequence affect integration most significantly. To try to dissect the effect of the signal sequence on integration from its effect on N-terminus localization, we generated Prl-PrP chimeras with point mutations in the Prl signal sequence known to affect signal cleavage timing and glycosylation. We hypothesized that





C

Infγ	MSYTTYFLAFQLCVTLC
Prl	MDSKGSSQKGSRLLLLLVVSNLLLCQGVVS
IgG	MDIWIFLFILSGTAG
GH	MAADSQTPWLLTFSLLCLLWPQEAGA
Ost	MRLAVVCLCLFGLASC
Lep	MRCGPLCRFLWLWPYLSYVEA
PrP	MANLSYWLLALFVAMWTDVGLC
ANP	MSSFTITVSFLLVLVFQFPGQTRA
βLac	MDIQHFRVALIPFFAAFCLPVFA
Ang	MRKRAPQSEMAPAGVSLRATILCLLAWAGLAAG

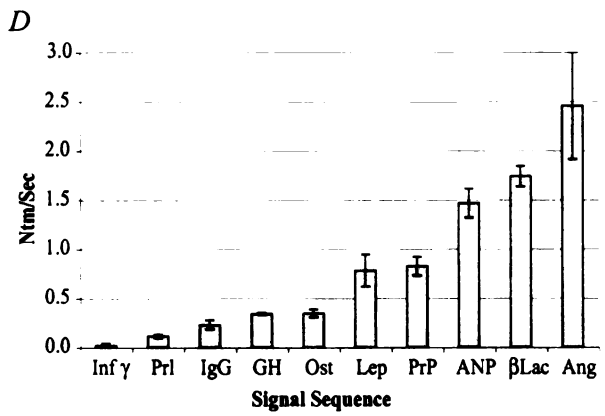


Figure 4: Many Signal Sequences alter ^{Ntm}PrP integration.

(A) PrP topology is determined by two different steps during biogenesis. After the signal sequence targets the nascent chain to the translocon (*I*) the N terminus localizes to either the ER lumen (*IIa*) or to the cytosol (*IIb)*. Integration of the transmembrane domain is the second step. If the N terminus is localized to the ER lumen and integration occurs ^{Ntm}PrP is generated (*IIIa)*. If integration does not occur the entire protein is translocated into the ER lumen generating ^{Sec}PrP (*IIIb)*. ^{Ctm}PrP is generated when the N terminus is localized to the cytosol and the transmembrane domain integrates (*IIIc)*. Failure to integrate when the N terminus is in the cytosol results in the entire chain remaining in the cytosol (*IIId)*. (B) The different topologic forms of PrP are can be identified after proteolysis based on their size: signal cleaved and signal uncleaved (pre) ^{Sec}PrP are completely protected, while ^{Ntm}PrP and ^{Ctm}PrP are protected to different extents. Constructs with various signal sequences (C) on the PrP mature domain were transcribed and translated *in vitro* in the presence of microsomal membranes. After isolation of the microsomal membranes, samples were split and incubated in the presence (+) or absence (-) of PK and separated by SDS PAGE. The fraction of chains synthesized in each topological form was quantitated and the ratio of ^{Ntm}PrP to ^{Sec}PrP from each reaction is graphed (D).

since these mutant signal sequences alter biosynthesis of Prl after closing of the ribosome-membrane junction, they would not alter the junction-sensitive localization of the PrP N-terminus, but only affect integration.

SNtoQT and SNtoMP point mutations in the hydrophobic region of the Prl signal sequence that have been shown to alter signal cleavage timing and glycosylation of the Prl mature domain also affect integration when present in the Prl-PrP chimera. We analyzed these mutants further to try to understand how the signal sequence affects integration. We found that the observed differences in topology are not due to altered closure of the ribosome-membrane junction. Nor can they be attributed to differences in timing of signal cleavage. By manipulating the dissociation of the nascent chain from the ribosome and translocon, we find that the SNtoQT and SNtoMP mutant Prl signal sequences promote integration, while the wild type Prl signal sequence is a weak integration effector. The ability of the signal sequence to promote integration is not irreversibly established early in translocation.

It is widely accepted that TM domains influence integration of one another in multi-spanning membrane proteins (Heinrich and Rapoport 2003; Skach et al. 1993), but in single-spanning membrane proteins the only disparate region known to influence integration is a sequence located directly upstream of a TM domain termed the stop transfer effector (Falcone et al. 1999; Yost et al. 1983). There is some controversy in the field as to how single-spanning membrane proteins integrate. The spontaneous, passive model predicts that the partitioning of the TM domain into the lipid bilayer is due solely to the hydrophobicity of the TM domain and is unregulated both spatially and temporally. In contrast, the scheduled, protein-mediated model proposes that either or both inter- and

intra-protein interactions influence or regulate integration kinetics. We used this newly established system that allows us to look specifically at the effect of the signal sequence on integration to differentiate between the two models of membrane integration. Based on our results we conclude that the PrP TM domain stays associated with the translocon throughout biosynthesis—whether or not the chain integrates.

Experimental Procedures

Plasmid Construction

Plasmids containing the PrP mature domain with the leptin (Lep), angiotensinogen (ang), osteopontin (Ost), interferon-γ (Inf-γ), and atrial naturatic peptide (ANP) signal sequences were generously provided by D. Mitra and R. Hegde. All Prl-PrP signal mutants and subdomain chimeras were generated by directed mutagenesis of Prl-PrP (Rutkowski et al. 2001). The SN mutant library was generated by directed mutagenesis using primers with degenerate sequences (MHN forward). The +120 constructs were generated by digesting PrP (+120) (Kim and Hegde 2002) and Prl-PrP or SNtoQT PrP with *Bsu36I* and *PvuII*. The fragment containing the Prl or SNtoQT signal sequence was then ligated into the PrP (+120) vector. The Prl-PrP (–GPI) STOP was generated by directed mutagenesis of the Ser 240 to TGA. The 104 aa truncation was generated by *NaeI* digestion. All other truncations were generated by PCR.

In vitro translation, translocation and integration assays

In vitro transcription and translation were performed as described previously (Chuck and Lingappa 1992). Translations were carried out at 34°C for 30 minutes unless otherwise noted. Glycosylation was inhibited by 0.2mM tripeptide competitor (Chuck and Lingappa 1992). Where indicated 1.6mM aurintricarboxcylic acid (ATA) or 1mM puromycin and 500mM Potassium acetate were added. Microsomal membrane isolation and ammonium sulfate precipitation have been described previously (Rutkowski et al. 2003). Samples were proteolized at 4°C for 45 minutes with 0.2-0.4 mg/mL proteinase K

(PK). The protease was inactivated by incubation with 10mM PMSF for 5 min and boiling in 10 volumes of 0.1M Tris Acetate pH 8, 1% SDS. Immunoprecipitation of inactivated proteolysis samples was performed in 10 volumes of TXSWB (1% Triton X-100, 100 mM NaCl, 50 mM Tris [pH 8], 10 mM EDTA). After a one hour incubation with the RO13 antibody, protein A beads (Bio-Rad) were added and the samples were incubated overnight at 4°C. Samples were rinsed three times in TXSWB prior to SDS PAGE. Carbonate extraction was performed as described previously (Chuck and Lingappa 1992) except that the samples were centrifuged for 40 min at 80,000 rpm, in a TL100.2 rotor (Beckmann). Both pellets and supernatants were then precipitated with 10% tricarboxcylic acid (TCA).

Miscellaneous

15% tricine and 15% tris-glycine gels were used for SDS-PAGE. Autoradiographs were scanned using an Agfa Arcus II flatbed scanner and quantitated using NIH Image 1.63. Graphs with error bars represent the mean and standard deviation calculated from triplicate reactions.

Results

Signal Sequences affect integration of ^{Ntm}PrP

Previous studies revealed that substitution of the PrP signal sequence with that of Prl results in a significant decrease in ^{Ntm}PrP integration. To determine if the ability of the Prl signal sequence to influence ^{Ntm}PrP integration was unique, or a general property of a subset of signal sequences we decided to look at several other signal sequence substitutions. We examined the effect of the immunoglobulin G (IgG), β lactamase (βlac), growth hormone (GH), leptin (Lep), angiotensinogen (ang), osteopontin (Ost), interferon-γ (Inf-γ), and the atrial naturatic peptide (ANP) signal sequences on integration of the wild type mature domain of PrP (see figure 4C for sequences). It was already known that many of these signals alter the levels of ^{Ctm}PrP in the wild type context (Rutkowski et al. 2001) or in the presence of the A120L mutation (Kim et al. 2002). For the purposes of the present study we needed a way to discriminate between effects on localization of the N terminus and effects on integration, both of which can alter ^{Ntm}PrP levels. Toward that end, we quantitated the percent of PrP nascent chains generated in each topologic form and then calculated the ratio of percent ^{Ntm}PrP to percent ^{Sec}PrP. This allowed us to distinguished between changes in the localization of the N terminus and integration because the former will increase both ^{Sec}PrP and ^{Ntm}PrP equally and thus the ratio of the two conformers will remain constant, while a change in the latter will alter this ratio. A small value such as that of Prl-PrP (0.11) indicates that very few chains with the N terminus in the lumen are integrating (see Figure 4D). In contrast, ANP-PrP generates more ^{Ntm}PrP than ^{Sec}PrP and has a larger ratio of 1.46. Looking at the results in

figure 4D as a whole it is clear that signal sequences do affect integration of ^{Ntm}PrP in a way that is distinct from any affect they have on localization of the N terminus. Like the Prl signal sequence, in the presence of the Inf-γ, IgG, GH, and Ost signal sequences fewer PrP nascent chains whose N termini are in the ER lumen, integrate. The leptin signal sequence causes similar levels of integration to the PrP signal sequence, while the ANP, βLac and Ang signal sequences cause increased integration.

The N terminal and hydrophobic domains of the signal sequence have the largest effect on integration

To gain a better understanding of the features of the signal sequence that affect ^{Ntm}PrP integration we chose to study the Prl-PrP construct in more detail. Prl-PrP was selected for several reasons. Very few Prl-PrP chains integrate and, although Inf-γ makes less ^{Ntm}PrP than Prl-PrP it also makes significantly more ^{Ctm}PrP. The near absence of ^{Ctm}PrP in Prl-PrP means almost all nascent chains are distributed between ^{Ntm}PrP and ^{Sec}PrP, the two populations of most interest in the present study. In addition, Prl signal sequence targeting and translocation initiation has been extensively studied (Jungnickel and Rapoport 1995; Mothes et al. 1998; Voigt et al. 1996). Mutations in the Prl signal sequence that affect maturation of the Prl mature domain have been identified and are available to study the effects of the Prl signal on PrP biogenesis (Rutkowski et al. 2003).

The first question we wanted to address was what region or regions of the signal sequence are important for mediating the effect on ^{Ntm}PrP integration. We identified the central hydrophobic domain (H domain) and the N and C terminal hydrophilic domains of the PrP and Prl signal sequences and then generated constructs with every possible

combination of Prl and PrP signal sequence regions (see figure 5A). Constructs were named based on the region in the N, H, and C domains, with the letters "P" and "L" representing regions from the PrP and Prl signal sequences respectively (for example under this nomenclature Prl-PrP is called LLL). We then quantitated the ^{Ntm}PrP / ^{Sec}PrP value for each construct. In all cases the constructs with the Prl hydrophobic domain integrated less (had lower ^{Ntm}PrP / ^{Sec}PrP values) than constructs with the PrP hydrophobic domain (compare figure 5B upper and lower panels). In addition, the Prl N and H domains together integrated poorly regardless of the C-terminal domain used. The signal sequence is poorly cleaved from the construct with the Prl N and H domain but the PrP C domain (LLP), which is surprising because the PrP signal cleavage site is in tact. However, ^{Ntm}PrP with the signal sequence attached can be detected (figure 5B; see below for the effect of signal cleavage on ^{Ntm}PrP integration). In contrast to the N and H domains, replacement of the PrP C terminal domain with that of Prl (PPL) has a small effect on integration. These results suggest that the Prl N and H domains have the most significant impact on ^{Ntm}PrP integration, and are consistent with previous findings on the effect of signal sequence domains on protein biogenesis (Rutkowski et al. 2003).

Characterization of the effects of mutations in the N terminal and hydrophobic domains of Prl-PrP on ^{Ntm}PrP integration

We reasoned that if the N and H domains were responsible for the effect of the Prl signal sequence on ^{Ntm}PrP integration then mutations in these regions would also be able to alter ^{Ntm}PrP integration. We made several mutations in each region of the Prl signal sequence and quantitated the ratio of ^{Ntm}PrP to ^{Sec}PrP (see in figure 6A and B). In the N

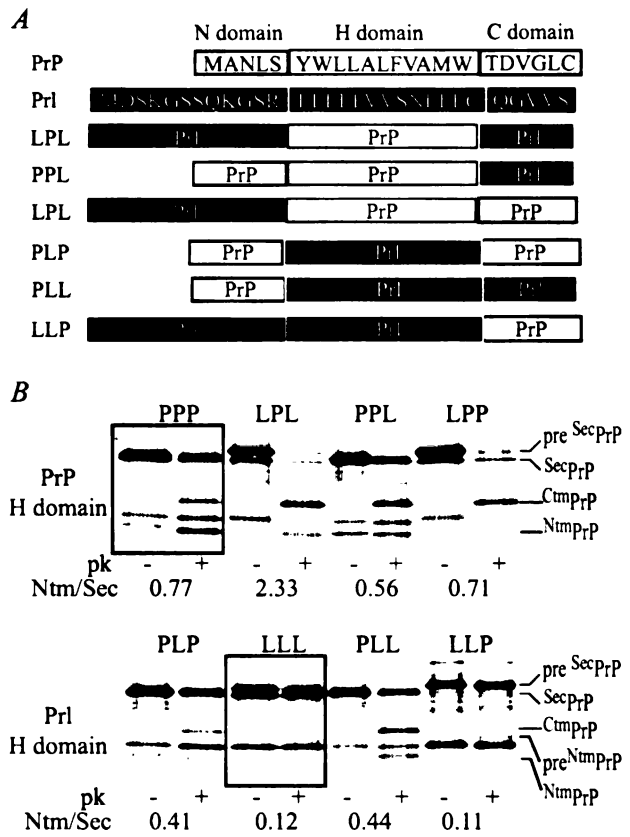


Figure 5: Identification of regions of the signal sequence responsible for altering NtmPrP integration.

(A) The N, H, and C domains of the PrP and Prl signal sequences are shown. To assess the affect of each domain on NtmPrP integration we generated chimeras in which the Prl and PrP signal sequence domains were interchanged. Each construct is named according to the sequence in the N, H, and C domains, with P representing a region from PrP and L representing a region from Prl. Each construct was transcribed, translated and treated as described in figure 1. Samples were separated by SDS-PAGE on 15% tricene gels and the NtmPrP / SecPrP value for each construct is indicated below the gels. A box indicates the complete PrP (upper panel) and Prl (lower panel) signal sequences. With the exception of LLP, which has a signal sequence that remains uncleaved, all calculations were performed by quanitating only the signal cleaved material. Representative data is shown.

domain we mutated Lys9, a residue that is highly conserved among Prl signal sequences from different species, to an oppositely charged amino acid, Asp, and to a hydrophobic amino acid, Leu. To our surprise, these very different mutations both caused approximately half as many more chains to integrate. We wondered whether changing the Ser and Arg at positions 11 and 12 in the N domain to two neutral hydrophilic residues (Gln and Thr) or two hydrophobic residues (Leu and Leu) would affect ^{Ntm}PrP integration. Like at position 9, these very different substitutions both resulted in higher ^{Ntm}PrP / ^{Sec}PrP values. These results support the conclusion that the N domain can affect integration.

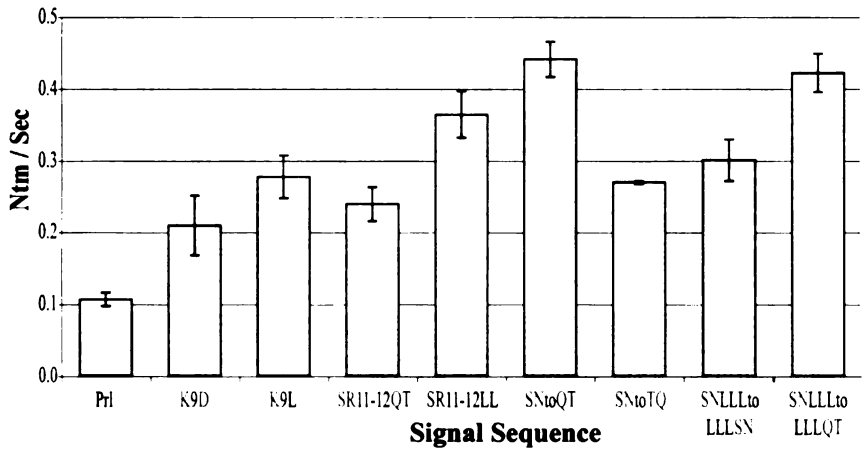
Because the H domain of PrP is less hydrophilic than the H domain of Prl (see figure 5A), we hypothesized that making the Prl H domain less hydrophobic would lead to increased integration of ^{Ntm}PrP. To test this we changed some of the leucines in the Prl H domain to less hydrophobic amino acids, Ala or Phe. The construct in which the first five Leu were changed to Ala failed to target to the ER. Substituting Leu 14-16 with Phe or Leu 22-24 with Ala did significantly increase the ^{Ntm}PrP / ^{Sec}PrP value of Prl-PrP (see figure 6A). These results suggest that the hydrophobicity of the Prl signal sequence might have a significant impact on ^{Ntm}PrP integration.

An interesting characteristic of the Prl H domain is that the string of hydrophobic residues is interrupted by two neutral hydrophilic amino acids, Ser and Asn at positions 20 and 21. Mutation of these residues in the context the Prl mature domain has been shown to affect nascent chain biogenesis by altering the association of the nascent chain with translocon proteins and varying the timing of signal cleavage. In addition, mutation of these residues can affect post-translational modification of the prolactin mature

A

Signal Sequence		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Prolactin	0.11	M	D	S	K	G	S	S	Q	K	G	S	R	L	L	L	L	L	V	V	S	N	L	L	L	C	Q	G	V	V	S
K9D	0.21									D																					
K9L	0.28									L																					
SR11-12QT	0.24											Q	T																		
SR11-12LL	0.36											L	L																		
L22-24A	0.32																						A	A	A						
L14-16F	0.36														F	F	F														
L13-17A	NA													A	A	A	A	A													
VVSN18-21SNLL	1.09																		S	N	L	L									
VVSN18-21QTLL	0.61																		Q	T	L	L									
SNLLL20-24LLLSN	0.30																				L	L	L	S	N						
SNLLL20-24LLLQT	0.42																				L	L	L	Q	T						

B



C

position 20	position 21	Ntm / Sec
S	D	0.07
H	H	0.09
S	N	0.12
N	Q	0.12
P	K	0.12
P	N	0.13
P	H	0.14
N	H	0.15
Q	N	0.16
P	Q	0.17
L	P	0.19
H	P	0.19
Q	Q	0.21
K	K	0.24
T	Q	0.27
I	N	0.27
P	P	0.27
T	T	0.31
T	L	0.34
P	T	0.35
M	P	0.37
S	T	0.38
L	L	0.44
Q	T	0.44
I	T	0.49
T	I	0.51
L	M	0.55
T	P	0.56
M	L	0.56
Q	L	0.66
K	L	0.81

D

Position 20	N	H	S	P	Q	I	L	T	M	K
n	2	2	3	5	4	2	3	5	2	2
Ntm / Sec	0.13	0.14	0.19	0.20	0.37	0.38	0.39	0.40	0.47	0.53

Position 21	D	H	N	K	Q	P	T	I	M	L
n	1	3	4	2	4	5	5	1	1	5
Ntm / Sec	0.07	0.13	0.17	0.18	0.19	0.32	0.39	0.51	0.55	0.56

■ = hydrophobic
■ = hydrophilic, neutral
■ = hydrophilic, charged

Figure 6: Mutations in the Prl signal sequence affect ^{Ntm}PrP integration.

