Title
Characterization of the molecular determinants of prion protein topology

Author
DeFea, Kathryn Anne

Publication Date
1994

Peer reviewed|Thesis/dissertation

Characterization of the Molecular Determinants of Prion Protein Topology

by

Kathryn Anne De Fea

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Endocrinology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Characterization of the Molecular Determinants of Prion Protein Topology

Kathryn Anne De Fea

Thesis Advisor: Vishwanath Lingappa

Abstract

Recently, a number of proteins containing topogenic sequences that direct unconventional mechanisms of translocation across the Endoplasmic Reticulum (ER) membrane have been identified, suggesting a role for regulatory events occurring at the ER membrane in the expression of a functional protein. The prion protein (PrP), a brain glycoprotein associated with the pathogenesis of a number of neurodegenerative disorders is one example. In brain and transfected cell lines, PrP is observed as a fully translocated molecule, tethered to the plasma membrane by a glycolipid anchor. Pulse/chase studies in the *Xenopus* Oocyte described here, reveal that this fully translocated topology is achieved via a transient transmembrane intermediate that spans the membrane twice. Information directing this transmembrane topology was found to reside in a single domain (codons 86-143 of hamster PrP), consisting of the first membrane spanning domain (TM1) and the hydrophilic sequence preceding it (STE). Small deletions and point mutations of STE-TM1 that either abolish the transmembrane intermediate or that result in an accumulation of transmembrane chains have been identified. Assays using alkaline perturbants and reversible cross-linking agents demonstrate such mutations can result in transmembrane molecules that are poised in an aqueous pore interacting with channel proteins Similarities between the sequence and function of STE-TM1 and another recently identified topogenic sequence, the pause-transfer, are suggestive of a class of functionally divergent proteins utilizing a common mechanism to achieve their final topologies. Finally, studies indicating that conformational changes in a region of PrP overlapping STE-TM1 may be involved in the pathogenesis of prion related diseases suggest a role for the translocational mechanism described here in either the normal or disease associated function of PrP.

Acknowledgments:

I want to thank all of the people who made completion of this work possible. Of course, much thanks and appreciation is due to my thesis advisor Vishu Lingappa, whose irrepressible enthusiasm and creativity first drew me to his lab, and whose patience and encouragement throughout periods of frustratingly slow progress kept me from giving up. Special thanks are also due to Spencer Yost, who helped out in some of the initial stages of what grew to be the core of my thesis work, and who taught me site-directed mutagenesis (as well as a supply of morbid jokes). Both my enjoyment of science, coaching for my Oral Exams and insights into alternative approaches to my project were provided by all of the members of the Lingappa lab-both former and present. In particular, numerous brainstorming sessions with Dori on the subject of PrP topology and ideas and suggestions from Bill Skach were indispensable to my work. Special thanks to Bill for reading over parts of my thesis! Thanks also to Manu for agreeing to bring the transgenic mouse project to its fruition. I wish to thank my thesis committee members Robert Fletterick and Bill Mobley for showing the utmost confidence in my abilities as a scientist. Finally, I owe gratitude to my closest friends Jeanmarie, Clara, Theresa, Michelle and my parents, Betty and Bob De Fea, for their support during my graduate training.