(A) The Prl signal sequences is shown with the central hydrophobic domain shaded grey. Below are the mutants that were examined including the ^{Ntm}PrP / ^{Sec}PrP value. (B) Constructs with the specified signal sequences were translated and treated as in figure 1. Samples were analyzed in triplicate. The average ^{Ntm}PrP / ^{Sec}PrP value and standard deviation are graphed. (C) The library of SN mutants was generated by directed mutagenesis with degenerate primers. Mutants SN to ST, TQ, QN, MP, NH, and LL were generated by mutagenesis using specific primers. Samples were analyzed as described above and the ^{Ntm}PrP / ^{Sec}PrP value is shown. (D) The average ^{Ntm}PrP / ^{Sec}PrP value for all SN mutants with specific residues in the 20th (top) and 21st (bottom) positions are shown. The n value represents the number of constructs that were used to calculate the average. Amino acids are color coded according to their properties: hydrophobic residues are black, neutral hydrophilic residues are blue and charged hydrophilic residues are red.

domain with an engineered glycosylation site (Rutkowski et al. 2003). We mutated Ser20 and Asn21 of Prl-PrP to Gln and Thr because this mutation was known to have a dramatic effect on Prl biogenesis. We hypothesized that this mutation would also affect PrP biogenesis and cause increased ^{Ntm}PrP integration. Figure 6B shows that SNtoQT PrP has an ^{Ntm}PrP / ^{Sec}PrP value that is almost four times larger than Prl-PrP. We also generated SNtoTQ, SNtoQN, and SNtoST mutants. SNtoQN had little effect on integration, while SNtoTQ and SNtoST had increased ^{Ntm}PrP / ^{Sec}PrP values, although not as high as SNtoQT (see figure 6B and C). Like Ser and Asn, Gln and Thr are neutral hydrophilic residues. The ability of SNtoQT to significantly affect ^{Ntm}PrP integration suggests that the identity of residues 20 and 21, and not the interruption of the hydrophobic domain, influences ^{Ntm}PrP integration.

In addition to the SNtoQT mutant, two other mutations whose effects on the Prl mature domain have been well characterized are SNtoMP and SNtoNH. On the Prl mature domain the SNtoNH signal sequence behaves similarly to wild type, whereas, the SNtoQT and SNtoMP signal sequences alter biosynthesis (Rutkowski et al. 2003). When we examined the effect of the mutations on PrP biogenesis we found that, like the wild-type Prl signal sequence, the SNtoNH signal sequence generated almost entirely ^{Sec}PrP. In contrast, both SNtoMP and SNtoQT PrP generated significantly more ^{Ntm}PrP (see shaded regions in figure 6C). Thus the result of Prl signal sequence mutations engineered onto a different passenger, PrP, are faithful to their effects on Prl (Rutkowski et al. 2003).

To determine whether changing the location of the Ser and Asn in the Prl H domain altered ^{Ntm}PrP integration, we generated constructs which shifted the Ser and Asn two residues to the left [VVSN(18-21)SNLL] or three positions to the right [SNLLL(20-

24)LLLSN]. VVSN(18-21)SNLL altered both the localization of the N terminus (evidenced by a significant increase in ^{Ctm}PrP; data not shown) and integration of the transmembrane domain (^{Ntm}PrP / ^{Sec}PrP value of 1.09), while SNLLL(20-24)LLLSN primarily affected integration. We also generated similar constructs that included QT in place of SN both which also had higher ^{Ntm}PrP / ^{Sec}PrP values than Prl-PrP (see figure 6A and B). The results of the SN shifted mutants are somewhat surprising because in contrast to the L(14-16)F and L(22-24)A mutants, these mutations affect the ^{Ntm}PrP / ^{Sec}PrP value without changing the overall hydrophobicity of the signal sequence.

Mutation of SN(20-21) in the Prl signal sequence affects ^{Ntm}PrP integration

The initial substitutions we made at positions 20 and 21 suggested that the identity of the amino acids at these positions was important for the low ^{Ntm}PrP integration rate of Prl-PrP. Since Ser and Thr, and Asn and Gln are very similar amino acids, which have very different effects on ^{Ntm}PrP integration we decided many more mutants were needed to try to determine what amino acid characteristics are important to mediate the effect of the signal sequence on integration. Toward that end, we generated a library of constructs with random mutation of the SN residues and assayed the effects of the mutations on integration (see figure 6C). The presence of charged hydrophilic residues resulted in lower ^{Ntm}PrP / ^{Sec}PrP values with one exception. No trend is as evident with neutral, hydrophilic or hydrophobic residues, however.

Because we could draw few conclusions from the raw data, we analyzed the mutant library to try to determine the effect of specific amino acids at each position. To do this we averaged the ^{Ntm}PrP / ^{Sec}PrP values of all mutants with the same amino acid in

the 20th position, and separately, the 21st position. For example, we averaged the ratios of QN, QQ, QT and QL to get a value of 0.37. Figure 6D shows the results of this analysis. The results for some amino acids vary. For example, Pro in position 20 has a low average ^{Ntm}PrP / ^{Sec}PrP value while in position 21, it has a much higher ^{Ntm}PrP / ^{Sec}PrP value. In contrast, Asn and His have low ^{Ntm}PrP / ^{Sec}PrP values and Thr has high ^{Ntm}PrP / ^{Sec}PrP values regardless of location. In general, hydrophobic amino acids have higher average values and hydrophilic amino acids have lower average values, but there are several exceptions to this trend. We can conclude that the identity of the amino acid at position 21 more strongly correlates to hydrophobicity because the average ^{Ntm}PrP / ^{Sec}PrP values of mutants with hydrophobic amino acids at position 21 are all quite high.

Signal sequence effects on integration are not mediated by expected mechanisms

Since this is the first clear demonstration that the primary structure of the signal sequence can influence integration *per se*, we wanted to determine whether this affect is mediated by previously described or novel mechanisms. The state of the ribosome membrane-junction has been associated with regulating the localization of the N terminus (Rutkowski et al. 2001). Prl-PrP has a closed ribosome-membrane junction, which means that shortly after targeting to the membrane, a tight junction is formed between the ribosome and the translocon. In contrast the signal sequence from β-lactamase on PrP (β-lac PrP) has an open junction (Rutkowski et al. 2001). This was assayed by translating truncated mRNAs that have no stop codon. As a result, the ribosome remains associated with the nascent chain for a relatively long time (Perara et al. 1986). Upon addition of

protease, the "closed" ribosome-membrane junction protects Prl-PrP, but β-lac PrP is accessible to protease through the "open" ribosome-membrane junction.

The SNtoQT mutation on the Prl mature domain affects biogenesis without altering closure of the ribosome-membrane junction (Rutkowski et al. 2003). By extension, we predicted that the mutant signal sequences would have closed ribosome-membrane junctions on PrP. When we tested the protease protection of 104 aa truncations we found that differences in ribosome-membrane junction closure did *not* correspond to the effect of the mutations on integration (figure 7A and B). SNtoNH PrP, which integrated as poorly as Prl-PrP, had a more open ribosome-membrane junction. However, SNtoMP PrP, which integrates relatively well, had a more closed ribosome-membrane junction than SNtoQT PrP. Thus there is no clear correlation between the effect of the mutations on integration and closure of the ribosome-membrane junction. The same is true if we look at the signal sequence substitutions. Both βlac-PrP and IgG-PrP have open ribosome-membrane junctions (Rutkowski et al. 2001) but from figure 4 it is clear that they have opposite effects on ^{Ntm}PrP integration. Therefore, we conclude that the observed differences in ribosome-membrane junction closure had no significant impact on integration.

The signal sequence directs the interactions with the translocon, which in turn influence the maturation of the nascent chain. In the case of the mutant signal sequences on the Prl mature domain, one consequence of regulated signal-translocon interactions was significantly altered timing of signal cleavage (Rutkowski et al. 2003). In that case, the SNtoMP and SNtoQT signal sequences began to be cleaved at an early truncation, but the Prl and SNtoNH signal sequences were not cleaved until the chains were much longer

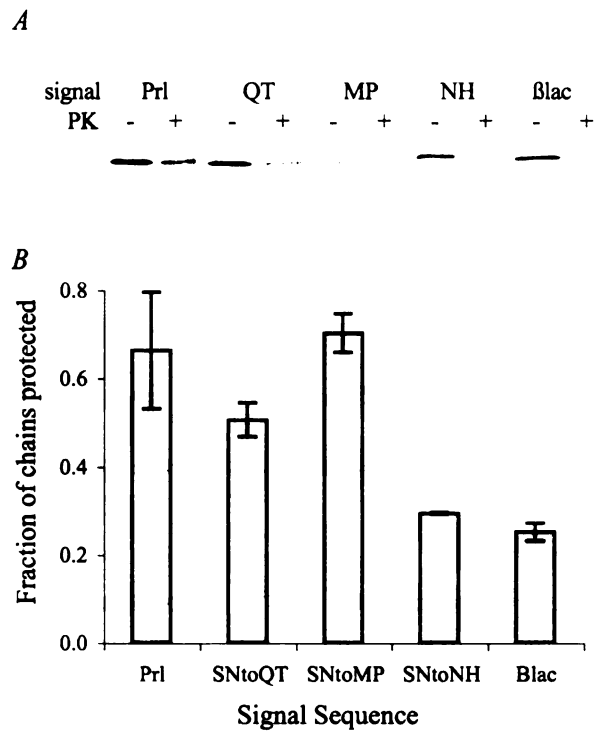


Figure 7: Signal sequence mutations have different effects on ribosome-membrane junction closure.

(A) Constructs with the indicated signal sequences were digested with NaeI. Upon transcription and translation, 104 aa nascent chains were generated that remain associated with the ribosome. The microsomal membranes were isolated and the samples were incubated in the presence (+) or absence (-) of PK and separated by SDS PAGE. (B) The intensity of the bands – and + PK was quantitated and fraction of chains protected is graphed.

(Rutkowski et al. 2003). Since it was already clear that the mutations could affect the interaction of the PrP nascent chain with the translocon (resulting in different topological distributions), we expected that on PrP, the signal sequences would also be cleaved at different times. When we assayed signal cleavage timing we found, to our surprise, that none of the mutations significantly altered the timing of signal cleavage (see figure 8). Although the amount of cleavage varies slightly, all chains begin to be cleaved at the same length. There do appear to be subtle differences in the extent of signal cleavage observed at later points but we found the overall timing of signal cleavage was similar for Prl-PrP and the mutant signal sequence constructs.

Signal cleavage has little effect on ^{Ntm}PrP integration

It seemed possible that the subtle differences in signal cleavage might be responsible for the observed effect on ^{Ntm}PrP integration. To minimize the impact of this difference on our data we performed almost all calculations using the quantitated data from only the signal cleaved ^{Sec}PrP and ^{Ntm}PrP material. However when we included the uncleaved material we still saw significant differences, for example, between Prl-PrP and SNtoQT PrP (data not shown). To test more definitively the effect of signal cleavage on ^{Ntm}PrP integration we generated two pairs of Prl-PrP and SNtoQT constructs. The first mutation we made improved signal cleavage by replacing the first three amino acids of the PrP mature domain (KKR) with those from Prl (TPV) (see figure 9A). These constructs are denoted Prl-PrP +3 and SNtoQT PrP +3. The second pair had a mutation of the Prl cleavage site from GVVS to WPVP (uncl Prl-PrP and uncl SNtoQT-PrP). This mutation prevents signal cleavage as shown in figure 9A. Both making the signal

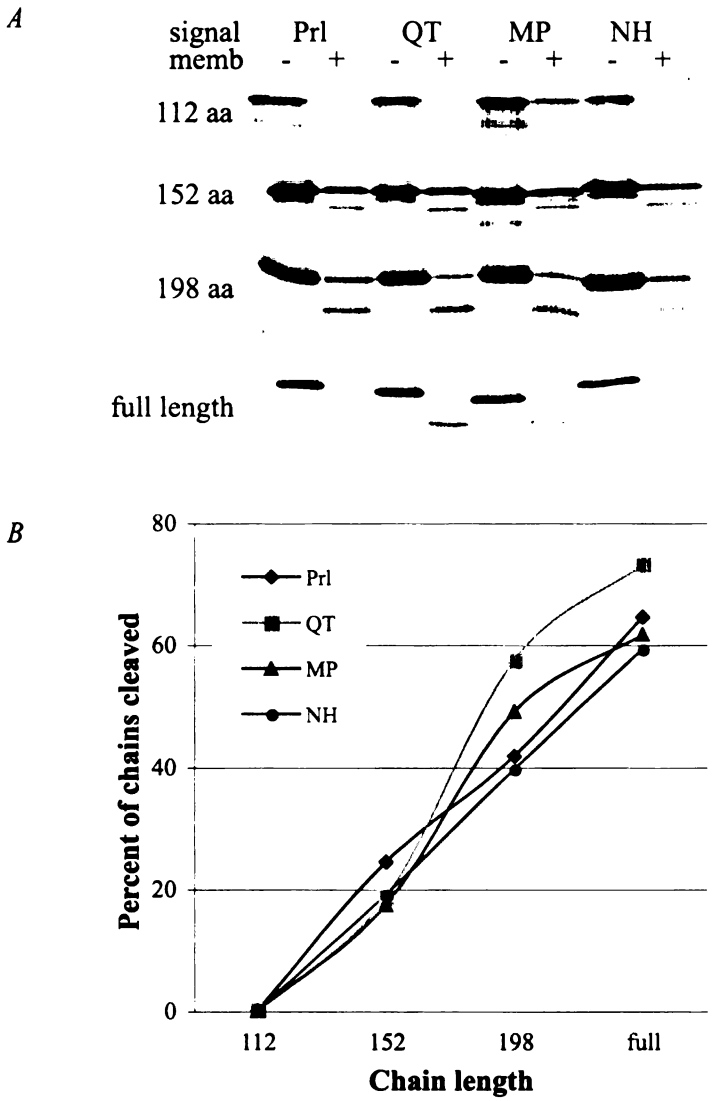


Figure 8: Signal sequence mutations do not alter the timing of signal cleavage.

(A) PrP nascent chains with the Prl or mutant signal sequences of the indicated length were translated in the presence (+) or absence (-) of microsomal membranes (memb). Samples were precipitated with 50% ammonium sulfate (- membranes) or pelleted through a sucrose cushion (+ membrane), solubilized, and separated by SDS PAGE on 15% tricene gels (112aa and 152 aa) or 15% tris-glycine gels (198aa and full length). (B) The intensity of the signal cleaved and signal uncleaved bands in the presence of membranes was quantitated and the percent signal cleavage was calculated.

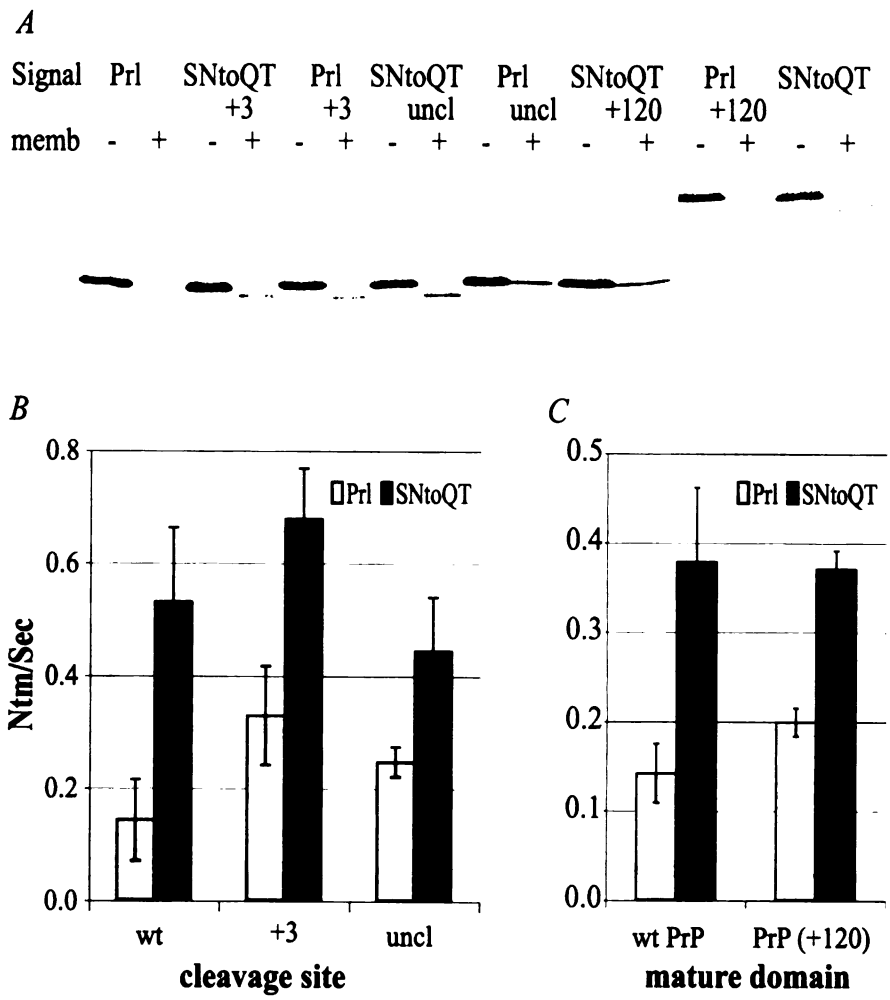

A
Signal Prl SNtoQT Prl SNtoQT Prl SNtoQT Prl SNtoQT
+3 +3 uncl uncl +120 +120
memb - + - + - + - + - + - + - + - +
B
0.8
0.6
0.4
0.2
0.0
Ntm/Sec
Prl SNtoQT
wt
+3
uncl
cleavage site
C
0.5
0.4
0.3
0.2
0.1
0.0
Prl SNtoQT
wt PrP
PrP (+120)
mature domain

Figure 9: Neither signal cleavage nor the distance separating the signal sequence and the transmembrane domain are critical for signal sequence regulation of integration.

Three pairs of constructs were made. In the +3 construct the first three amino acids of the PrP mature domain (KKR) were changed to the first three amino acids of the Prl mature domain (TPV). The last four amino acids of the Prl signal sequence were changed from GVVS to WPVP in the uncl constructs. The +120 constructs have a 120 aa insertion from globin between the signal sequence and the transmembrane domain. (A) The indicated constructs were transcribed in vitro and translated in the presence or absence of microsomal membranes to assess signal cleavage. The samples were then separated by SDS PAGE on 15% tris-glycine gels. (B) Constructs with the indicated signal sequences and cleavage sites were transcribed and then translated in the presence of microsomes and treated as described in figure 1. The percent of each topologic form was quantitated and the average the ^{Ntm}PrP / ^{Sec}PrP value of samples assayed in triplicate is graphed. (C) Prl and SNtoQT PrP constructs with or without the 120 aa insertion were assayed as described above, except after proteolysis the samples were immunoprecipitated with an antibody to the N terminus of PrP. These samples were then separated by SDS PAGE on 15% tricene gels. The percent SecPrP and NtmPrP was quantitated and the average ratio of the two values from triplicate samples is graphed.

uncleavable and more cleavable may have a small affect on Prl-PrP integration, but neither mutation alters the relationship between Prl-PrP and SNtoQT-PrP (figure 9B). Therefore, the observed differences in signal cleavage cannot account for the effect of the signal on integration. It is interesting that the signal has a similar effect whether it remains attached or is cleaved.

The distance between the signal sequence and the TM domain is not significant for signal sequence regulation of ^{Ntm}PrP integration

It has previously been reported that increasing the distance between the signal sequence and the transmembrane domain can alter the localization of the N terminus (Kim and Hegde 2002). We wondered whether increasing the distance between the signal sequence and the TM domain might also decrease the effect of the signal sequence on integration. If that were the case then when the distance is increased we would expect the Prl signal and the SNtoQT signal to generate similar levels of ^{Ntm}PrP. To test this we put the Prl and SNtoQT signal sequences on a construct that has 120 aa of the cytosolic protein globin inserted in the domain between the signal sequence and the TM domain (generating the Prl-PrP (+120) and SNtoQT-PrP (+120) constructs). This insertion doubles the distance between the signal sequence and the transmembrane domain. To properly identify the +120 ^{Ntm}PrP and ^{Sec}PrP fragments we immunoprecipitaed the samples with an antibody to the PrP N terminal region prior to analysis. We then compared the ^{Ntm}PrP / ^{Sec}PrP values (see figure 9C). A few more Prl-PrP chains appear to integrate when the globin spacer is present but the large discrepancy between the Prl and SNtoQT signal sequences is still observed. This indicates that the distance between the

signal sequence and the TM domain has little impact on the ability of the signal sequence to affect integration of ^{Ntm}PrP.

The SNtoQT and SNtoMP mutant signal sequences promote integration

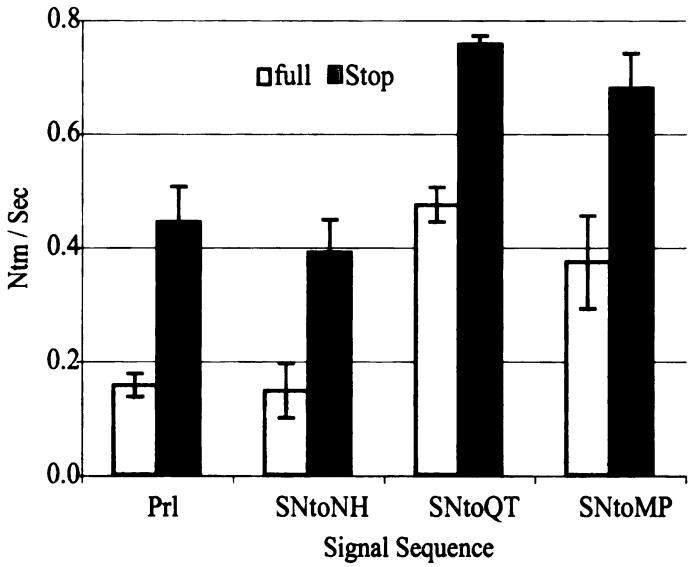
Absent any other explanation for the effect of the signal sequence on integration, three different scenarios could account for the observed results. For simplicity we consider here only the Prl and SNtoQT signal sequences but the logic applies equally well to others. In scenario 1 the Prl signal sequence could prevent integration (for example by occupying the TM domain recognition site), while SNtoQT PrP integrates by default because the signal sequence does a poor job of preventing integration. Alternatively, in scenario 2, the SNtoQT signal sequence promotes integration (perhaps by recruiting appropriate factors or facilitating necessary interactions) and the Prl signal sequence generates ^{Sec}PrP by default because it is a "weak" integration effector. In scenario 3 the Prl signal sequence prevents integration while the SNtoQT signal sequence promotes integration. In this case neither signal sequence is weak, but rather each is effective at different functions. We reasoned that giving a nascent chain more time to integrate would allow us to distinguish between the three possibilities. If a nascent chain had an extended period of time to integrate, the weak signal sequence would have an increased opportunity to block or cause integration. In other words, if the Prl signal sequence were a poor mediator of integration (secretory by default), then giving the chain more time would result in increased integration. Alternatively, if the SNtoQT signal sequence weakly prevented integration (integrated by default), then giving the signal sequence more time would allow it to better prevent integration. If one signal sequence

prevented integration while the other promoted it (scenario 3), there would be no change in the fraction of chains integrated given increased time.

To generate experimental conditions where nascent chains had a prolonged time to integrate, we translated truncated mRNAs generated from each mutant construct, allowed the translation to go to completion, released the nascent chains by treatment with puromycin and high salt, and analyzed topology. Truncation greatly increased the amount of time the nascent chain remained at the translocon, and thus the time during which it could potentially integrate. Figure 10A shows that upon truncation at 262 aa (full length minus the stop codon) constructs with the Prl, SNtoNH, SNtoQT, and SNtoMP signal sequences all made significantly more ^{Ntm}PrP than identical chains that possesses a stop codon. The results were similar for all chain lengths examined (data not shown). This data supports scenario 2, which indicates that the SNtoQT and SNtoMP signal sequences strongly promote integration while the wild type Prl and SNtoNH signal sequences are weak integration effectors.

To ensure that ^{Ntm}PrP made upon truncation and release of the Prl-PrP nascent chain was actually integrated into the lipid bilayer, we tested if it could be extracted by sodium carbonate at pH 11.5, a treatment that disrupts protein-protein, but not protein-lipid interactions (Fujiki et al. 1984). Before we could do this we first had to generate constructs that lacked the GPI-anchor at the C- terminus, which makes ^{Sec}PrP resistant to extraction. The Prl-PrP (–GPI) STOP construct (with a stop codon at 240) and the Prl-PrP (–GPI) truncation (truncated at codon 240; contains no stop codon) behaved similar to the full length constructs in all assays (data not shown). Under conditions in which the control secretory protein was extracted but the control membrane protein was retained

A


0.8
0.6
0.4
0.2
0.0
Ntm / Sec
full
Stop
Prl
SNtoNH
SNtoQT
SNtoMP
Signal Sequence

B

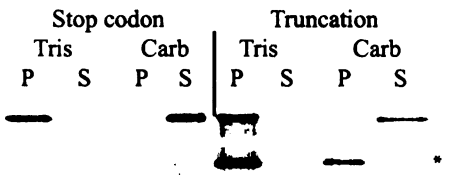

Stop codon
Truncation
Tris
Carb
Tris
Carb
P
S
P
S
P
S
P
S

Figure 10: SNtoQT and SNtoMP signal sequences promote integration.

(A) PrP nascent chains with the indicated signal sequences were synthesized in the presence of microsomal membranes either as full-length nascent chains, containing the authentic stop codon, or as intermediates truncated at the last amino acid. After translation all reactions were treated with puromycin and high salt to release nascent chains still associated with the ribosome. Proteolysis was then used to assess the generation of ^{Ntm}PrP (as described in figure 1). The average ^{Ntm}PrP / ^{Sec}PrP value from the full length and truncated chains is graphed. (B) In order to assess whether the ^{Ntm}PrP generated in (A) is integrated, a new Prl-PrP construct was made which had a stop codon immediately prior to the GPI anchor cleavage site. An identical construct truncated at the GPI cleavage site (no stop codon) was generated by PCR. The samples were translated and released with puromycin and high salt as in (A) prior to membrane isolation and proteolysis. After addition of the protease inhibitor, microsomal membranes were again isolated. The samples were then split and centrifuged in a Tris-sucrose buffer or in carbonate. Both the supernatants (S) and the resuspended pellets (P) were precipitated with TCA prior to separation by SDS-PAGE. The * indicates ^{Ntm}PrP.

(data not shown) we observed significant carbonate extraction resistant ^{Ntm}PrP chains generated by the truncated construct, but few chains integrated in the presence of a stop codon (see figure 10B). This verifies that the ^{Ntm}PrP made by Prl-PrP upon truncation and puromycin release is actually integrated.

Integration promotion is not an early function of the signal sequence

Previously described effects of signal sequences on biogenesis, like regulation of ribosome-membrane junction closure and gating, occur early in biogenesis. However, the increased ability of truncated chains to integrate suggested to us that the signal sequence was acting later in biogenesis to promote integration. To explore this question in a different way we took advantage of the observation that translation temperature could affect integration. Specifically, translation at 34°C or 37°C generated more ^{Ntm}PrP than translation at 26°C or 29°C (see figure 11A). Using temperature manipulation we were able to separate the effects of the signal sequence on early events in translocation from later events. Chains truncated at the C-terminus were translated at 26°C for 10 minutes, which was just long enough for the first full-length chains to appear (data not shown). Then ATA was added to inhibit further initiation (Blobel and Dobberstein 1975). The samples were split and half of each reaction was returned to 26°C, while the other half was shifted to 37°C for 10 minutes to complete elongation. The chains were then released by addition of puromycin and high salt and incubated at 26°C. In this way, translation and translocation initiation of all nascent chains occurred at 26°C. If the ability of the TM domain to integrate were irreversibly established by the signal sequence early in biogenesis, we expected that all chains initiated under the same conditions would yield

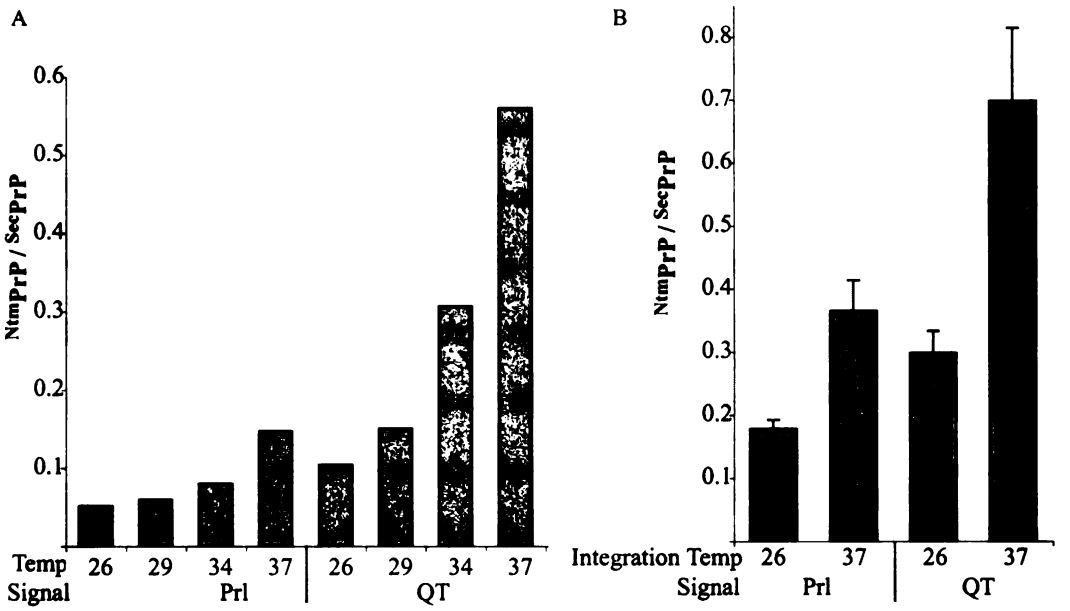


Figure 11: Signal sequences influence integration independent of their role in early translocation events.

(A) Full length Prl-PrP and SNtoQT PrP were translated at the indicated temperatures and treated as in fig. 1. (B) Full-length chains truncated at the C-terminus, were translated for ten minutes at 26°C, then ATA was added to inhibit further translation initiation. The reactions were split and half of each sample was incubated at the indicated temperature. After 10 minutes, puromycin and high salt were added. All samples were returned to 26°C for 5 minutes and then topology was assayed as in fig. 1. In both A and B the ratio of NTMPrP: SecPrP is graphed.

similar levels of ^{Ntm}PrP, regardless of the temperature at which they elongate. However, if the signal sequence impacts events late in biogenesis, then chains elongated at a higher temperature would be expected to generate more ^{Ntm}PrP. We found that elongation temperature did affect nascent chains with both the Prl and the SNtoQT signal sequences. Figure 11B shows the ^{Ntm}PrP to ^{Sec}PrP ratio observed when Prl-PrP and SNtoQT PrP were elongated at 26°C and 37°C. In both cases, the reactions elongated at 37°C generated more ^{Ntm}PrP than the reactions left at 26°C. Although SNtoQT PrP yields significantly more ^{Ntm}PrP than Prl-PrP, the ^{Ntm}PrP to ^{Sec}PrP ratio with both signal sequences nearly doubles at 37°C. From this we conclude that functions of the signal sequence early in biogenesis do not significantly impact the ability of the TM domain to integrate. This supports the conclusion that the signal sequence's integration effector function is distinct from its role in the early events of translocation.

Discussion

In this study we have demonstrated that the signal sequence can affect integration of the PrP TM domain. Previously the signal sequence was thought to influence just the first step in PrP topology determination, namely, the localization of the N terminus, while only the TM domain and adjacent residues were thought to affect the second step, integration. We found that different signal sequences have different effects on ^{Ntm}PrP integration. In addition, we found that mutation of one of these signal sequences, the ^{Sec}PrP favoring Prl-PrP, actually increased the integration of the TM domain. The mechanism by which the signal sequence has an effect on integration is due neither to the closure of the ribosome-membrane junction, nor to an alteration in signal cleavage timing, two parameters previously shown to be important for early functions of the signal sequence. Cleavage of the signal sequence appears to neither promote nor prevent ^{Ntm}PrP integration. In addition we found that doubling the distance between the signal sequence and the transmembrane domain has little effect on integration. When we assayed the ability of nascent chains given more time to associate with the translocon to integrate we found that all constructs generated more ^{Ntm}PrP. This led us to the conclusion that some signal sequences intrinsically promote integration better than others. By temporally separating early effects of the signal sequence from its effects on integration we were able to confirm that the ability of the signal sequence to affect integration occurs after initiation and the early stages of translocation have been completed. The implications of these results for understanding both signal sequence function and membrane protein integration are discussed below.

PrP has been an important model protein for understanding translocational regulation. It was one of the first proteins found to be synthesized in multiple topological forms (Hay et al. 1987a; Hay et al. 1987b; Hegde et al. 1998a; Lopez et al. 1990). More recently it was used to demonstrate that the signal peptide is not a degenerate sequence, but one that can affect how a nascent protein is synthesized (Kim et al. 2001; Rutkowski et al. 2001). Previously, the presence of two oppositely oriented TM forms prevented the use of PrP as a model for understanding integration. Here we separate the effect of the signal sequence on N terminus localization from the regulation of integration. Substitution of two residues in the Prl signal sequence led to an increase in the fraction of chains synthesized as ^{Ntm}PrP. Therefore we were able to increase integration without affecting the localization of the N terminus. This is significant because the range of signal sequence effects has previously been limited to the region of the mature domain adjacent to the signal sequence (Kim et al. 2002; Rutkowski et al. 2003).

Neither of the previously reported mechanisms thought to explain the role of the signal sequence in regulating protein biogenesis explain the effects described here. In looking at both the ribosome-membrane junction and the timing of signal cleavage we found that signal sequences had different effects when connected to PrP than they had on the Prl mature domain. This observation is consistent with previous reports that the signal sequence and the mature domain act together to affect biosynthesis (Andrews et al. 1988; Kim et al. 2002). The observation that the state of the ribosome-membrane junction does not correlate with signal sequence effects on integration reported here suggests that the protein-protein interactions affecting junction closure, which are critical for localization of the N terminus (Rutkowski et al. 2001), are distinguishable from those controlling

integration. The finding that the effect of the signal sequence on integration is not irreversibly established early in biogenesis also supports this conclusion.

Because the H domain is important for mediating the effects of the signal sequence on integration and the Prl H domain is much more hydrophobic than the PrP H domain, we hypothesized that hydrophobicity of the H domain would be important for mediating the effect of the signal sequence on ^{Ntm}PrP integration. When we looked at the effects of mutations that significantly alter hydrophobicity we found two examples [mutants L(14-16)F and L(22-24)A] where making the Prl signal sequence less hydrophobic resulted in increased ^{Ntm}PrP integration. Surprisingly, however, we also found mutations that affect ^{Ntm}PrP integration without significantly changing the signal hydrophobicity [SNLLL(20-24)LLLSN and VVSN(18-21)SNLL]. In addition, the signal sequences shown in figure 4 that affect ^{Ntm}PrP integration vary widely in hydrophobicity. Taken together these results suggest that the overall hydrophobicity of the H domain does not mediate the effect of the signal sequence on integration. Analysis of the SN mutant library does suggest, however, that the presence of a hydrophobic amino acid at position 21 may promote integration.

How the signal sequence influences ^{Ntm}PrP integration is not yet clear. Two observations suggest that a specific consensus sequence for signal sequence integration regulation will not likely emerge. First, we found that while several signal sequences increase ^{Ntm}PrP integration, other signal sequences have the opposite effect, and there are no obvious similarities between the signal sequences that have comparable effects (figure 4). In addition, a large variety of mutations both in the N and H domain of the Prl signal

sequence were able to improve integration efficiency. If there were a simple consensus sequence we would have expected mutation at some residues to have no effect.

Implications of signal sequence regulation on understanding membrane protein integration

In addition to being the first example of an N terminal signal sequence influencing an event as temporally and spatially distant in biosynthesis as integration of the PrP transmembrane domain, this is also the first example of integration of a TM domain being influenced by a region as distant as the signal sequence. These results have important implications for understanding single-spanning membrane protein integration. There has long been debate about how TM domains enter the bilayer. Initially the question was whether hydrophobic domains entered the membrane spontaneously or through a protein channel (Blobel 1980; Engelman and Steitz 1981). After evidence accumulated that membrane proteins pass vectorally into, and transfer laterally out of the translocon (High et al. 1993; Singer et al. 1987; Thrift et al. 1991), the debate shifted slightly, but the questions of how and when a TM domain, once in the aqueous channel, passes into the hydrophobic interior of the lipid bilayer is still at issue.

Evidence has been presented in support of two different models. The first model (figure 12A) predicts that integration is passive and spontaneous. In this case the TM domain, once in the translocon, has access to lipids (Martoglio et al. 1995) and as the chain gets longer the TM domain associates more strongly with lipids, and less with the translocon channel protein Sec61α (Mothes et al. 1997). In this model the equilibrium constant for passive partitioning of the TM domain into the lipid bilayer is influenced

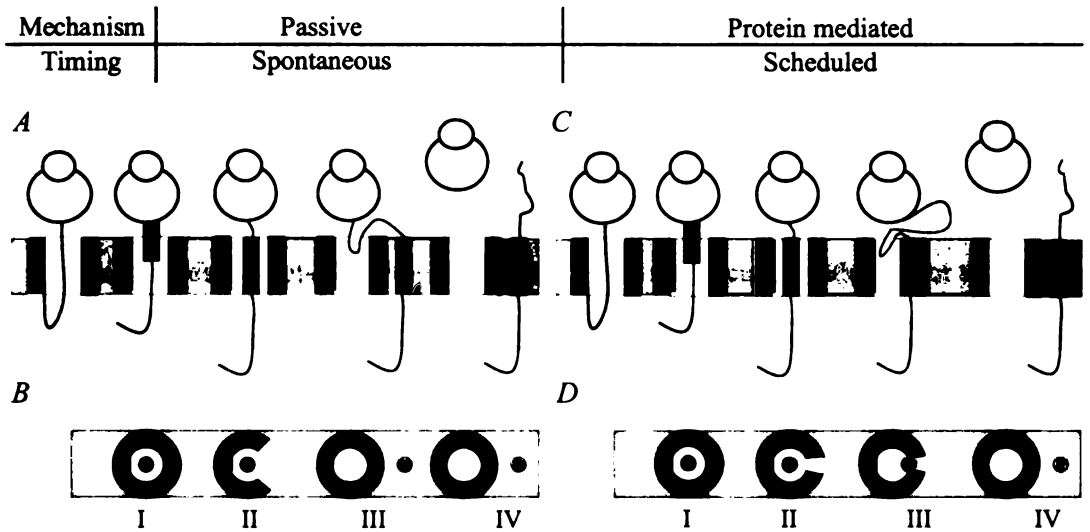


Figure 12: Two models of TM domain integration.