Characterization of the Molecular Determinants of Prion Protein Topology

Kathryn Anne De Fea

Abstract

Recently, a number of proteins containing topogenic sequences that direct unconventional mechanisms of translocation across the Endoplasmic Reticulum (ER) membrane have been identified, suggesting a role for regulatory events occurring at the ER membrane in the expression of a functional protein. The prion protein (PrP), a brain glycoprotein associated with the pathogenesis of a number of neurodegenerative disorders is one example. In brain and transfected cell lines, PrP is observed as a fully translocated molecule, tethered to the plasma membrane by a glycolipid anchor. Pulse/chase studies in the *Xenopus* Oocyte described here, reveal that this fully translocated topology is achieved via a transient transmembrane intemediate that spans the membrane twice. Information directing this transmembrane topology was found to reside in a single domain (codons 86-143 of hamster PrP), consisting of the first membrane spanning domain (TM1) and the hydrophilic sequence preceding it (STE). Small deletions and point mutations of STE-TM1 that either abolish the transmembrane intermediate or that result in an accumulation of transmembrane chains have been identified. Assays using alkaline perturbants and reversible cross-linking agents demonstrate such mutations can result in transmembrane molecules that are poised in an aqueous pore interacting with channel proteins Similarities between the sequence and function of STE-TM1 and another recently identified topogenic sequence, the pause-transfer, are suggestive of a class of functionally divergent proteins utilizing a common mechanism to achieve their final topologies. Finally, studies indicating that conformational changes in a region of PrP overlapping STE-TM1 may be involved in the pathogenesis of prion related diseases suggest a role for the translocational mechanism described here in either the normal or disease associated function of PrP.

Figures and Tables: **page**

Table of Contents:

Chapter 1: INTRODUCTION

Most nascent proteins destined for the cell surface or extracellular environment, as well as some intracellular compartments, are initially targeted to the Endoplasmic Reticulum (ER). Such proteins contain discrete sequences within the nascent polypeptide that interact with specific cytosolic and ER receptors, directing targeting to, translocation across, and, in the case of membrane proteins, integration into the ER membrane. Because their action defines the final topology of a given protein, these sequences are referred to as topogenic sequences. Through these interactions, most proteins assume their final topology at the ER (1). Subsequent trafficking events occur through vesicular budding into a series of membrane bound organelles (e.g. the Golgi Network, lyzosomes or microsomes) and ultimately, in the case of secretory and membrane proteins, to the plasma membrane (2). Thus, the only hydrophobic barrier most proteins must cross as they exit the cell is at the Endoplasmic Reticulum, and regulatory events occurring at this level are of utmost importance to their final topology .

I. Topogenic Sequences

For many years, there has been considerable debate as to whether protein translocation occurs through the lipid bilayer, or whether a proteinaceous pore exists in the ER membrane capable of conducting macromolecules as large as proteins across it. Increasing evidence for the latter hypothesis has emerged recently. First, an ion conducting channel in the ER membrane has been identified (3). Second, several of the protein components of this channel, collectively referred to as the translocon, have been isolated and cloned (4-9). Third, a number of secretory mutants have been identified that lack one or more of these protein components (10-12). The translocation machinery, much of which comprises this channel, consists of a variety of cytosolic and ER membrane protein factors

(6). Interaction of a nascent chain with translocation machinery depends on the topogenic sequences within the polypeptide.

Several classes of topogenic sequences have been identified and functionally characterized. A large majority of both secretory and transmembrane proteins contain cleavable amino (N-) terminal signal sequences (S) that target them to the ER (13, 14). Conventional transmembrane proteins contain additional topogenic sequences known as stop transfer sequences (ST) that direct the chain to stop translocation and integrate into the membrane, with subsequent domains residing in the cytosol (15-17). Some transmembrane proteins do not contain N-terminal or cleavable S sequences, and the membrane targeting signal also serves as a transmembrane domain. Such sequences are known as signal anchor (SA) sequences and are found anywhere within the polypeptide chain (14, 17). Two types of SA sequences have been functionally defined: Type I SA sequences translocate the N-terminus of the molecule, while Type II SA sequences translocate the C-terminus (6, 18, 19). The exact mechanism by which ST or SA sequences direct translocational arrest and membrane integration has not been determined. It is hypothesized that translocation is arrested through interaction of the ST sequence with specific receptors located within the translocation channel, in a manner such that the channel is disassembled and the chain becomes integrated into the lipid bilayer at that region (6). In the case of multispanning membrane proteins, topogenic events may be directed by several S and ST or SA sequences in succession. Internal S and SA sequences direct retargeting of a protein to the membrane. An internal SA sequence can also serve as the next membrane domain; alternatively, translocation may continue until an ST sequence, directing the protein to stop translocation, is reached. With this mechanism a separate S and ST sequence pair or SA sequence is required for each membrane domain of a polytopic membrane protein (20). Other mechanisms of generating polytopic membrane proteins may also exist.

Unlike some other commonly observed protein sequence motifs, the S, ST and SA sequences bear no discernible primary sequence homology to each other, although general

domains of topogenic sequences have been defined based on functional and statistical criteria. To date identification of these topogenic sequences has relied largely on hydropathy analysis to identify regions of a protein likely to span the membrane. In the case of S, ST and SA sequences, a core of relatively hydrophobic amino acids is characteristically observed (21-23). In addition, several theories regarding the makeup of the regions flanking the hydrophobic core have been proposed, suggesting a role for charged residues in the orientation of SA sequences and the ability of ST sequences to carry out their function (21).

Recently, a new set of topogenic sequences, known as pause-transfer sequences, have been identified. A pause-transfer sequence can direct a nascent polypeptide to pause transiently as it traverses the ER membrane and then resume translocation, passing on into the lumen. This phenomenon was originally identified for Apolipoprotein B (Apo B), and an element (B'), consisting of codons 78 to 111 of mature Apo B, was found to direct this behavior (24, 25). Subsequently, similar elements in Glucose Regulated Protein 94 (GRP 94) and Prion Protein (PrP) were identified (26), all of which not only demonstrate similar mechanistic phenomena but whose pause-transfer sequences also contain a rough sequence similarity to B'. In the case of PrP, this sequence appears to be a variation both on the theme of a pause-transfer and of a classic stop transfer sequence, establishing itself as a unique topogenic element. Molecular characterization of this sequence and the mechanism of translocation it directs, as well as possible physiological consequences of these events, will be extensively discussed in chapters 2 and 3.

II. Translocation as a receptor-mediated process

Before delving more deeply into the various activities directed by topogenic sequences, it is helpful to consider the machinery involved in conventional protein translocation and the means by which such sequences might direct these events. In this context, translocation across the ER membrane can be thought of as a receptor mediated

process in which the topogenic sequence is the ligand and the translocation machinery is the receptor (6).