A and B represent the side and bottom view, respectively, of the passive, spontaneous model of membrane integration. Here, after the signal sequence initiates translocation, the TM domain enters the translocon (I), is exposed to lipids (II), and partitions into the bilayer (III) independent of translation termination (IV). C and D represent similar views of the protein-mediated, scheduled model of integration. In this model after signal sequence initiation of translocation the transmembrane domain enters the translocon (I), the nascent chain has limited access to lipids (II) and remains associated with proteins in the translocon (III) until late in biosynthesis, often until termination (IV). The red box and red circle represent the TM domain. The translocon is black and the lipids are gray. In this illustration the signal sequence (shown here as a blue rectangle) is cleaved prior to integration however, the data presented in figure 6 indicate that signal cleavage is not essential.

only by the hydrophobicity of the TM domain (Heinrich et al. 2000). The function of the Sec61 complex is thought to be the removal of the barrier presented by the charged phospholipids head groups (Heinrich et al. 2000). The alternate model shown in figure 12B is that integration of the TM domain is scheduled and protein mediated. The TM domain is thought to move from the center of the pore and association with Sec61 stepwise towards the outside of the translocon where it associates with another translocon component, the translocating chain-associated membrane protein (TRAM) (Do et al. 1996). Translocon proteins are thought to bind to the TM domain until late in translocation (McCormick et al. 2003).

Partitioning into the lipid bilayer cannot be solely due to the properties of the TM domain because in the case of Prl-PrP and SNtoQT PrP both TM domains are identical. In fact, the entire chains are identical except two residues in the signal sequence. Thus, it is highly unlikely that integration of the PrP TM domain is passive. The mechanism by which the signal sequence promotes integration is not yet clear. One possibility is that the signal sequence mediates intra-protein interactions that make it easier for the TM domain to partition into the lipid bilayer. This mechanism might be similar to the way TM domains in multi-spanning membrane proteins facilitate integration of other TM domains (Ota et al. 2000; Skach et al. 1993). Alternatively, the signal sequence could promote integration by preventing non-productive interactions or conversely, by promoting productive interaction of the TM domain with translocon proteins like TRAM. These explanations are not mutually exclusive and the real mechanism may involve complex effects of the signal sequence through both intra- and inter- protein interactions.

From our studies we are also able to conclude that the PrP TM domain remains associated with the translocon until translation is terminated. This conclusion is based on the results presented in figure 10. If integration of the PrP TM domain happened spontaneously, then by the time the chain reached the truncation at 262 aa, the integration decision would have already been made and the TM domain of chains that failed to integrate would have been in the lumen. If that were the case, then leaving full-length chains associated with the translocon would not have increased integration. But since more Prl-PrP and SNtoNH PrP chains integrated upon truncation at the C-terminus (figure 10A), we conclude that the TM domain of truncated chains (and by extension the TM domains of full length chains) must remain at the translocon until termination in order to maintain their functional competence for integration. This must be true whether the TM domain integrates or not because the Prl-PrP and SNtoNH PrP full length chain seldom integrate.

The ability to sort out how and when integration occurs has been limited by the tools available. For the most part, crosslinking has been employed (both to proteins and lipids), which can be useful for demonstrating proximity, but can't prove that protein–protein interactions are relevant for a specific process. In fact, the act of crosslinking abolishes the functional ability of the chains to integrate. In contrast, the methods used here allowed us to manipulate the nascent chain in ways that did not prohibit integration. We altered the amount of time allowed for integration by translating truncated chains that were subsequently released from the ribosome. These released chains are able to integrate into the bilayer as demonstrated by resistance to carbonate extraction at high pH. We were also able affect integration by changing the co-

translational translocation reaction temperature. Using these methods to functionally manipulate integration in *trans*, in combination with the signal sequence constructs that manipulate integration in *cis*, we were able to circumvent the limitations of traditional approaches and to differentiate between the current models of integration. This was especially valuable because one practical implication of the surprising observation that increased length of time in the translocon alters the fraction of chains integrated is that chemical crosslinking of truncated substrates cannot be utilized to explore the present phenomenon. This is because, without truncation, it is not possible to sufficiently synchronize the population of nascent chains at a precise point, and yet upon truncation, the fraction of chains that integrate as ^{Ntm}PrP increases, undermining the basis for the distinction that such an experimental approach would endeavor to make.

Although the results of this study clearly support the scheduled, protein-mediated model of integration we propose that the two separate models presented in figure 12 can be combined. It is possible that both mechanisms of integration occur *in vivo* in a protein-specific fashion. Thus, protein mediated, scheduled integration can be utilized *in vivo* to regulate the biosynthesis of some membrane proteins, while the passive, spontaneous model is likely to be valid for others. This would explain why groups working with different substrates or under different conditions would have conflicting mechanistic observations. It is has been well established that both precision and diversity can be achieved by regulating multiple stages in the life of a protein. Integration is another level where regulation can occur. Just as not all signal sequences require TRAM or the translocon-associated protein complex (TRAP) for initiation of translocation (Fons et al. 2003; Voigt et al. 1996), not all TM domains require scheduled protein-mediated

integration. It makes sense that a simple example, such as a bacterial membrane protein, would integrate passively. Clearly different principles apply for the TM domain of PrP, and likely other complex eukaryotic membrane proteins. The idea we present is not that chains never integrate passively and leave the translocon before termination, but that the mechanism and timing of TM domain integration can be regulated.

Chapter III

Affects of the Stop Transfer Effector and Transmembrane Domain on ^{Ctm}PrP Integration

Summary

To date, mutations in the PrP signal sequence, the N terminal region of the TM domain, and the stop transfer effector (STE; the region immediately preceding the TM domain) have been found to alter the fraction of nascent chains that are synthesized as ^{Ctm}PrP, but little is known about how the STE functions or how these regions effect TM domain orientation. In this study we found that the PrP STE is functionally distinct from the IgM STE. The two STEs appear to be adjacent to different proteins in the translocon during integration. Together these results suggest that the two STEs utilize different mechanisms to promote integration. Using chimeric constructs we examined the effect of the STE on integration and orientation of foreign TM domains in PrP. We found that most TM domains integrate as CtmPrP regardless of the STE. To further characterize the features of the PrP TM domain that influence ^{Ctm}PrP integration we generated a panel of chimeric proteins in which different regions of the PrP TM domain were replaced with the corresponding regions from the IgM TM domain. The topology of these proteins appeared to depend primarily on the identity of the hydrophobic region, which we found also mediates the requirement of the TM domain for the STE. Surprisingly, in the presence of the PrP H domain, the PrP C terminal region was essential for ^{Ctm}PrP integration. Through mutational analysis we determined that the identity of the amino acid at residue 129 in this region appears to have the most significant impact.

Introduction

Most proteins that enter the translocon are thought to either pass entirely through the channel or to integrate into the lipid bilayer. The prion protein (PrP) was one of the first proteins shown to be made in multiple topologic forms (Hay et al. 1987a; Hay et al. 1987b; Hegde et al. 1998a; Lopez et al. 1990). During co-translational translocation at the ER, some PrP nascent chains pass completely through the translocon to generate secretory PrP (^{Sec}PrP), others integrate into the lipid bilayer with their N terminus in the ER lumen (^{Sec}PrP), while still others integrate in the opposite orientation (^{Ctm}PrP). The ability of PrP to be made in multiple topologic forms has made it an ideal candidate for the studying regulation of transmembrane (TM) domain integration. Understanding how biosynthesis of ^{Ctm}PrP is regulated has been especially important because ^{Ctm}PrP has been shown to cause neurodegeneration associated with prion disease (Hegde et al. 1998a).

Several regions of PrP have been found to affect integration of ^{Ctm}PrP. One of the most surprising of these is the signal sequence. Substitution of the PrP signal sequence with that from other proteins can increase or decrease the amount of ^{Ctm}PrP generated (Kim et al. 2002; Rutkowski et al. 2001). In addition, mutation of the PrP signal sequence can similarly affect ^{Ctm}PrP production (Kim et al. 2001). The signal sequence is thought to alter ^{Ctm}PrP levels by affecting the localization of the N terminus – not by affecting integration(Kim et al. 2001). Increasing the fraction of nascent chains whose N terminus is in the lumen increase ^{Sec}PrP and ^{Ntm}PrP. Conversely, increasing the fraction of chains with their N terminus in the cytosol increases levels of ^{Ctm}PrP. Signal sequences have been found to affect integration of ^{Sec}PrP (see chapter II), but not ^{Ctm}PrP.

Only two regions have been identified that affect ^{Ctm}PrP integration: the stop

transfer effector (STE) and the TM domain. How these regions together determine PrP topology, however, is unclear. The PrP STE was one of the first regions identified, outside the hydrophobic TM domain, found to promote integration (Yost et al. 1990). Deletion of the PrP STE causes all nascent chains to be made as ^{Sec}PrP. Outside its native context, the STE was also found to promote integration of a hydrophobic sequence that otherwise does not integrate. Interestingly, this sequence does not appear to be 100% efficient. When synthesized *in vitro*, a large fraction of nascent chains do not integrate. In addition, several prion disease-associated mutations have been found in the STE (P102L, P105L. KHtoII), all of which presumably strengthen STE function because they increase production of ^{Ctm}PrP. All of these mutations increase the hydrophobicity of the STE, in some cases, even replacing the charged amino acids. This suggests that the STE is not simply acting as a charged region and facilitating orientation according to the positive inside rule (von Heijne 1992).

STEs have been identified in one other protein to date: IgM (Falcone et al. 1999). Both immediately precede the TM domain and contain several charged amino acids. Although the functions of the two STEs are similar their sequences are not. While the PrP STE contains several Lys and His, the IgM STE contains several Glu. In fact, studies of the IgM STE revealed that the negative charges are important for promoting integration (Falcone et al. 1999), yet the mutations which strengthen PrP STE activity all introduce more hydrophobic amino acids. Exactly how either of these sequences facilitate integration is unknown, however, they are thought to interact with a receptor in the translocon (Mize et al. 1986; Yost et al. 1990). Possible candidates include the translocating-chain-associated membrane protein (TRAM), a 54 kD glycoprotein, and a

69 kD membrane protein, which have been found adjacent to the IgM STE in the translocon. It seems unlikely, however, that two STEs with such different sequences will utilize the same receptor.

The TM domain also plays an important role in determining topology. Unlike conventional TM domains, the PrP TM domain contains many glycines and alanines. It has been speculated that the ability of PrP to be made in multiple topologic forms is due in large part to the unique features of the TM domain (Lopez et al. 1990). Several disease-associated mutations in the PrP TM domain have also been found to significantly increase the generation of ^{Ctm}PrP. These mutations are all in the N terminal region of the transmembrane domain, suggesting that this region is essential for PrP topology determination. This is surprising because the hydrophobic core of the TM domain is conventionally considered to have the largest impact on integration. Although it is now well established that the signal sequence influences localization of the N terminus, a process which affects orientation of the TM domain, the ability of mutations in the STE and TM domain to cause significant increases in ^{Ctm}PrP levels suggests that these regions influence both integration and orientation. Yet, little work has been done to determine what features of the STE and TM domain are important for orientation and integration of the nascent chain.

The goal of the research presented below was to address unanswered questions about TM domain orientation and integration such as: are the IgM and PrP STEs functionally interchangeable? Do these STEs interact with the same proteins in the translocon? How do TM domains from other proteins orient and integrate when placed in PrP? Do these TM domains require the STE? How do different regions of the PrP TM

domain influence integration? The answers to these questions provide insight into the regulation of both TM domain integration in general and ^{Ctm}PrP biosynthesis in particular.

Experimental Procedures

Plasmid construction

Constructs used in *in vitro* translation were derived from the pSP64 haPrP, constructs published previously (Hegde et al. 1998a). To generate the TM domain chimeric constructs, the amyloid precursor protein (APP), immunoglobulin M (IgM), vesicular stomatitis virus glycoprotein (VSVG), T-cell specific surface protein (CD28), Cop-coated vesicle membrane protein p24 (p24), interleukin 2 receptor (IL2r) and asialoglycoprotein receptor (ASGP-R) TM domains were PCR amplified, digested, and ligated into NdeI and NsiI sites outside the PrP TM domain. All other mutants and chimeras were generated by directed mutagenesis.

In vitro translation, translocation and crosslinking assays

In vitro transcription and translation were performed as described previously (Chuck and Lingappa 1992). Prion protein topology was assayed as described in chapter II. For crosslinking studies, prior to transcription the DNA was digested with HincII. After translation the microsomal membranes were isolated and resuspended in 1-2 volumes of fresh 1x physiologic salt buffer (50 mM Hepes pH 7.5, 0.1M KoAc, 50mM MgoAc, 0.25M sucrose). The samples were then split and one aliquot was treated with 0.2 μM disuccinimidyl suberate (DSS) for 30 min at room temperature. The reaction was quenched by addition of $1/5^{th}$ volume of 0.5M glycine, 0.4M Tris-acetate pH 8, 35mM EDTA, 20mM DTT, and 20 mg/ml RNAse A and left for an additional 10 min at room temperature before separation by SDS PAGE.

Miscellaneous

15% tricine and 15% tris-glycine gels were used for SDS-PAGE. Autoradiographs were scanned using an Agfa Arcus II flatbed scanner and quantitated using NIH Image 1.63.

Results

Stop transfer effectors are not interchangeable

Both of the STEs identified to date are thought to interact with a protein receptor to facilitate integration of the TM domain (Falcone et al. 1999; Mize et al. 1986; Yost et al. 1990). Because the IgM and PrP STEs have very different primary structures we thought it unlikely that they both utilized the same receptor. We hypothesized that if the STEs utilized the same mechanism to promote integration then the two STEs should be interchangeable. To test this we generated a chimeric construct in which we replaced the PrP STE with that of IgM. We then assayed topology of the PrP nascent chains by proteolyzing isolated microsomes following *in vitro* translation. For comparison we also assayed topology of wild type PrP and ΔSTE PrP, a construct in which the PrP STE has been removed. Figure 13A shows the results. As expected, deletion of the PrP STE results all chains being made as ^{Sec}PrP. Surprisingly, so does substitution of the PrP STE with that of IgM. No ^{Ctm}PrP is detectable from either construct. This suggests that the two STEs are not interchangeable.

The failure of the IgM STE to functionally replace the PrP STE suggests that the two STEs interact with different receptors to promote integration. Two unidentified proteins have been found to be adjacent to the IgM STE during integration by chemical crosslinking methods in addition to TRAM. To look for proteins adjacent to the PrP during integration we generated two constructs: one in which all the Lys in the protein had been changed to Arg (Kout), and one in which only the Lys in the STE were retained (KSTE). We then used chemical crosslinking with a homobifunctional amine crosslinker

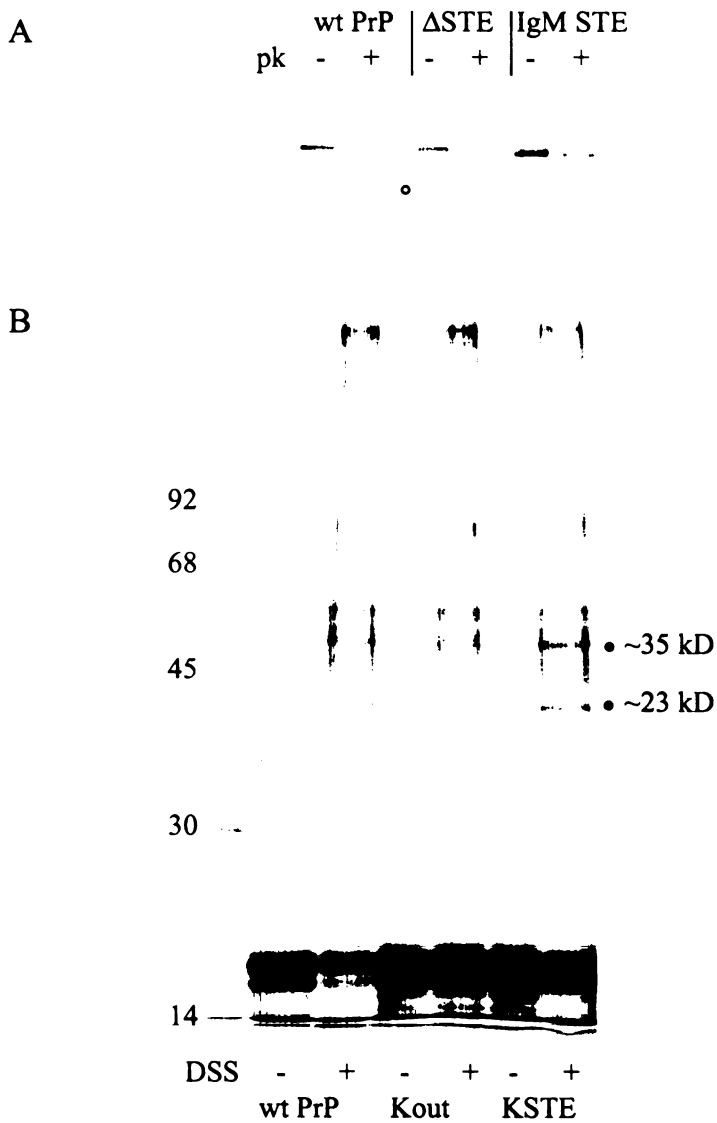


Figure 13: The PrP STE is functionally distinct from the IgM STE.

(A) A construct was made in which the STE from IgM replaced the PrP STE. The topology of this chimera, as well as wild type PrP and a PrP STE deletion mutant was assayed by proteolysis and SDS PAGE following in vitro transcription and translation. ^{Ctm}PrP is marked with an open circle. (B) Constructs in which the Lys were replaced with Arg in the entire protein (Kout) or everywhere except the STE (KSTE) were generated for crosslinking analysis. 180 amino acid truncated nascent chains were incubated in the presence or absence of the crosslinker DSS prior to seperation by SDS PAGE.

to identify proteins that crosslinked to the Lys in the PrP STE. At a length of 180 amino acids we found ~23kD and ~35kD crosslinks that are specific to the PrP STE (see figure 13B). These proteins appear to be in close proximity to all four Lys in the STE because the crosslinks were present when we examined crosslining to chains with only one Lys in each of the four positions (data not shown). The proteins that crosslink with the PrP STE are too small to be the same as the proteins found in proximity to the IgM STE supporting the conclusion that the two sequences utilize different mechanisms to facilitate integration.

How do the STE and TM domain affect orientation?