Early events in protein translocation into the ER

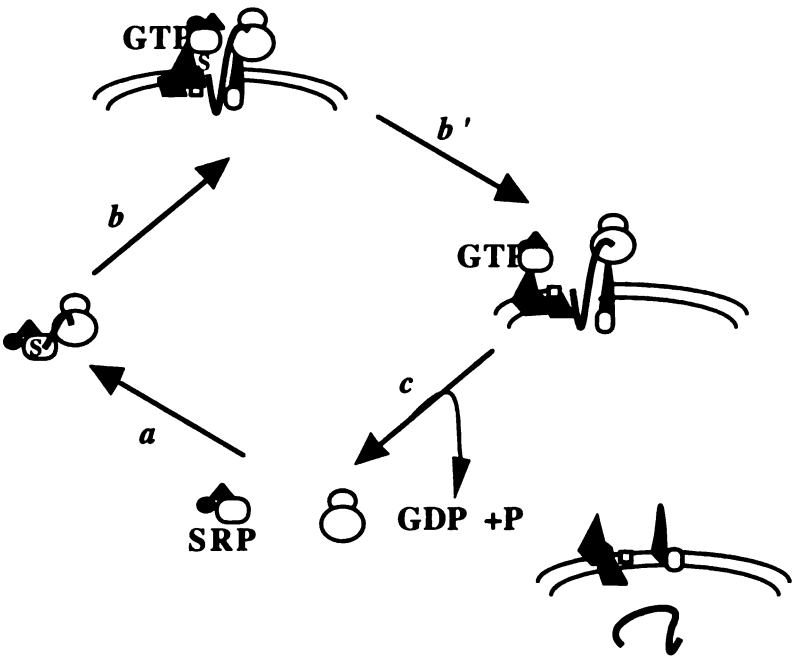
Targeting of a nascent chain to the ER is accomplished by the interaction between an S or SA sequence with a specific receptor protein, Signal Receptor Particle (SRP), probably the most well-characterized component of the translocation machinery. SRP is a cytosolic ribonucleoprotein consisting of 2 heterodimers (68-70kD and 9-14kD), 2 monomers (54kD and 19kD) and a 7S RNA containing homologous repeat domains (6, 27-32). The 54kD subunit contains the signal sequence binding site and contains two functional domains: an N-terminal G domain , so named because of a putative GTP binding site and a C-terminal M-domain that is particularly methionine rich (33). The amphipathic alpha-helical nature of the M domain is predicted to allow a tertiary structure containing a superficial hydrophobic groove, which could provide a binding site for the hydrophobic signal sequence (34,35). In addition, the flexibility of the "methionine bristle" created by the methionine rich M-domain could provide a high degree of variability of size and shape of the hydrophobic groove, allowing for interaction with the diverse primary sequence found in signal sequences (36, 37). Binding of a signal sequence is thought to stabilize SRP54 in a nucleotide free state (46). Studies using photoactivatible cross-links incorporated by nascent chains have shown that the M-domain mediates attachment of SRP54 to the rest of the SRP subunits, binding directly to the 7S RNA . The role of the 68-70 kD heterodimer is unclear, although it appears to interact with the complex via the 7S RNA also. The 19kD monomer appears to facilitate the binding of SRP 54 to RNA. (36-38). While SRP directs ER targeting of most secretory and transmembrane chains, there are a few exceptions. SRP receptor (SR), an ER membrane bound complex, specifically recognizes SRP bound to the ribosome-attached nascent peptide chain. SR consists of two subunits, SRα, a 69kD protein that binds SRP, and SRß, a 30kD integral membrane protein (39-43, 49). SRα contains a GTP binding site and intrinsic GTPase activity (44).

In the simplest example of SRP dependent protein translocation (Fig. 1), that of a classic secretory protein with a cleavable, N-terminal signal sequence, the ribosome bound nascent chain/SRP complex is first targeted to the ER membrane by the affinity of SRP for its receptor, SR (Fig. 1a). The interaction of SRP with SR has been proposed to result in a conformational change in SRP54, triggering GTP binding while reducing its binding affinity for the signal sequence (45)(Fig. 1b). SRP and SR then dissociate from the translocation apparatus and subsequent GTP hydrolysis results in dissociation of SRP from SR, leaving SR and free SRP ready for a new cycle (46) (Fig. 1c). In other studies, GTP-bound SRα also appears to be required for progression through this targeting cycle (47, 48). SRα may undergo a similar cycle of GTP binding and hydrolysis stimulated by its binding to SRP. The ribosome/nascent chain complex, now at the ER membrane, is fully translocated into the ER and the signal sequence is cleaved by signal peptidase either during or immediately after translocation (6).

Translocation channel

The events directing the post-targeting steps of translocation are considerably less clear than the initial SRP-SR mediated events. Crosslinking techniques, utilizing photoactive aminoacyl tRNA lysine analogs, incorporated into an SRP-arrested nascent chain, have allowed identification of several putative translocon components (50). One of the early protein components identified was termed Signal Sequence Receptor (SSR) because it was proposed to interact with the free signal sequence upon its release from SRP, to initiate the process of translocation. SSR consists of two subunits, a 35kD integral membrane protein (SSRα) and a second component of similar size (SSRß) (51). SSRα crosslinked to a signal sequence , as well as more distal portions of a nascent polypeptide chain, and studies using an antisera to SSRα to inhibit translocation provided early evidence of a role for SSR in translocation (5b). However, subsequent studies using an *in vitro* reconstituted translocation system, found that microsomes that had been