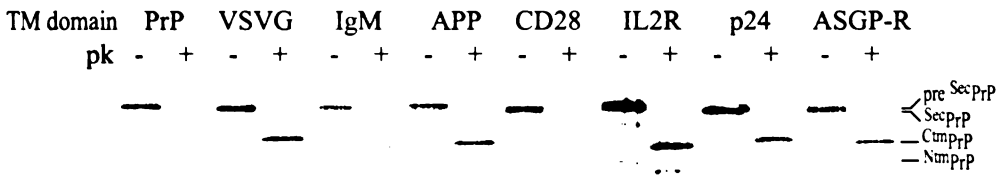
Most integral membrane proteins are only thought to integrate in a single orientation. This orientation is thought to be determined in large part by the charged residues found just outside of the TM domain (von Heijne 1992). Because PrP can be made in multpiple topologies, we hypothesized that foreign TM domains would orient the same way in a PrP chimeric construct as in their native context. To test this we replaced the PrP TM domain with the TM domains from proteins that integrate in either the Ntm (VSVG, APP, IgM, CD28, IL2r, and p24) or the Ctm (ASGP-R) orientation (sequences shown in figure 14A). To our surprise we found that all chimeras examined generate predominantly ^{Ctm}PrP (see figure 14B).

This result is interesting for two reasons. First, it supports the conclusion that although the signal sequence influences localization of the nascent chain, it is not the only determinant of TM domain orientation. Second, the finding that TM domains all integrate in one orientation regardless of their orientation in their native protein suggests

A

	PrP	AGAAAAGAVVGGLGGYMLGSAMS
	VSVG	SFFFIIGLIIGLFLVLRVSIYL
	APP	AIIGLMVGGVVIATVIVITLVML
	IgM	TTASTFIVLFLLSLFYSTTVTLF
	CD28	WVLVVVGGVLACYSLLVTVAFIIFW
	IL2r	VAVA GCVFLLISVLLLSGLT
	P24	VVLWSFFEALVLVA MTLGQIYYL
	ASGP-R	FSLLALSFNILLLVVI

B



C

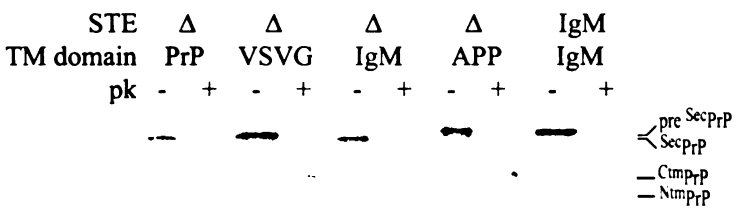


Figure 14: Foreign TM domains in PrP integrate with or without an STE.

The sequences of foreign TM domains used in chimeric constructs are shown in (A). The topology of the chimeric constructs in the presence (B) or absence (C) of the PrP STE was assayed as described in figure 13A. In (C) the final construct contains the IgM STE and TM domain in PrP.

that the contextual information outside the PrP TM domain directs chains to integrate with the C terminus in the lumen. One obvious potential source of information is the STE. Perhaps the more hydrophobic foreign TM domains don't need the STE to facilitate integration, but the positive charges in the STE direct PrP orientation according to the positive-inside rule. To test these questions we generated three ΔSTE constructs with the VSVG, APP, and IgM TM domains and analyzed the topologic distribution of the resultant proteins. The observed levels of ^{Ctm}PrP decreased slightly in these constructs, but there appeared to be no effect on nascent chain orientation (see figure 14D). This suggests that the STE is utilized to direct neither the integration nor the orientation of these TM domains. These results also support the assumption that the relatively decreased hydrophobicity of the PrP TM domain is in part responsible for the toplogical heterogeneity of PrP.

The observation that the IgM TM domain integrates in the PrP chimera is surprising because in its native context it requires an STE. However, the similar levels of integration in the presence or absence of the PrP STE supports the conclusion that the two STEs are not interchangeable. We next wondered how the IgM STE would affect integration of the IgM TM domain in PrP. Although it seemed unlikely that the IgM STE would affect the already high levels of integration, we thought it possible that the negatively charged IgM STE would alter the nascent chain orientation. When we examined the topology of this chimeric construct, however, we found almost exclusively ^{Ctm}PrP (see figure 14). This suggests that neither the PrP nor the IgM STE direct TM domain orientation.

A systematic examination of the features of the TM domain that influence ^{Ctm}PrP integration.

Because all the foreign TM domains examined generate predominantly ^{Ctm}PrP we wondered what specific features of the PrP TM domain are responsible for regulating integration. Was it indeed that the PrP TM domain is less hydrophobic than the others? Or are there special features of the PrP TM domain that make it more capable of achieving multiple topologies. To address these questions we began by generating additional chimeric constructs in which the N terminal, hydrophobic, and/or C terminal regions of the IgM TM domain were substituted for that of PrP (see figure 15A). We did this in constructs with either the PrP or the IgM STE. This also allowed us to address the question of what features of the IgM TM domain are necessary to overcome the requirement for the PrP STE. Constructs were named with the letters representing each domain present. For example, I-PIP has the IgM STE and H domain, and the PrP N and C domains.

We assayed the topology of the chimeric proteins and calculated the fraction of chains synthesized as ^{Sec}PrP or ^{Ctm}PrP. Figure 15B shows the fraction of ^{Sec}PrP normalized based on wild type ^{Sec}PrP levels. This means that constructs with a value larger than one make more ^{Sec}PrP than wild type and constructs with a value less than one make more ^{Ctm}PrP than wild type. Regardless of the STE, constructs containing the IgM hydrophobic domain generated the most ^{Ctm}PrP. In contrast, constructs with the PrP hydrophobic domain generated wild type levels of ^{Ctm}PrP or less. Because the IgM H domain is much more hydrophobic than that from PrP this result supports the hypothesis

A

	STE	N terminal	Hydrophobic	C terminal
IgM	KSTEGEVNAEEEGFENLW	TTASTFI	VLFLLSLF	YSTTVTLF
PrP	KPSKPKTNMKHM	AGAAAAGA	VVGGLGG	YMLGSAMS

B

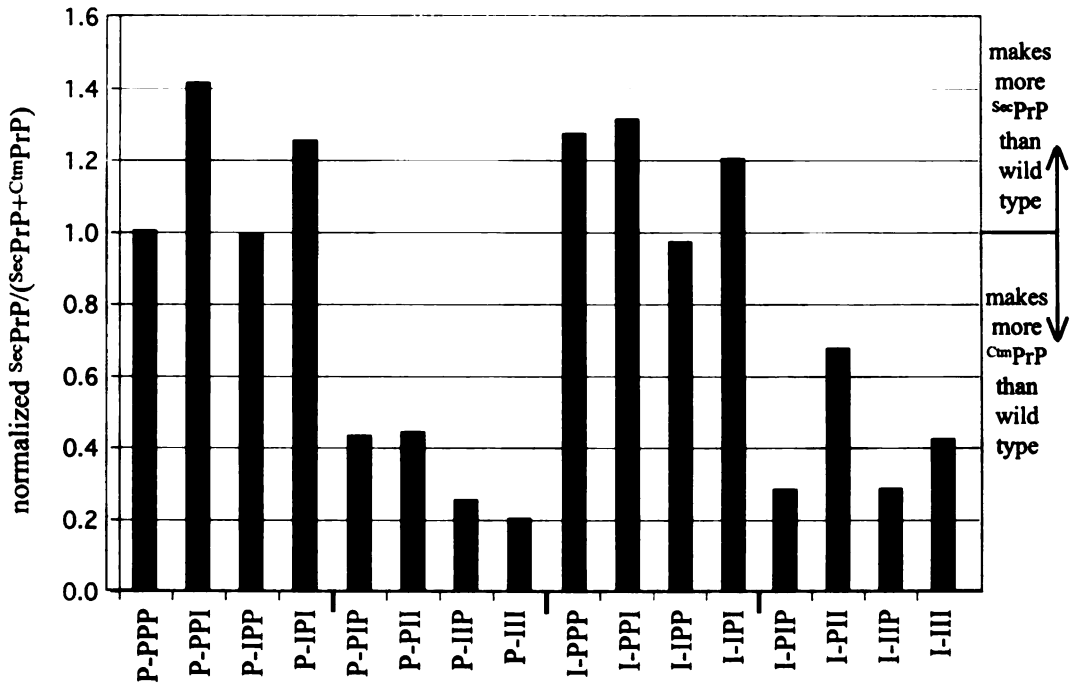


Figure 15: The hydrophobic and C terminal regions of the PrP TM domain influence topology.

(A) The sequence of the IgM and PrP STEs and TM domains are shown. The N terminal, hydrophobic and C terminal regions are indicated. (B) To test the influence of each region of the TM domain on integration we generated chimeric constructs in which the N terminal, hydrophobic and/or C terminal regions of IgM were substituded for those of PrP in the presence of the PrP or IgM STE. These constructs were named according to the identity of the sequence at each position (for example, wild type PrP is P-PPP). The topology of the chimeras was assayed as described in figure 13A, then the ^{Sec}PrP and ^{Ctm}PrP was quantitated and the fraction of ^{Sec}PrP was calculated. These values, normailzed relative to wild type PrP, are graphed.

that the unique hydrophobicity of the PrP TM domain is essential for generating multiple topologies.

A special feature of the hydrophobic region of the PrP TM domain appears to be mediating the requirement for the PrP STE. I-PPP, which is the IgM STE construct also shown in figure 13, makes little ^{Ctm}PrP. However, substitution of the PrP H domain with that of IgM alone, in construct I-PIP, or in combination with other substitutions (I-IIP, I-PII, and I-III) eliminates the need for the PrP STE.

In addition to I-PPP, which has a value above 1.2, four other constructs also made essentially no ^{Ctm}PrP. These constructs all had the PrP H domain, but the IgM C domain. P-PPI, which has only one region different from wild type, makes the least amount of ^{Ctm}PrP. This result suggests that a previously unappreciated region of the PrP TM domain, the C terminal region, can affect ^{Ctm}PrP integration.

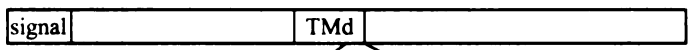
Characterization of a new region of the PrP TM domain that affects membrane integration

To verify that the C terminal region of the PrP TM domain is important for ^{Ctm}PrP integration we wanted to try to substitute this region with the eight C terminal amino acids from other TM domains. We searched through sequence databases and identified two candidate TM domains: the sixth predicted TM domain from the human growth hormone inducible transmembrane protein (GHIT) which has some similarities to the C terminal region of the PrP TM domain, and the last eight amino acids of the TM domain from the human myelin associated glycoprotein (MAG) which is more similar to IgM. In addition we substituted the C terminal region of the PrP TM domain with eight central

hydrophobic amino acids and, separately, the last eight amino acids of the ASGP-R. None of these chimeras generated high levels of ^{Ctm}PrP (see figure 16). Although these results confirm that a feature of the C terminal region of the PrP TM domain is important for ^{Ctm}PrP integration, they do not enable us to draw any additional conclusions. We were surprised that the GHIT sequence, which has several of the same amino acids as those found in PrP, caused such a significant reduction in the levels of ^{Ctm}PrP integration.

We next decided to use mutational analysis to try to further characterize the affect of this region on ^{Ctm}PrP integration. For most mutations, we substituted a few amino acids with those found in one of the other constructs. For example, since GHIT is so similar to PrP we mutated just the first two or, separately, the last three amino acids. The substitutions and results are shown in figure 16. We found that the only constructs that generated wild type levels of ^{Ctm}PrP were those that have Met at position 129. Constructs with a Ser in this position have the lowest levels of ^{Ctm}PrP. From these results we conclude that the identity of the amino acid at position 129 is important for regulating CtmPrP integration.

A



Name	128	129	130	131	132	133	134	135	% ^{CTM}PrP
PrP	Y	M	L	G	S	A	M	S	17
IgM	Y	S	T	T	V	T	L	F	4
GHIT	S	S	L	G	S	M	F	L	4
MAG	I	V	C	Y	I	T	Q	T	7
Asialo hydrophobic	N	I	L	L	L	V	V	I	7
Asialo C term	V	V	I	C	V	T	G	S	7
TAMT	Y	M	L	G	T	A	M	T	17
SS	S	S	L	G	S	A	M	S	4
MFL	Y	M	L	G	S	M	F	L	20
VTLF	Y	M	L	G	V	T	L	F	24
STT	Y	S	T	T	S	A	M	S	2
Δ TM	-	-	-	-	-	-	-	-	0

B

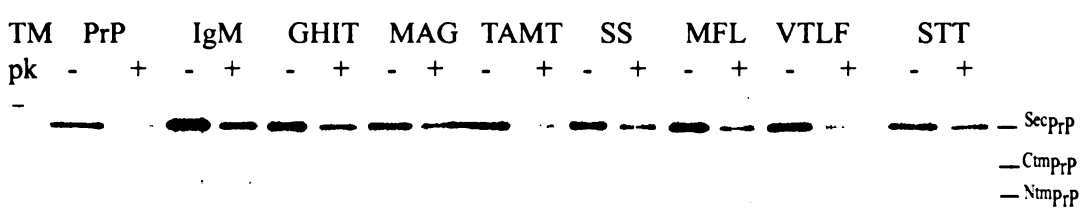


Figure 16: The identity of the amino acid at position 129 influences integration of the PrP TM domain.

Mutant constructs were generated in which the C terminal region of the PrP TM domain was replaced with the sequences indicated in (A). The topology of these constructs was assayed as described in figure 13A and is shown in (B). The % ^{Ctm}PrP was quantitated for each construct and it is indicated to the right of the sequences in (A).

Discussion

In this study we have examined the features of the PrP STE and TM domain that mediate CtmPrP integration. We draw several conclusions from our results: (1) the IgM and PrP STE utilize different machinery to mediate stop transfer; (2) in addition to the signal sequence and TM domain, another region of PrP influences TM domain orientation; (3) the hydrophobic region of the PrP TM domain mediates the requirement for the PrP STE; and (4) in the presence of the PrP hydrophobic region, the C terminal region and more specifically, the identity of the amino acid at position 129 influences ^{Ctm}PrP integration.

STE function is currently poorly understood. To date only two STEs have been characterized. Evidence has been found supporting the idea that each of these STEs interact with a protein receptor at the translocon. However, the results presented here suggest that the two STEs function in different ways. We find different proteins adjacent to the PrP STE during integration than those observed in studies of the IgM STE, although it is possible that the use of different crosslinking procedures may account for the different crosslinks found to the IgM STE versus the PrP STE. This seems unlikely, however, because the two STEs have very different primary structure and we found that the IgM STE is not able to functionally replace the PrP STE. More detailed mutational analysis of the PrP STE is necessary to determine what features of this region are important for mediating STE function.

Many aspects of the PrP nascent chain affect orientation. As mentioned previously, the signal sequence has been shown to direct localization of the N terminus, which in turn can affect TM domain orientation. In addition, the ability of mutations in

the TM domain to increase ^{Ctm}PrP levels suggests that the sequence of the TM domain itself can influence orientation. However, the fact that many of these TM domains integrate in the opposite orientation in their native context suggests that the observed topology cannot be attributed to the foreign TM domain itself. The results presented here suggest that a third and separate region of PrP also provides topologic information. This region does not appear to be the STE because the topology of the chimeric chains is not influenced by deletion of the positively charged STE, or the presence of the negatively charged IgM STE.

Previously, only the N terminal region of the PrP TM domain had been found to influence integration. The hydrophobic region of the PrP TM domain is much less hydrophobic than corresponding regions of other TM domains. It has been speculated that this difference in part accounts for the ability of PrP to be made in different topologic forms(Yost et al. 1990). The results from examining chimeric constructs with either the IgM or the PrP hydrophobic domain clearly demonstrate that replacing the PrP hydrophobic domain can cause increased ^{Ctm}PrP integration. However, further studies will be necessary to determine whether the hydrophobicity of this region alone, or some other feature of this unique sequence is responsible for mediating altered topologies. It is interesting to note that the C terminal region, and presumably residue 129, only affects PrP integration in the presence of the PrP hydrophobic domain.

Implications of residue 129 for prion disease

It has not escaped our attention that in human PrP there is a disease-associated polymorphism at residue 129. While some people are Met homozygotes, others have one

or two alleles with a Val at position 129. The identity of the amino acid has been associated with a propensity to develop sporatic and infectious prion diseases. In addition, several disease causing mutations in PrP are predominantly found on an allele with a Met at position 129, while others mutations are only found with a Val at that position. Increased levels of ^{Ctm}PrP have been associated with familial, sporadic, and infectious prion diseases. The data presented here suggests that the different polymorphic forms of PrP may have different propensities to form ^{Ctm}PrP. This may in part explain the observed association of the polymorphism with disease.

Chapter IV

A Disease Associate Polymorphism in the Prion Protein Affects Generation of the Neurodegeneration promoting ^{Ctm}PrP

Summary

The human prion protein (PrP) contains a Met/Val polymorphism at position 129 which is associated with inherited, sporadic, and infectious prion diseases. Recent evidence suggested that the identity of the amino acid at this position can influence the generation of ^{Ctm}PrP, a transmembrane form of the protein which is known to be associated with disease. To test whether Met and Val at position 129 affect ^{Ctm}PrP biosynthesis differently, we generated constructs and assayed ^{Ctm}PrP production *in vitro* and *in vivo*. We found in both systems the presence of a Met at 129 results in higher levels of ^{Ctm}PrP. In vivo, M129 PrP also made more ^{Ctm}PrP in the presence of a proteasome inhibitor. Substitution of other amino acids at 129 can have both positive and negative effects on ^{Ctm}PrP biosynthesis. Since several disease causing mutations are found commonly with one allelic form (either Met or Val). We examined the effect of the polymorphism in the mutant contexts and observed little difference between the ^{Ctm}PrP levels of the wild type and mutant constructs, with one exception. D178N in the presence of M129 generates more ^{Ctm}PrP than wild type M129 PrP.

Introduction

Prion diseases are a unique class of neurodegenerative disorders that can be inherited, spontaneous, or infectious. Transmission of prion diseases is not thought to occur through bacteria or viruses, but rather by an infectious protein (PrP^{Sc}) that converts the normally folded prion protein (PrP^{C}) into the infectious form. This is the prion hypothesis put forward by Prusiner (Prusiner 1982).

Prion diseases include Creutzfeldt-Jakob disease (CJD), fatal familial insomnia (FFI), and Gerstmann-Sträussler-Scheinker disease (GSS) in humans, as well as animal diseases such as scrapie in sheep, bovine spongiform encephalopathy (BSE; also called mad cow disease), and chronic wasting disease in deer and elk. While FFI and GSS are familial prion diseases, CJD can be familial (fCJD), sporadic (sCJD), acquired iatrogenicaly (iCJD; for example through treatment with hormones extracted from human cadaveric pituitary glands), or acquired by infection with BSE (vCJD). PrP^{Sc} is found in the brains of all patients, regardless of how they acquired the disease, however, in sCJD different subtypes of PrP^{Sc} have been identified based on the glycosylation state and conformation (as assayed by protease resistance). The presence of specific subtypes of PrP^{Sc} correlate with different disease symptoms (Gambetti et al. 2003; Hill et al. 2003). Interestingly, the subtype of PrP is serially transmissible to transgenic mice expressing human PrP (Hill et al. 2003).

A feature of prion biology that has long been unexplained is the existence of a disease-associated polymorphism at position 129 of PrP. Although the distribution is varies in different populations (Mead et al. 2003), in the United Kingdom approximately 37% of the population has a Met at this position, 12% have a Val and 51% are Met/Val

heterozygotes (Collinge et al. 1991). The identity of the amino acid at this position has a significant impact on prion diseases. For example, all of the people who have been diagnosed with vCJD are Met homozygotes (Collinge et al. 1996; Hill et al. 1999; Hill et al. 1997). In addition, as many as 74% of patients with sCJD are Met homozygotes (Gambetti et al. 2003). The disease phenotype of people with a mutation from Asp 178 to Asn depends on the residue at 129 of the mutant allele: if 129 is Met, the person develops FFI, if 129 is Val, the person develops CJD. In addition, the time course of the FFI is longer for Met/Val heterozygotes, than for homozygotes. One explanation that has been put forward is heterozygotes are less susceptible to disease because Met129 PrP^{Sc} is better able to convert Met129 PrP^{C} and that Val129 PrP^{Sc} is better able to convert Val129 PrP^{C}. Bovine PrP has a Mat at position 129, which could account for the protective effect of the Val allele. However, little data has been found to support either conclusion. A recent study examined the conformation and stability of Met and Val 129 PrP fragments synthesized in E.coli and found no differences between the two (Hosszu...Collinge JBC 2004). In many cases, the subtype of PrP^{Sc} found in the brains of sCJD patients correlates with the identity of the amino acid at position 129 but the explanation for this is also unknown.

Although it has been clearly demonstrated that PrP^{Sc} is the infectious agent for prion diseases, it is not clear how the presence of PrP^{Sc} causes the cell death associated with these neurodegenerative diseases. It is likely that the presence of PrP^{Sc} on the cell surface affects intracellular production of PrP, which is required to generate the disease phenotype. One line of evidence that supports this idea stems from the observation that upon infection of an animal with a specific subtype of PrP^{Sc}, that animal then makes the

identical subtype of PrP^{Sc} (Hill et al. 2003). Presumably, in order to duplicate the glycosylation state of the infecting PrP^{Sc}, the extracellular PrP^{Sc} must somehow direct the cell that all nascent PrP chains should receive N-linked glycosylation only at specific residues. Another finding that supports the idea that the effect of extracellular PrP^{Sc} on intacellular PrP production is important for disease is the observation that PrP null mice infected with PrP^{Sc} never come down with disease. In addition, subclinical prion infections have been recently been recognized in animals who, upon exposure to PrP^{Sc}, generate high titres of infectious prions, but never develop clinical symptoms of disease (Hill& Collinge 2003 #2). This argues that the presence of high levels of PrP^{Sc} alone does not cause disease.

The mechanism by which extracellular PrP^{Sc} transmits a signal into the cell is unclear, but an intracellular form of PrP that affects disease has been identified. It is thought to be a transmembrane form of PrP designated ^{Ctm}PrP. This is one of three different topologic variants of PrP that are generated during co-translational translocation of PrP nascent chains at the endoplasmic reticulum (ER). PrP contains an N-terminal signal sequence that directs the nascent chain to the translocon shortly after translation is initiated. At the translocon the N-terminus of the signal sequence can be localized either to the cytosol or the ER lumen. PrP also contains a potential transmembrane (TM) domain. Localization of the N terminus to the cytosol followed by utilization of the TM domain generates ^{Ctm}PrP (see figure 17A). Localization of the N Terminus to the ER lumen generates ^{Ntm}PrP if the TM domain is integrated or ^{Sec}PrP if it is not (see figure 17A). ^{Sec}PrP is associated with the membrane through a GPI anchor at the C terminus. Both PrP^{C} and PrP^{Sc} are thought to be alternately folded conformations of ^{Sec}PrP.

Transgenic mice expressing high levels of PrP with ^{Ctm}PrP favoring mutations were found to come down with prion diseases spontaneously, suggesting that ^{Ctm}PrP might be the intracellular conformer responsible for neurodegeneration (Hegde et al. 1998a). Mice expressing less mutant PrP came down with disease much later, while mice expressing very low levels of the mutant protein never got sick (Hegde et al. 1999). In addition examination of the brains of hamsters infected with moPrP^{Sc} demonstrated increased levels of ^{Ctm}PrP (Hegde et al. 1999). Examination of tissue from human patients with familial GSS revealed high levels of ^{Ctm}PrP in addition to PrP^{Sc} (Hegde et al. 1998a). More evidence for the role of ^{Ctm}PrP came from examination of mice expressing the prion protein with mutations that prevent the formation of ^{Ctm}PrP. Like PrP null mice, these mice showed no symptoms of neurodegeneration upon infection with high levels of ^{Ctm}PrP (Hegde et al. 1999).

In the course of our investigation of the properties of the TM domain responsible for ^{Ctm}PrP integration (see chapter III) we found that mutations at position 129 can alter ^{Ctm}PrP biogenesis. Here we demonstrate that the levels of ^{Ctm}PrP made by M129 and V129 PrP are different both *in vivo* and *in vitro*. In addition, we found that the mutation D178N also has a subtle effect on ^{Ctm}PrP levels. These results support a new hypothesis to explain the correlation between the Met/Val polymorphism and disease.

Experimental Procedures

Plasmid Constructions

Constructs used in *in vitro* translation were derived from the pSP64 haPrP, moPrP and huPrP constructs published previously (Hegde et al. 1998a; Kim et al. 2002). Transfection vectors were in the pCDNA3.1 vector. All mutants were generated by directed mutagenesis.

In vitro Translation, Translocation, and Proteolysis

In vitro transcription and translation were performed as described previously (Chuck and Lingappa 1992). Translations were carried out at 34°C or 37°C for 30 minutes. Glycosylation was inhibited by 0.2mM tripeptide competitor (Chuck and Lingappa 1992). Microsomal membrane isolation has been described previously (Rutkowski et al. 2003). Samples were proteolized at 4°C for 45 minutes with 0.2-0.4 mg/mL proteinase K (PK). The protease was inactivated by incubation with 10mM PMSF for 5 min and boiling in 10 volumes of 0.1M Tris, 1% SDS.

Cell-Culture

CHO1 cells were cultured in Ham's F-12 complete medium (University of California San Francisco Cell Culture Facility) with 10% fetal bovine serum and antibiotics. Transfections were performed using lipofectamine Plus (Invitrogen). For each 35mm well, 3ug of DNA was incubated with 125 µL serum free media and 7.5 µL Plus reagent for 15 minutes. In parallel, 5 µL Lipofectamine was incubated with 125 µL

serum free media. The two solutions were then mixed and incubated for an additional 15 minutes at room temperature. This mixture was then added to the well containing 1 mL of fresh serum free media. After three hours the media was replaced with complete media. Prior to pulsing with ^{35}S Met, the cells were incubated for 30 minutes in Ham's F-12 Cys/Met free media with 10% dialyzed FCS and antibiotics. Where indicated 150 μM ALLN in DMSO was added to culture media. Cells were labeled with 30 μL ^{35}S EasyTag (Perkin Elmer) per well 24 hours after transfection for a duration of four hours.

Cell harvesting, lysis and proteolysis

After setting on ice for ten minutes cells were harvested by scraping and transferred to a 15 mL Falcon tube. They were then pelleted for 5 min at 3,000 rpm (2,060g) and 4°C in a Beckman GS6R centrifuge. The media was removed and the cells were resuspended in 600 μL of 4°C 10mM HEPES pH 7.5 and frozen. For homogenization the cells were thawed and passes 10 times through a 27 gauge needle. $1/10^{th}$ volume of 10x physiologic slat buffer (0.5M HEPES ph 7.5, 1M KoAc, 50 mM MgoAc) was then added. For proteolysis, samples were split into three 200 μL aliquots. One sample was treated with both 8 μL of 10 mg/mL proteinase K (pk) and 10 μL 20% Triton X-100, another was treated only with pk and the third was left untreated. All samples were incubated on ice at 4°C for 1 hour and then transferred to four volumes of boiling 1% SDS, 0.1M Tris pH 8.0.

Immunoprecipitation and PNgase F treatment

400 μL proteolized sample was added to 1.25 mL 1.5x Doc TXSWB (75mM KoAc, 75mM Tris Acetat pH 8.0, 0.75% deoxycholate, 1.5% triton X-100), 20 μL packed protein G beads and 2 μL 13A5 ascites. Samples were incubated on the rotator overnight at 4°C. The beads were then washed three times with 1x Doc TXSWB and one time with 10 mM Tris Acetate pH 8. The beads were then boiled to evaporate liquid, and boiled in 50 μL 0.5-1% SDS, 0.1M Tris acetate pH 8 to release the PrP from the beads.

For PNGase F treatment 50 μL of 2% BME, 0.1M Tris pH6.8 was added to each sample and the tubes were incubated for 15 minutes at 37 °C, then boiled for 2 min. Then 2.5 μL 20% Triton X-100 was added and the samples were mixed by vortexing. The samples were split into two 45 μL aliquots, to one of which 1 μL of PNGaseF was added. The samples were then left at least 6 hours at 37 °C.

Miscellaneous

15% tricine and 15% tris-glycine gels were used for SDS-PAGE. Autoradiographs were scanned using an Agfa Arcus II flatbed scanner and quantitated using NIH Image 1.63. Graphs with error bars represent the mean and standard deviation calculated from triplicate reactions.

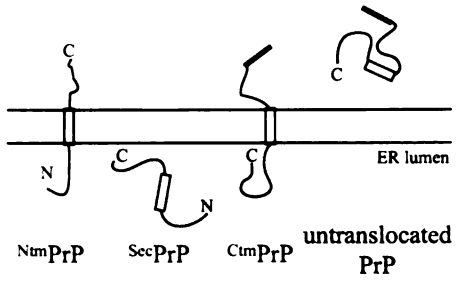
Results

The identity of the amino acid at the polymorphic residue 129 affects generation of ^{Ctm}PrP

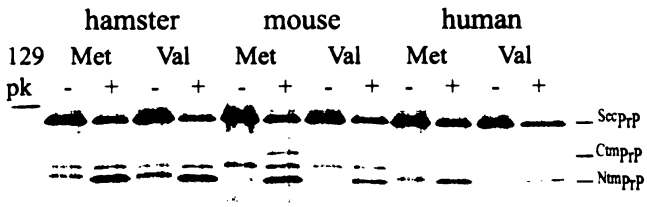
Our studies of the regions of PrP important for ^{Ctm}PrP biogenesis suggested that the identity of the amino acid at position 129 of the TM domain could alter ^{Ctm}PrP integration. Because the polymorphism at residue 129 impacts prion diseases we decided to compare the levels of ^{Ctm}PrP made by hamster M129 and V129 PrP constructs *in vitro*. The two constructs were transcribed and translated *in vitro*. To asses the topology of the resulting nascent chains the microsomal membranes were isolated, resuspended and incubated in the presence or absence of proteinase K. Following inactivation of the protease, ^{Ntm}PrP, ^{Ctm}PrP, and ^{Sec}PrP were separated by SDS PAGE on tricene gels. The fraction of chains synthesized in each topology was determined after quantitation of the autoradiogram. Figure 17B and C demonstrates that hamster M129PrP generates more ^{Ctm}PrP than V129PrP. This is the first demonstration that the Met/Val polymorphism affects nascent protein biosynthesis.

The differences we observe cannot be accounted for simply by different levels of incorporation of ^{35}S Met. There are nine Met residues outside the signal sequence of hamster PrP, not including position 129 so the difference between incorporation in Met129 PrP and Val129PrP should be slight. In addition, the calculation of % ^{Ctm}PrP is calculated as 100 x [^{Ctm}PrP/(^{Ntm}PrP + ^{Ctm}PrP + ^{Sec}PrP)]. Any difference in labeling between Met129 and Val129 constructs might decrease the overall labeling of nascent chains, but it would not alter the fraction of chains found in each topologic form.

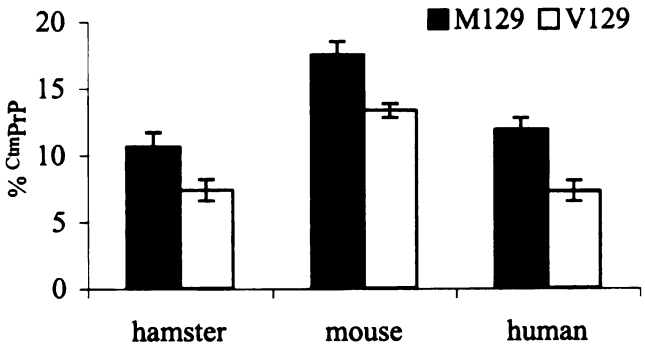
A



B



C



D

^{Ctm}PrP levels	Amino acid	Amino acid properties
Most ^{Ctm}PrP	Met, Leu	Large hydrophobic
Reduced ^{Ctm}PrP	Val, Ala, Ser, Cys, Tyr	Small hydrophobic and Polar, uncharged
Little ^{Ctm}PrP	Arg, Glu, Gly	Charged and Small polar uncharged

Figure 17: Methioine generates more ^{Ctm}PrP than Valine at position 129 of PrP.

(A) PrP can be made in different topologic forms at the ER membrane. Shown here are: ^{Ntm}PrP, which spans the membrane and has its N terminus in the lumen; ^{Ctm}PrP, which integrates in the opposite orientation; ^{Sec}PrP, which is fully translocated across the ER membrane; and untranslocated PrP which remains in the cytosol. (B) PrP from hamster, mouse and human, with either a Met or a Val at position 129 was transcribed and translated in vitro in the presence of canine microsomal membranes. After isolating the microsomes samples were incubated in the presence or absence of proteinase K (pk), and then separated by SDS PAGE. (C) The data in (B) was quantitated and the fraction of chains in each topologic form was calculated. The % ^{Ctm}PrP for each construct is graphed. (D) Residue 129 of human PrP was replaced with other amino acids and the resulting topology was assayed. This table shows the correlation between the amount of ^{Ctm}PrP generated and the properties of the amino acids tested.

We next wanted to verify that the difference we observed between Met129 PrP and Val129 PrP was also true for PrP from other organisms. This is especially significant because the only organism found to posses a polymorphism at 129 is humans (Schatzl et al. 1997). We generated mouse and human PrP constructs with either the Met or the Val at position 129. Upon analysis of the topologic distribution we found that Met129 PrP consistently generated more ^{Ctm}PrP than Val129 PrP (see fig. 17B and C).

Our previous results suggested that other substitutions at position 129 may be able to reduce ^{Ctm}PrP integration even further. Toward that end we generated the M129S construct in hamster PrP. Ser at position 129 almost eliminates ^{Ctm}PrP (data not shown). To look at further effects of amino acid substitutions we chose to generate several mutants in the human PrP construct. We substituted residue 129 with the following amino acids and assessed their effects on ^{Ctm}PrP biogenesis: Ala, Arg, Cys, Glu, Gly, Leu, Ser and Tyr (see figure 17D). Three of these substitutions, Arg, Glu, and Gly, essentially eliminate ^{Ctm}PrP production. Surprisingly, Ser, which significantly alters ^{Ctm}PrP levels of hamster PrP generates similar levels of ^{Ctm}PrP as does Val129 in human PrP. Ala, Cys, and Tyr also generated ^{Ctm}PrP levels similar to Val129. In contrast, M129L generates more ^{Ctm}PrP than M129. The ability of varied substitutions at 129 to dramatically alter ^{Ctm}PrP levels supports the conclusion that the identity of the amino acid at position 129 is important for ^{Ctm}PrP biogenesis.

The disease associated mutation D178N also affects generation of ^{Ctm}PrP *in vitro*

Many disease-associated mutations in PrP are consistently found in the presence of either Met or Val at position 129. For example, the mutation V180I, which causes

CJD is usually found on the same gene with a Met at position 129 (Gambetti et al. 2003). We wondered whether the reason these mutations were found with a single polymorphic form of the prion disease was due to an effect on ^{Ctm}PrP biogenesis. To test this hypothesis we generated the many mutant constructs, several of which are shown in figure 18 with either a Met or a Val at position 129. A previous study had already examined the effect of individual mutations on ^{Ctm}PrP levels, but no one had ever looked at the effect of the polymorphism in conjunction with these mutations. We consistently observe that regardless of the mutation, Met 129 constructs generate more ^{Ctm}PrP than Val129 constructs (see fig 18A). For most mutations, however, we could find no convincing effect of the mutations on ^{Ctm}PrP integration. Mutations previously known to affect ^{Ctm}PrP levels, like A117V and P105L, showed little or no difference in ^{Ctm}PrP levels between Met and Val constructs.

One mutation that did seem to have a small effect on ^{Ctm}PrP production was D178N. Met129, D178N appears to make more ^{Ctm}PrP than Met129 PrP. However, the levels for Val129, D178N are similar to those seen for Val129 PrP (see figure 18B). To try to determine whether the difference observed between wt and D178N PrP in the presence of Met at 129 is real we repeated the experiment at an elevated temperature (37°C), which causes levels of ^{Ctm}PrP to rise. The results, shown in figure 18C, demonstrate a subtle, but statistically significant, increase in the levels of ^{Ctm}PrP in the D178N mutant.

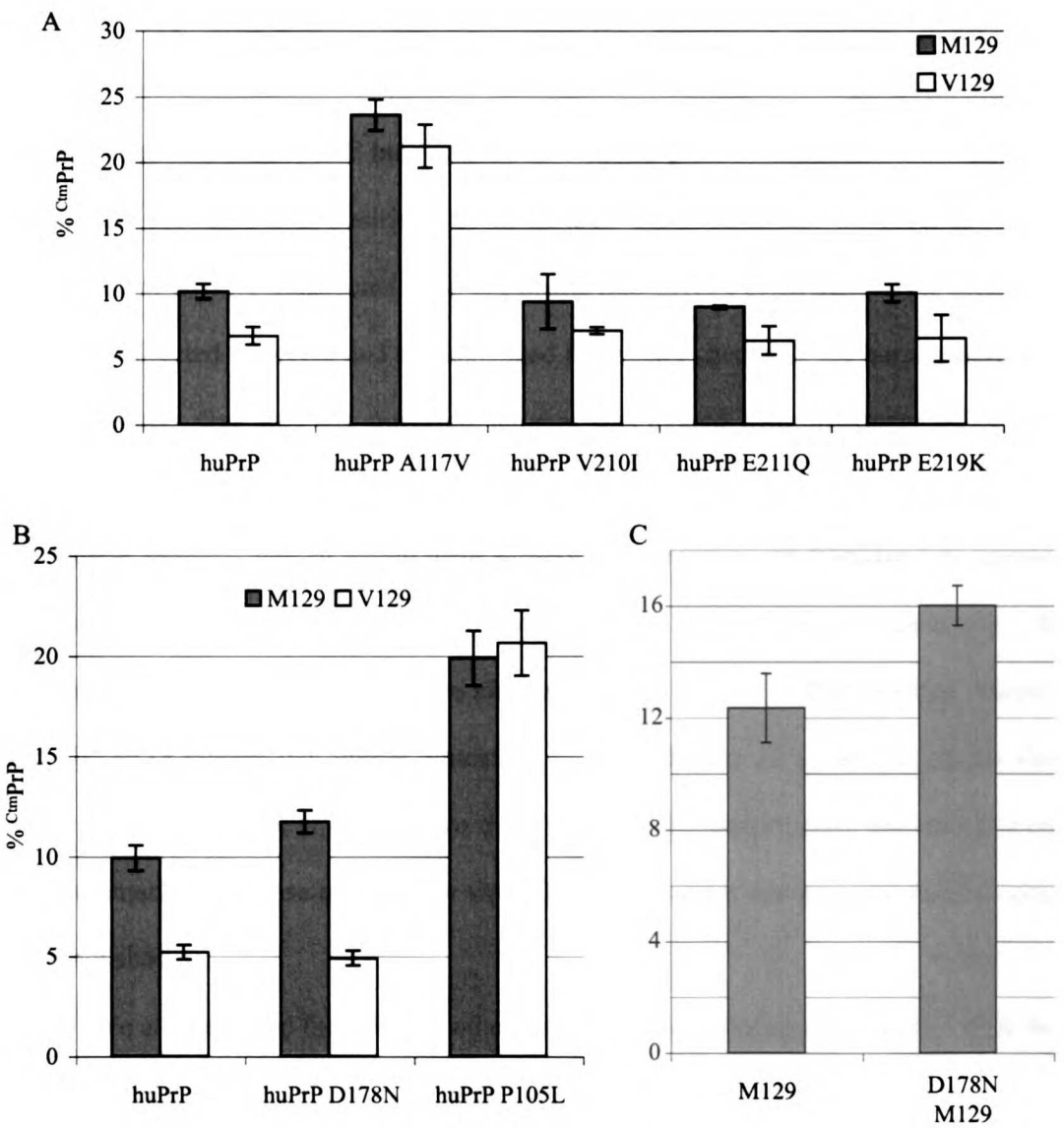


Figure 18: The effect of the 129 polymorphism on CtmPrP in the presence of disease causing mutations.