Figure 1: Receptor Mediated Model for Translocation into the ER



a. SRP binds to the signal sequence of a nascent chain as it emerges from the ribosome and targets it to the ER. where it binds to its receptor, SR. SRP 54 undergoes a conformational change and binds GTP. c. GTP binding decreases the affinity of SRP54 for the signal sequence and the ribosome/nascnet chain complex dissociates from SRP/SR and continues translocaiton into the lumen. c. GTP hydrolysis catalyzed by the GTPase activity of SRP54 results in dissociation from SR, generating free SRP and unoccupied SR ready for a new cycle.

immunodepleted of SSR showed similar translocation efficiency of preprolactin as non-depleted vesicles, whereas microsomes depleted of SR were not translocation competent (5a). Thus, the role of SSR in translocation is now somewhat more nebulous, as it does not appear to be essential for all proteins directed to the ER.

Another putative component of the translocation channel is the 88kD phosphoprotein, Calnexin, which was originally purified as part of a complex containing SSRα and β, and a 25kD protein of unknown function (52). Both calnexin and SSRα were found by calcium overlay studies to be the major Ca^{2+}-binding proteins of the ER (52). A possible role in Ca^{2+} dependent regulation of protein translocation has been proposed for this complex. Further studies suggested that calnexin associates transiently with a diverse group of newly synthesized proteins during folding events. Recently, calnexin was shown to associate specifically with incompletely folded glycoproteins, with its dissociation rate correlating with the rate of folding. Calnexin also shows prolonged association with slowly folding or misfolded proteins (53-55). Another ER lumenal protein (BiP), a 78kD member of the heat shock protein family, is found associated with both misfolded proteins as well as unassembled components of oligomeric complexes such as immunoglobulin heavy chain (56). Both calnexin and BiP appear to promote and maintain a proper conformation state of newly synthesized proteins as well as recognize proteins that are misfolded (57). There is considerable evidence to suggest that the recognition of misfolded proteins by BiP is the first step in a system designed to clear the ER of malfolded or aggregated proteins (58). Thus it appears that much of the machinery associated with translocation may consist of proteins that serve to increase the efficacy of both translocation and proper protein folding.

A 35kD transmembrane protein called TRAM, for translocating chain-associating membrane protein, was discovered in 1992 using nascent chain crosslinking techniques (4, 8). Unlike SSR, TRAM appears essential for translocation of a number of proteins. Sequence analysis of TRAM predicts eight membrane spanning domains which are

postulated to form part of the translocation channel; little else is known about TRAM at this point. Another protein, sec 61p also appears to be essential for translocation. Sec 61p was initially identified by genetic studies in *S. Cerevesiae* (11). Three ER membrane proteins, sec61P, sec 62p and sec 63p, as well as the resident ER lumen KAR2p protein (BiP), have been found in a complex that can be cross-linked to a translocationally arrested nascent polypeptide. The translocating chain appears to be specifically cross-linked to sec61p and BiP, with sec62p and sec 63p as well as ATP, being required for efficient association of sec61p with the nascent chain (59, 60). The role of BiP appears to be both at the point of association with the nascent chain as well as at some point in the later stages of translocation (59). Sec61p sequence analysis suggests a multispanning membrane topology with minimally exposed cytosolic and lumenal domains, making it a likely candidate for a major channel component. Because of the similarity in crosslinking patterns between yeast and canine microsomes, it seemed likely that analogous proteins exist in mammalian cells, and recently a sec 61p homolog has been cloned from MDCK cells (61).

Other ER proteins that appear to have a peripheral role in translocation, primarily via their effect on protein folding and isomerization, are Protein Disulfide Isomerase (PDI), Peptidyl-Prolyl Isomerase (PPI), ribophorins and a host of other as of yet unidentified proteins (62, 63, 64). Because translocation across the ER membrane is a common step required for a structurally and functionally diverse set of proteins, the presence of many proteins not required for conventional translocation may reflect different needs of various proteins containing divergent topogenic information.