(A and B) The topology of constructs with the indicated disease-causing mutations was assayed as described in figure 17 in the presence of a Met or a Val at position 129. The % ^{Ctm}PrP is graphed. (C) Wild type PrP and D178N PrP with Met at position 129 were translated at 37°C and the topologic distribution was assayed. The average %^{Ctm}PrP from six samples is graphed.

The Met/Val polymorphism effects ^{Ctm}PrP biosynthesis *in vivo*

To test whether the differences we observed in ^{Ctm}PrP levels produced *in vitro* were representative of ^{Ctm}PrP biosynthesis *in vivo* we generated transfection constructs with either Met or Val at position 129 of hamster PrP. To assess the level of ^{Ctm}PrP production *in vivo* we transfected these constructs into CHO-1 cells, labeled the cells with ^{35}S Met, collected the cells, and homogenized them. We then split the homogenate into three aliquots. One was left untreated, another was treated with proteinase K, and the third was treated with both proteinase K and detergent. PrP was immunoprecipitated from these samples, half of which were subsequently treated with PNGase F to remove glycosylation. The samples were then separated by SDS PAGE on tricene gels. In parallel with M129 and V129 PrP we also analyzed A120L (a ^{Ctm}PrP favoring mutant) and G123P (a mutant that makes almost no ^{Ctm}PrP). Figure 19 (A and B) shows that M129 PrP makes more ^{Ctm}PrP *in vivo* than Val129 PrP. Surprisingly the results were more dramatic than those observed *in vitro*. The observation was also true in COS cells (data not shown).

We also assessed the effect of other amino acid substitutions at position 129 in the *in vivo* system. Again the results were similar to those observed *in vitro*, but more dramatic (see figure 19A and B). M129L hamster PrP makes significantly more ^{Ctm}PrP than M129 PrP, almost as much as A120L. In contrast, M129S and M129G make very little detectable ^{Ctm}PrP.

One major difference between assaying ^{Ctm}PrP *in vitro* versus *in vivo* is that *in vivo* there is a functioning protein degradation system. We wondered what affect proteasomal degradation had on the levels of ^{Ctm}PrP we observed. To test this we treated

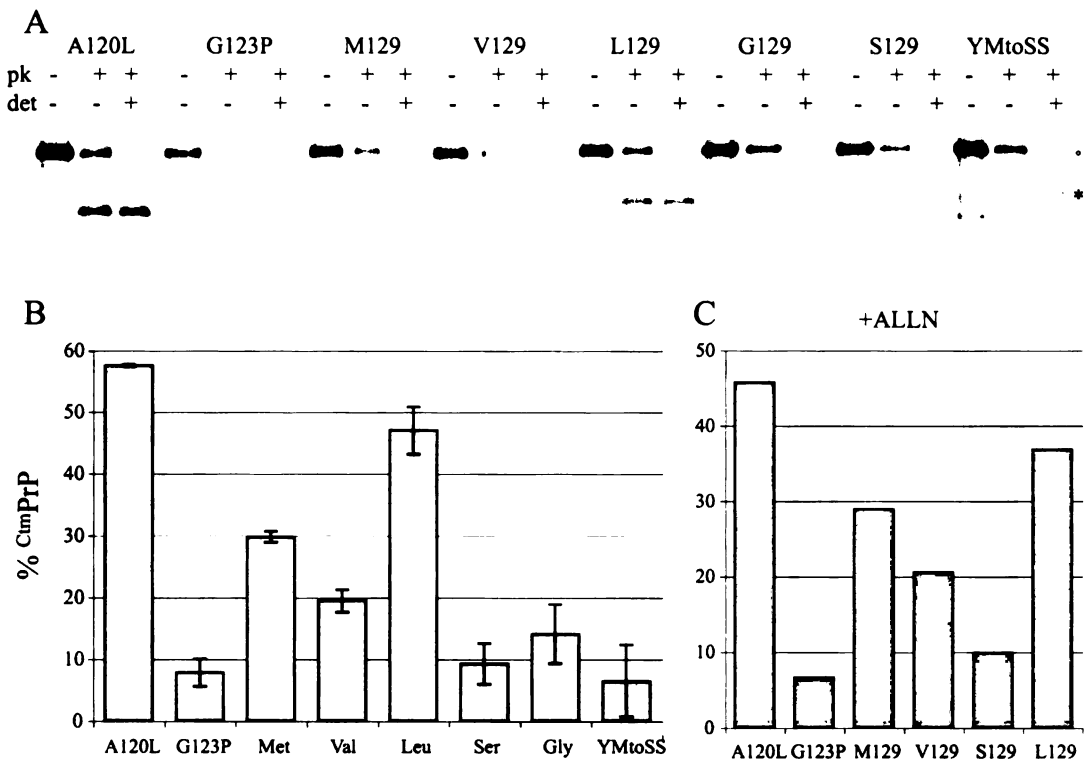


Figure 19: Residue 129 influence ^{Ctm}PrP generation in vivo.

(A) CHO-1 cells were transfected with PrP constructs containing the ^{Ctm}PrP favoring A120L mutation, the ^{Sec}PrP favoring G123P mutation, the indicated amino acids at position 129, or a mutation of residues 128 and 129 to Ser (YMtoSS). 24 hours after transfection samples were labeled with ^{35}S Met for four hours. Cells were then harvested and lysed. The samples were then split and incubated in the absence or presence of pk and detergent. Following inactivation of the protease, PrP was immunoprecipitated from the reactions and N-linked glycosylation was removed by treatment with PNGase F. The samples were then seperated by SDS PAGE. ^{Sec}PrP is indicated by the open circle and ^{Ctm}PrP is indicated by the asterisk. (B) The fraction of chains synthesized as ^{Sec}PrP and ^{Ctm}PrP from (A) was quantitated and the average %^{Ctm}PrP from duplicate samples is graphed. (C) Samples were prepared and quantitated as in (A) and (B) except cellswere treated with 150μM ALLN to inhibit the proteasome. Representative data is shown.

cells with 150 μM ALLN for 13 hours prior to harvesting (Yedidia et al. 2001). Upon treatment with ALLN the total levels of several proteins, including PrP, were increased as expected. In addition, we saw an accumulation of unglycosylated and singly glycosylated PrP. When we assayed the levels of ^{Ctm}PrP in treated cells we found that the difference between the levels of ^{Ctm}PrP generated by M129 and V129 expressing cells still remained (see figure 19C). These results suggest that proteasomal degradation does not contribute to the differences in levels of ^{Ctm}PrP observed *in vivo* for M129 and V129 PrP.

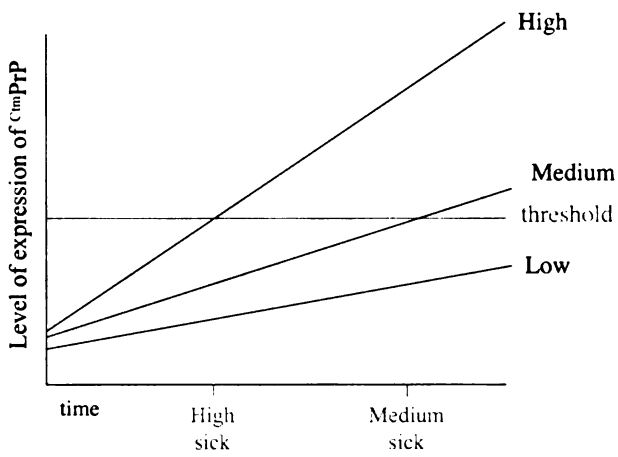
Discussion

In this study we examined the effect of the disease associated Met/Val polymorphism at position 129 in the prion protein on production of ^{Ctm}PrP. *In vitro* we found more ^{Ctm}PrP produced by M129 PrP than V129 PrP. The importance of the identity of the residue at 129 was reinforced by the observation that mutation of 129 to other amino acids could have both positive and negative effects on ^{Ctm}PrP biosynthesis. Although most of the mutants we examined showed no correlation between the Met/Val polymorphism and ^{Ctm}PrP, we did find that the D178N mutation in the presence of Met, but not Val, generates more ^{Ctm}PrP than wild type. *In vivo* we again observed more ^{Ctm}PrP made by M129 PrP than V129 PrP, in the presence and absence of proteasome inhibitors. Below we discuss the implications of these results for understanding the cell biology of prion diseases.

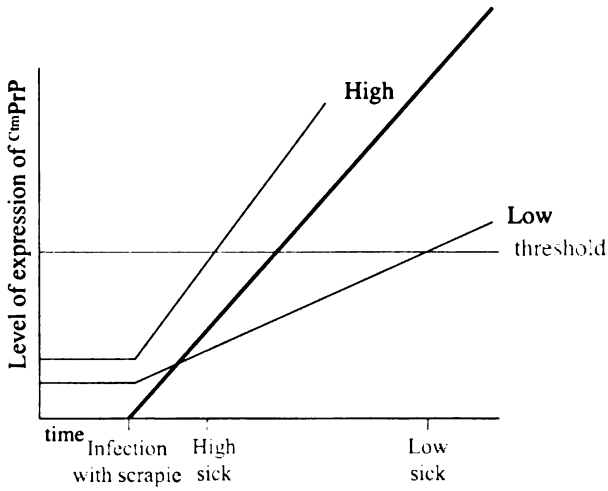
How can such a small change in ^{Ctm}PrP levels explain the dramatic correlation between the polymorphism and disease?

Although the differences in the levels of ^{Ctm}PrP generated by M129 PrP and V129 PrP are statistically significant, it seems unlikely that the observed differences in ^{Ctm}PrP levels could explain, for example, the dramatic observation that all of the people who have been diagnosed with vCJD are Met homozygotes (Collinge et al. 1996; Zeidler et al. 1997). Yet, it is already known that differences in ^{Ctm}PrP levels affect both the incubation period and time course of prion diseases. One model of how ^{Ctm}PrP causes disease is the threshold model diagramed in figure 20A and B. This assumes that a cell tolerates low

A


Level of expression of CmPrP
High
Medium
threshold
Low
time
High sick
Medium sick

B


Level of expression of CmPrP
High
Low
threshold
time
Infection with scrapie
High sick
Low sick

C

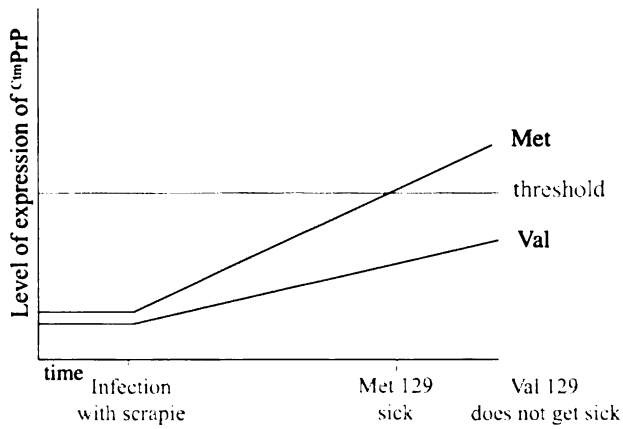

Level of expression of CmPrP
Met
threshold
Val
time
Infection with scrapie
Met 129 sick
Val 129 does not get sick

Figure 20: The threshold model explains why a small change in the propensity to make ^{Ctm}PrP can dramatically affect disease.

(A) Level of expression of ^{Ctm}PrP favoring mutants in transgenic mice has been shown to correlate with the timing of disease onset (Hegde et al. 1999). This can be explained by a threshold model for ^{Ctm}PrP mediated disease which predicts that disease only occurs when ^{Ctm}PrP levels exceed a threshold. Low level expressors never get sick because they never produce enough ^{Ctm}PrP to exceed the threshold. (B) Upon infection with PrP^{Sc}, high level expressors of a ^{Ctm}PrP favoring mutant develop disease much more quickly and which much less PrP^{Sc} accumulation that low level expressors (Hegde et al. 1999). According the threshold model this is because the low level expressors take much longer to exceed the CtmPrP threshold. (C) By extension, the threshold model predicts that the reason Met homozygotes develop vCJD is because they produce higher levels of ^{Ctm}PrP. Upon infection, the levels of ^{Ctm}PrP expressed by these people increases at a quicker rate than for heterozygotes or Val homozygotes. After several years Met homozygotes exceed the threshold and develop disease. It is not yet clear if Val carriers will remain well, or develop disease with a much longer incubation period.

levels of ^{Ctm}PrP expression. However, when ^{Ctm}PrP levels exceed a certain threshold, the ^{Ctm}PrP triggers cell death.

This model is supported by two sets of experiments with transgenic mice. First was a set of experiments done with transgenic mice expressing different levels of the ^{Ctm}PrP favoring mutant KHtoII. Mice expressing high levels of the mutant PrP develop disease at a young age and die quickly. Medium level expressors also develop disease but at a much older age. In contrast, no ^{Ctm}PrP is ever detectable in low level expressors and they never develop disease. These observations suggest that the high level expressors quickly exceed the ^{Ctm}PrP threshold, medium level expressors take much longer to exceed the threshold, and low level expressors never generate enough ^{Ctm}PrP to trigger neurodegeneration.

The second set of experiments that supports the threshold model looked at the time course of prion disease upon infection of transgenic mice expressing high and low levels of the KHtoII mutant PrP. Upon infection with PrP^{Sc}, high-level expressors developed disease very quickly and had relatively low titers of PrP^{Sc} in their brains upon death. In contrast, the low level expressors developed disease much later but had very high levels of PrP^{Sc} upon death. These results suggest that the low expressors got sick much later than the high level expressors, because it took longer for their ^{Ctm}PrP levels to exceed the threshold, in spite of the fact that their PrP^{Sc} levels continued to increase.

The threshold model can be used to explain how the subtle difference in ^{Ctm}PrP production between M129 PrP and V129 PrP could contribute to the observed difference in vCJD. Upon infection, PrP^{Sc} is thought to cause an increase in the amount of ^{Ctm}PrP made by the cell. Because V129 PrP has a lower propensity to make ^{Ctm}PrP, the amount

of ^{Ctm}PrP in heterozygotes and Val homozygotes never exceeds the disease threshold and those people never gets sick. This hypothesis seems more plausible when you consider that the predicted mean incubation time for vCJD is eleven years (Cooper and Bird 2003). Perhaps the incubation time for heterozygotes is 30 or 60 years.

The threshold model can also be used to explain how the subtle difference in ^{Ctm}PrP levels between M129 PrP and M129, D178N could contribute to FFI. The average age of disease presentation for FFI is 49 years. It seems likely that it takes many years for the levels of the mutant ^{Ctm}PrP to exceed the threshold. If the normal allele has a Val at position 129 the duration of the disease is between 2 and four times as long, perhaps because the levels of ^{Ctm}PrP accumulate more slowly.

But how would such a threshold be maintained? One possible mechanism is proteasomal degradation. To establish a threshold a certain fraction of ^{Ctm}PrP may be regularly degraded by the proteasome. This could explain why animals expressing low levels of the ^{Ctm}PrP favoring mutation show no detectable ^{Ctm}PrP. Upon mutation or infection, the level of ^{Ctm}PrP can increase and exceed the capacity of the degradation machinery. When enough ^{Ctm}PrP accumulates, it triggers neurodegenration.

The correlation we describe here between the Met/Val polymorphism and ^{Ctm}PrP may be very significant for understanding prion diseases. However, it does not explain why many disease mutations are only found with a specific polymorphic variant at position 129. It is also unclear why the D178N mutation causes FFI if the mutant allele has Met at position 129, but it causes fCJD if the there is a Val at 129. These mutations and the Met/Val polymorphism probably affect some other aspect of prion biology. Since the normal function of PrPC is not yet known, it is plausible that these are loss of

function mutations that affect normal PrPC. The results reported here also cannot explain the association of the polymorphism with the presence of specific sub types of PrP^{Sc} in sCJD. It is completely possible that in addition to affecting intracellular ^{Ctm}PrP levels, the polymorphism also affects the conversion of PrP^{C} to PrP^{Sc}.

Further studies are needed to fully appreciate the role of the Met/Val polymorphism in prion diseases. It will be interesting to see whether transgenic mice expressing the M129L and M129G mutations are less or more resistant to prion infections, respectively. In addition it will be very interesting to compare the levels of ^{Ctm}PrP and PrP^{Sc} in transgenic mice expressing human M129 or V129 PrP upon infection with PrP^{Sc} from BSE. This line of study has great potential to further elucidate the cell biology of prion diseases, which in turn could promote screening for compounds to treat or prevent disease.

Chapter IV

Conclusions, Perspectives, and Future Directions

Cell biology of membrane protein biogenesis

We are just beginning to understand how protein biosynthesis is regulated at the translocon. Much of what we have learned to date has come from the study of fewer than ten model membrane proteins. From this research there have already been conflicting results and discoveries that different proteins require proteins like TRAM and TRAP to different extents. As research extends to more interesting and more challenging substrates, even greater variety and regulation will be uncovered. A few of the exciting questions on the horizon are discussed below.

The research presented here describes sever different ways in which integration of the prion protein can be regulated. The ability to manipulate integration suggests that translocation can be tailored for individual proteins. Yet, few proteins have been found to demonstrate the topologic heterogeneity of PrP. That may be due in part to the limited number of substrates that have been examined, and the common assumption that most proteins are only synthesized in one form. Because multiple topologies are not expected, they may often go unrecognized. Regulation of integration could affect the boundaries of the transmembrane domain, something we don't currently have the tools to detect, which could have major implications for protein function. Sometime in the future perhaps it will be possible to develop a high throughput mechanism to screen for proteins that are synthesized in multiple forms.

One of the limitations of the in vitro cell free reconstitution system is that we generally use components from a few select tissues to study translocation. This has been due in part to the observation that preparations from specific tissues perform more robustly in the in vitro system. It is very plausible that tissue specific mechanisms are

utilized to regulate translocation and integration that cannot be seen in our system. For example, few cytosolic components have been implicated in translocational regulation. This may be because the reticulocytes, which provide the cytosol for the reactions, do not secrete many proteins, and therefore do not contain cytosolic factors involved in translocation.

It is still unclear what proteins are associated with the translocon and when. Several core components, such as Sec61, have been identified, but there are also a great number of proteins that have been found by crosslinking studies to be in proximity to the nascent chain at the translocon, that have yet to be identified (Falcone et al. 1999; Hegde and Lingappa 1996). We tried to use mass spectrometry techniques to identify crosslinked proteins, but the two major obstacles have been the purity and quantity of material needed. Mass spectrometry sequencing is often performed on proteins eluted from SDS polyacrylamide gels, but in order to isolate a band from a gel it is currently necessary to visualize the protein, which requires nanogram quantities of protein. Generally a crosslinking efficiency of five percent is considered high (Gorlich et al. 1991). Use of radioactive material allows us to visualize very small amounts of crosslinked material. To try to recover enough material we successfully scaled up our reaction volume as much as 100 fold, but a large proportion of the crosslinked complex was lost during subsequent purification steps. Hopefully in the near future new techniques will be developed that can be used to isolate and prepare small quantities of proteins for sequencing analysis. Knowing the identity of the many proteins that have been found near the translocon may reveal new mechanisms of regulation of translocation and integration.