Modifications occurring during protein transport

Most proteins sequestered in the lumen of the ER before being secreted are glycoproteins. Glycosylation occurs on the lumenal side of the ER membrane in a sequence of events involving the transfer of an "activated" oligosaccharide to the NH2 group of an asparagine residue occurring in one of two sequences: Asn-X-Thr or Asn-X-

Ser (where X is any amino acid). The oligosaccharide, consisting of N-Acetylglucosamine, mannose and glucose, is built up sugar by sugar on a donor lipid molecule, dolichol phosphate. Dolichol is activated to form dolichol phosphate in an ATP dependent step. In the cytosol, nucleotide-sugar intermediates are formed, which then donate their sugar to the dolichol phosphate in an orderly sequence involving the formation of a pyrophosphate bridge. The oligosaccharide is then translocated to ER lumen in an enzyme catalyzed isomerization step, and the energy released from the pyrophosphate bond is used for transfer to the protein. After transfer to the protein, the first in a series of steps responsible for the extensive carbohydrate remodeling occurs. The three glucose residues are removed by a specific glucosidase in the lumen of the ER . Trimming of the mannose side chains continues in the Golgi and addition of complex oligosaccharides can occur (65,66). These events can be used experimentally to determine the location of newly synthesized glycoproteins in the secretory pathway. Treatment of proteins with Endoglycosidase H (Endo H), will remove core carbohydrate side chains added in the ER, but once complex carbohydrates have been added in the Golgi, this enzymatic activity is prevented. The acquisition of Endo H resistance is a commonly used assay for a glycoprotein that has proceeded on to the Golgi apparatus (67).

Addition of C-terminal glycolipid anchors, is also thought to occur in the ER. The exact mechanism is unknown but it involves the recognition and cleavage of a C-terminal signal and addition of a glycosyl-phosphatidylinositol (GPI) anchor. The glycolipid moiety is transferred to the C-terminus of a diverse set of proteins. Cleavage of the anchor by a phosphatidylinositol specific phospholipase C or D (PIPLC or PIPLD), releases soluble protein extracellularly (68, 69). The purpose of this means of reversible membrane anchoring is unclear, but it is probably multifunctional.

Another important modification that takes place in the ER is the formation of disulfide bonds. This reaction is catalyzed by the enzyme PDI, a dimer of identical subunits having an approximate molecular weight of 57kD. This abundant enzyme is required for

the folding of a large number of nascent polypeptide chains (62). Additionally, many other modifications of protein conformation probably occur in the ER. Events occurring at this point in biosynthesis most certainly have ramifications on the ultimate fate of a protein following the secretory pathway.

III. Protein conformation and translocation

An essential feature of overall protein function is its proper folding into secondary, tertiary and quaternary structures. Both specific sequences within the protein, as well as the hydrophobicity of the surrounding milieu or interactions with other proteins, can affect the folding of a given protein. A diverse group of proteins that appear to catalyze protein folding reactions, have been identified, referred to as chaperones. The role of both cytosolic and ER chaperones, some of which may also be considered channel components, in translocation across the ER has been demonstrated previously (70-72). Studies in *S. Cerevesiae* showed that the 70kD member of the heat shock protein family (hsp70) was essential for translocation and cell growth . The action of hsp70 may be important in maintaining a certain secondary structure as the nascent chain emerges from the ribosome, conducive to interaction with translocon components (73, 74). Translocation channel components discussed above, particularly those that appear to be non-essential for translocation *in vitro*, may play a similar role as the chain passes from the aqueous environment of the cytosol to the highly concentrated environment of the translocation channel.

The prion protein (PrP), the topogenic features of which are the subject of this work, has been shown to assume either a fully translocated or transmembrane topology depending on the presence or absence, respectively, of a cytosolic factor (75, 76). This factor has not yet been identified and the mechanism by which it carries out its function

remains unknown, but our recent data suggests it is likely to be involved in promoting translocation into the lumen from a transmembrane intermediate. It is possible that folding events occurring during translocation may have an important role PrP function. Before proceeding with a discussion of PrP translocation, it is necessary to discuss some of the fundamentals of protein folding events and sequences that direct them.