The study of multi-spanning membrane proteins has been limited to date due to the size and complexity of these proteins. One challenge has been that *in vitro* translation efficiency decreases as the length of the nascent chain increases. Because of this, and the complexity multiple membrane spanning domains, much of the research of multi spanning membrane proteins has been with truncated proteins and chimeras. If synthesis of nascent chain in vitro is not feasible, perhaps new methods need to be developed to study integration in vivo. Many multi-spanning membrane proteins have been implicated in disease, so understanding how their biosynthesis is regulated may have a significant impact on human health.

Few membrane proteins function in isolation. Many form complexes with soluble proteins and/or other membrane proteins. Yet, little research has been done to date to look when and how these complexes form. Perhaps there is a "waiting room" outside the translocon where proteins are chaperoned until other members of their protein complex retrieve them. Like many other processes in the cell, the integration and biosynthesis of some membrane proteins may be independent, while others are only properly made in the presence of complex forming partners.

The translocon contains several multi-spanning membrane proteins, yet while the identity and function of translocon components has been studied for several years, little has been done to study how the translocon proteins are synthesized and integrated. Sec61 α has been found to target to the ER in an SRP dependent manner and to associate with Sec61 α, β and TRAM during integration (Knight and High 1998). But it is not clear whether nascent translocon proteins become a part of the translocon during integration at the ER. Perhaps they exit the translocon laterally and form new translocons or join other

translocon complexes? Recent research indicates that translocon proteins do not oligomerize de novo to translocate each nascent chain, but remain in an oligomeric state when not engaged by the ribosome (Snapp et al. 2004). This observation makes it seem more likely that incorporation of a nascent translocon protein would need a facilitated mechanism for assembly, either inside or outside the translocon.

^{Ntm}PrP *in vivo*: a new window into translocational regulation.

Previous studies of PrP biogenesis have focused on biosynthesis of ^{Ctm}PrP, in large part because of its association with disease. Although ^{Ctm}PrP and SecPrP have both been found *in vivo* little work has been done to look for ^{Ntm}PrP *in vivo*. Recently, however, research has suggested that ^{Ntm}PrP may be present in human platelets. The first evidence that PrP^{C} is present in platelets simply demonstrated that PrP is present on the surface of platelets and that upon activation, PrP reactive to the 3F4 antibody was released from the platelets (Perini et al. 1996a; Perini et al. 1996b). Because the glycosylation sites of PrP are C terminal to the TM domain, ^{Ntm}PrP does not get glycosylated. In theory this property can be used to differentiate SecPrP and ^{Ctm}PrP from ^{Ntm}PrP, however, it is unclear from those studies whether the released chains are sensitive to PNGase F treatment.

The first results suggesting that ^{Ntm}PrP specifically might be present in platelets looked not at the PrP that was released upon activation, but the PrP that was bound to the surface (Holada et al. 1998). By quantitating the amount of 3F4 reactive PrP on the surface of resting and activated platelets it was found that the amount of PrP increases greater than two-fold upon activation. The time course for the up regulation of PrP on the

surface of activated platelets appears to be very rapid (comparable to P-selectin), followed by a slow decrease in the amount of PrP on the surface, which may be related to active transport of the PrP back into the cells. This quick cellular response suggests that there must be an intrecellular reserve of PrP. To test this, platelets were fractionated and western blotted with 3F4. Platelet activation causes a decrease in the PrP in the organellar fractions and an increase in the PrP in the membrane fraction. In addition, although PIPLC did remove a control protein from the surface, it did not remove PrP. This result suggests that an integral membrane form of PrP (^{Ntm}PrP) is present on the cell surface. SecPrP, which is anchored to the membrane by a GPI anchor, would be released by PIPLC treatmen.

Additional evidence has been found that supports the hypothesis that ^{Ntm}PrP is present in platelets and other blood cells (Barclay et al. 1999). PrP was detected on the surface of platelets using 3F4 or an N terminus directed antibody, but not using a C terminus directed antibody. The strongest evidence in support of the presence of ^{Ntm}PrP in platelets came from the study of platelets from patients with paroxysmal nocturnal hemoglbinuria (PNH). This disease is caused by failure of GPI anchoring, which means that no SecPrP should be found in cells from these patients. Upon activation the PNH platelets demonstrated an increase in PrP on the cell surface that was immunoreactive only to antibodies at the N-terminus. Because only integral membrane proteins could be present on the surface of PNH patient platelets, this strongly supports the idea that upon activation ^{Ntm}PrP, not ^{Ctm}PrP, is up regulated at the cell surface.

The presence of ^{Ntm}PrP in platelets could provide an exciting new arena in which to study topology determination. While other cells generate predominantly SecPrP, how

do platelets up regulate ^{Ntm}PrP production? Are there factors at the translocon, in the ER lumen, or in the sytosol that facilitate increased integration of ^{Ntm}PrP? Reconstitution of translocation using platelet microsomal membranes could be used to try to study this question. It is also unclear what role ^{Ntm}PrP plays in platelets? Is ^{Ntm}PrP present in other tissues? One challenge to biochemically identifying ^{Ntm}PrP is that unlike ^{Ctm}PrP, which is protease resistant, opposite fragments of intracellular and extracellular ^{Ntm}PrP are digested by protease. To identify ^{Ntm}PrP one must look either at the cell surface, or at intracellular compartments. The inability of ^{Ntm}PrP to be glycosylated could also be helpful in demonstrating that it is present in vivo. If ^{Ctm}PrP is not made in these cells, then studies focused on differentiating integrated and GPI anchored PrP could be employed.

Conclusion

There is much we do not even realize that we do not know about membrane protein biogenesis. New questions will emerge as more is uncovered about this process. To really understand cell biology we need to understand how many different factors contribute to regulating diversity. This requires both understanding how the minimal machinery operates, and how other elements in combination with the minimal machinery can yield different outcomes. The challenge now ahead for understanding regulation of integration is to learn how different proteins present at the translocon alter integration. Many cellular processes involve integral membrane proteins. Discovering how their biosynthesis is regulated will have implications for understanding diverse areas of biology.

References

Andrews D W, Perara, E, Lesser, C and Lingappa, V R (1988) Sequences Beyond the Cleavage Site Influence Signal Peptide Function. Journal of Biological Chemistry 263: 15791-8.

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: 7761-9.

Barclay G R, Hope, J, Birkett, C R and Turner, M L (1999) Distribution of cell-associated prion protein in normal adult blood determined by flow cytometry. Br J Haematol 107: 804-14.

Beckmann R, Spahn, C M, Eswar, N, Helmers, J, Penczek, P A, et al. (2001) Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell 107: 361-72.

Beguin 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: 24921-31.

Blobel G (1980) Intracellular protein topogenesis. Proc Natl Acad Sci U S A 77: 1496-500.

Blobel G and Dobberstein, B (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol 67: 835-51.

Booth P J and Curran, A R (1999) Membrane protein folding. Curr Opin Struct Biol 9: 115-21.

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: 379-89.

Chuck S L and Lingappa, V R (1992) Pause Transfer - a Topogenic Sequence in Apolipoprotein-B Mediates Stopping and Restarting of Translocation. Cell 68: 9-21.

Collinge J, Beck, J, Campbell, T, Estibeiro, K and Will, R G (1996) Prion protein gene analysis in new variant cases of Creutzfeldt-Jakob disease. Lancet 348: 56.

Collinge J, Palmer, M S and Dryden, A J (1991) Genetic predisposition to iatrogenic Creutzfeldt-Jakob disease. Lancet 337: 1441-2.

Cooper J D and Bird, S M (2003) Predicting incidence of variant Creutzfeldt-Jakob disease from UK dietary exposure to bovine spongiform encephalopathy for the 1940 to 1969 and post-1969 birth cohorts. Int J Epidemiol 32: 784-91.

Cserzo M, Wallin, E, Simon, I, von Heijne, G and Elofsson, A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng 10: 673-6.

Deber C M, Wang, C, Liu, L P, Prior, A S, Agrawal, S, et al. (2001) TM Finder: a prediction program for transmembrane protein segments using a combination of hydrophobicity and nonpolar phase helicity scales. Protein Sci 10: 212-9.

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: 369-78.

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: 3609-16.

Engelman D M and Steitz, T A (1981) The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell 23: 411-22.

Evans E A, Gilmore, R and Blobel, G (1986) Purification of microsomal signal peptidase as a complex. Proc Natl Acad Sci U S A 83: 581-5.

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: 33661-70.

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: 7856-64.

Fons R D, Bogert, B A and Hegde, R S (2003) Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. J Cell Biol 160: 529-39.

Fujiki Y, Rachubinski, R A and Lazarow, P B (1984) Synthesis of a major integral membrane polypeptide of rat liver peroxisomes on free polysomes. Proc Natl Acad Sci U S A 81: 7127-31.

Fulga T A, Sinning, I, Dobberstein, B and Pool, M R (2001) SRbeta coordinates signal sequence release from SRP with ribosome binding to the translocon. Embo J 20: 2338-47.

Gambetti P, Kong, Q, Zou, W, Parchi, P and Chen, S G (2003) Sporadic and familial CJD: classification and characterisation. Br Med Bull 66: 213-39.

Gilmore R, Blobel, G and Walter, P (1982a) Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. J Cell Biol 95: 463-9.

Gilmore R, Walter, P and Blobel, G (1982b) Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. J Cell Biol 95: 470-7.

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: 257-66.

Goder V and Spiess, M (2003) Molecular mechanism of signal sequence orientation in the endoplasmic reticulum. Embo J 22: 3645-53.

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, Kurzchalia, T V, Wiedmann, M and Rapoport, T A (1991) Probing the molecular environment of translocating polypeptide chains by cross-linking. Methods Cell Biol 34: 241-62.

Gorlich D and Rapoport, T A (1993) Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. Cell 75: 615-30.

Haigh N G and Johnson, A E (2002) A new role for BiP: closing the aqueous translocon pore during protein integration into the ER membrane. J Cell Biol 156: 261-70.

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 A during cotranslational protein translocation at the ER membrane. Cell 89: 535-44.

Hamman B D, Hendershot, L M and Johnson, A E (1998) BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation. Cell 92: 747-58.

Hanein D, Matlack, K E, Jungnickel, B, Plath, K, Kalies, K U, et al. (1996) Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. Cell 87: 721-32.

Hay B, Barry, R A, Lieberburg, I, Prusiner, S B and Lingappa, V R (1987a) Biogenesis and Transmembrane Orientation of the Cellular Isoform of the Scrapie Prion Protein. Mol Cell Biol 7: 914-20.

Hay B, Prusiner, S B and Lingappa, V R (1987b) Evidence for a Secretory Form of the Cellular Prion Protein. Biochemistry-Us 26: 8110-5.

Hegde R S and Lingappa, V R (1996) Sequence-specific alteration of the ribosome-membrane junction exposes nascent secretory proteins to the cytosol. Cell 85: 217-28.

Hegde R S and Lingappa, V R (1999) Regulation of protein biogenesis at the endoplasmic reticulum membrane. Trends Cell Biol 9: 132-7.

Hegde R S, Mastrianni, J A, Scott, M R, DeFea, K A, Tremblay, P, et al. (1998a) A transmembrane form of the prion protein in neurodegenerative disease. Science 279: 827-34.

Hegde R S, Tremblay, P, Groth, D, DeArmond, S J, Prusiner, S B, et al. (1999) Transmissible and genetic prion diseases share a common pathway of neurodegeneration. Nature 402: 822-6.

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: 621-31.

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: 233-44.

Heinrich S U and Rapoport, T A (2003) Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. Embo J 22: 3654-63.

High S, Andersen, S S, Gorlich, D, Hartmann, E, Prehn, S, et al. (1993) Sec61p is adjacent to nascent type I and type II signal-anchor proteins during their membrane insertion. J Cell Biol 121: 743-50.

Hill A F, Butterworth, R J, Joiner, S, Jackson, G, Rossor, M N, et al. (1999) Investigation of variant Creutzfeldt-Jakob disease and other human prion diseases with tonsil biopsy samples. Lancet 353: 183-9.

Hill A F, Desbruslais, M, Joiner, S, Sidle, K C, Gowland, I, et al. (1997) The same prion strain causes vCJD and BSE. Nature 389: 448-50, 526.

Hill A F, Joiner, S, Wadsworth, J D, Sidle, K C, Bell, J E, et al. (2003) Molecular classification of sporadic Creutzfeldt-Jakob disease. Brain 126: 1333-46.

Holada K, Mondoro, T H, Muller, J and Vostal, J G (1998) Increased expression of phosphatidylinositol-specific phospholipase C resistant prion proteins on the surface of activated platelets. Br J Haematol 103: 276-82.

Jungnickel B and Rapoport, T A (1995) A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. Cell 82: 261-70.

Keenan R J, Freymann, D M, Stroud, R M and Walter, P (2001) The signal recognition particle. Annu Rev Biochem 70: 755-75.

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 S J and Hegde, R S (2002) Cotranslational partitioning of nascent prion protein into multiple populations at the translocation channel. Mol Biol Cell 13: 3775-86.

Kim S J, Mitra, D, Salerno, J R and Hegde, R S (2002) Signal sequences control gating of the protein translocation channel in a substrate-specific manner. Dev Cell 2: 207-17.

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: 26132-40.

Knight B C and High, S (1998) Membrane integration of Sec61alpha: a core component of the endoplasmic reticulum translocation complex. Biochem J 331 (Pt 1): 161-7.

Kyte J and Doolittle, R F (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-32.

Li Y, Bergeron, J J, Luo, L, Ou, W J, Thomas, D Y, et al. (1996) Effects of inefficient cleavage of the signal sequence of HIV-1 gp 120 on its association with calnexin, folding, and intracellular transport. Proc Natl Acad Sci U S A 93: 9606-11.

Lin J and Addison, R (1995) A novel integration signal that is composed of two transmembrane segments is required to integrate the Neurospora plasma membrane H(+)-ATPase into microsomes. J Biol Chem 270: 6935-41.

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: 226-9.

Martoglio B and Dobberstein, B (1998) Signal sequences: more than just greasy peptides. Trends Cell Biol 8: 410-5.

Martoglio B, Hofmann, M W, Brunner, J and Dobberstein, B (1995) The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. Cell 81: 207-14.

McCormick P J, Miao, Y, Shao, Y, Lin, J and Johnson, A E (2003) Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. Mol Cell 12: 329-41.

Mead S, Stumpf, M P, Whitfield, J, Beck, J A, Poulter, M, et al. (2003) Balancing selection at the prion protein gene consistent with prehistoric kurulike epidemics. Science 300: 640-3.

Meyer D I and Dobberstein, B (1980a) Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum. J Cell Biol 87: 503-8.

Meyer D I and Dobberstein, B (1980b) A membrane component essential for vectorial translocation of nascent proteins across the endoplasmic reticulum: requirements for its extraction and reassociation with the membrane. J Cell Biol 87: 498-502.

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: 711-9.

Morgan D G, Menetret, J F, Neuhof, A, Rapoport, T A and Akey, C W (2002) Structure of the mammalian ribosome-channel complex at 17A resolution. J Mol Biol 324: 871-86.

Mothes W, Heinrich, S U, Graf, R, Nilsson, I, von Heijne, G, et al. (1997) Molecular mechanism of membrane protein integration into the endoplasmic reticulum. Cell 89: 523-33.

Mothes W, Jungnickel, B, Brunner, J and Rapoport, T A (1998) Signal sequence recognition in cotranslational translocation by protein components of the endoplasmic reticulum membrane. J Cell Biol 142: 355-64.

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: 29743-8.

Ota K, Sakaguchi, M, von Heijne, G, Hamasaki, N and Mihara, K (1998) Forced transmembrane orientation of hydrophilic polypeptide segments in multispanning membrane proteins. Mol Cell 2: 495-503.

Perara E, Rothman, R E and Lingappa, V R (1986) Uncoupling Translocation from Translation - Implications for Transport of Proteins across Membranes. Science 232: 348-52.

Perini F, Frangione, B and Prelli, F (1996a) Prion protein released by platelets. Lancet 347: 1635-6.

Perini F, Vidal, R, Ghetti, B, Tagliavini, F, Frangione, B, et al. (1996b) PrP27-30 is a normal soluble prion protein fragment released by human platelets. Biochem Biophys Res Commun 223: 572-7.

Popot J L and Engelman, D M (1990) Membrane protein folding and oligomerization: the two-stage model. Biochemistry-Us 29: 4031-7.

Popot J L and Engelman, D M (2000) Helical membrane protein folding, stability, and evolution. Annu Rev Biochem 69: 881-922.

Prusiner S B (1982) Novel proteinaceous infectious particles cause scrapie. Science 216: 136-44.

Prusiner S B (1998) Prions. Proc Natl Acad Sci U S A 95: 13363-83.

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. Journal of Biological Chemistry 263: 10470-80.

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 U S A 98: 7823-8.

Rutkowski D T, Ott, C M, Polansky, J R and Lingappa, V R (2003) Signal sequences initiate the pathway of maturation in the endoplasmic reticulum lumen. J Biol Chem 278: 30365-72.

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: 438-42.

Schatzl H M, Wopfner, F, Gilch, S, von Brunn, A and Jager, G (1997) Is codon 129 of prion protein polymorphic in human beings but not in animals? Lancet 349: 1603-4.

Singer S J, Maher, P A and Yaffe, M P (1987) On the transfer of integral proteins into membranes. Proc Natl Acad Sci U S A 84: 1960-4.

Skach W R, Calayag, M C and Lingappa, V R (1993) Evidence for an alternate model of human P-glycoprotein structure and biogenesis. J Biol Chem 268: 6903-8.

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 2 Internal Sequences. Journal of Biological Chemistry 268: 23552-61.

Snapp E L, Hegde, R S, Francolini, M, Lombardo, F, Colombo, S, et al. (2003) Formation of stacked ER cisternae by low affinity protein interactions. J Cell Biol 163: 257-69.

Snapp E L, Reinhart, G A, Bogert, B A, Lippincott-Schwartz, J and Hegde, R S (2004) The organization of engaged and quiescent translocons in the endoplasmic reticulum of mammalian cells. J Cell Biol 164: 997-1007.

Sonnhammer E 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: 175-82.

Thrift 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: 809-21.

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.

Van den Berg B, Clemons, W M, Jr., Collinson, I, Modis, Y, Hartmann, E, et al. (2004) X-ray structure of a protein-conducting channel. Nature 427: 36-44.

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: 487-94.

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: 226-35.

Yedidia Y, Horonchik, L, Tzaban, S, Yanai, A and Taraboulos, A (2001) Proteasomes and ubiquitin are involved in the turnover of the wild-type prion protein. Embo J 20: 5383-91.

Yost C S, Hedgpeth, J and Lingappa, V R (1983) A Stop Transfer Sequence Confers Predictable Transmembrane Orientation to a Previously Secreted Protein in Cell-Free Systems. Cell 34: 759-66.

Yost C S, Lopez, C D, Prusiner, S B, Myers, R M and Lingappa, V R (1990) Nonhydrophobic Extracytoplasmic Determinant of Stop Transfer in the Prion Protein. Nature 343: 669-72.

Zeidler M, Stewart, G, Cousens, S N, Estibeiro, K and Will, R G (1997) Codon 129 genotype and new variant CJD. Lancet 350: 668.