Secondary Protein Structures

The secondary structure of a given protein is, to some extent, determined by its primary amino acid structure. The interior residues of a protein often consist of hydrophobic side chains, a feature that drives the protein to pack hydrophobic residues into the interior, away from the aqueous environment of a cell, and creates a hydrophobic core and a hydrophilic surface. Such a folding event results in the formation of either alpha-helical structures or beta-pleated sheets. Which of these two secondary structures is assumed is dependent both on the amino acid sequence and the surrounding environment, but certain amino acid residues have the propensity to form one or the other (77).

Alpha helices have a predictable periodicity with 3.6 residues per turn. A glycine residue is often seen initiating and a proline residue is often seen terminating an α-helix. Side chains of residues in an α-helix project out from the helix and do not interfere with helix formation, with the exception of proline in the 1 or 4 position of a turn. The side chain of a proline residue binds to the N- group, disallowing the formation of a hydrogen bond and causing a bend in the helix. A number of other amino acid residues have a high propensity to break or disrupt an α-helix. Ala, Glu, Leu and Met are good helix formers, whereas Pro, Gly, Tyr, and Ser are poor helix formers (77).

The second major secondary structure is the ß-sheet. This structure is built from a combination of several regions of the polypeptide chain rather than one continuous region as is the case with an α-helix. The individual regions, known as ß-strands, are usually 5-10 residues long. There are two ways to form a pleated sheet: either the amino acids in the

aligned ß-strands all run in the same direction (amino terminal to carboxyl terminal), or the amino acids in successive strands run in opposite directions (amino terminal to carboxyl terminal and then carboxyl terminal to amino terminal). The former example is referred to as a parallel sheet, and the latter example is referred to as an antiparallel sheet. Most protein structures are built up from a combination of α-helices, ß-strands and connecting loop regions, referred to as supersecondary structures or motifs (77).

Transmembrane proteins, which have both hydrophobic and hydrophilic regions on their surfaces, are usually anchored in the membrane by one or more regions of α-helices composed predominantly of hydrophobic residues. To span the lipid bilayer, a minimum of about 20 amino acids is required (78). A number of mathematical models predicting hydrophobicity have been developed which are typically used to identify potential transmembrane regions. Each amino acid side chain within a transmembrane helix has a different hydrophobicity. For each position in a sequence the hydropathic index is calculated. The hydropathic index is the mean value of the hydrophobicity of the amino acids within a given region, usually about 19 residues long (79). Charged amino acid residues are rarely found within a transmembrane helix because the energy cost of maintaining a charged residue in a lipid environment would be too great. However, two residues of opposite charge close together inside the lipid membrane can be tolerated because of their neutralizing interactions. In addition to overall hydrophobicity, their are other structural determinants of the membrane spanning potential of a given amino acid sequence. ß-pleated sheet secondary structures are not conducive to membrane spanning potential, and several hydrophobic amino acids have ß-branched side chains and therefore a propensity to disrupt α-helices in aqueous solutions. Studies using short (less than 20 residues) peptides have been used to study these properties in greater detail. Studies in which alanine residues were replaced with either glycines or isoleucines suggested that a surrounding lipid environment can increase the proclivity of a hydrophobic peptide to form an α-helix regardless of the disruptive presence of ß-branched side chains. In a mixed

lipid/aqueous environment such as might be encountered in the translocation channel of the ER membrane, these same peptides showed a 50% reduction in the formation of α-helices (80). These findings would be consistent with what is known about the structure of topogenic elements such as signal and stop transfer sequences.

Structural features of topogenic sequences

Although signal sequences are divergent in their exact amino acid content, certain structural characteristics can be identified (Fig. 2A). The N-terminal basic region usually contains 1-2 charged residues; the core hydrophobic region usually contains 10-15 hydrophobic or neutral residues and the C-terminal cleavage region usually consists of 5-7 amino acids and often ends with an alanine or glycine (6, 81). Recent studies have suggested that the hydrophobic domain is capable of both α-helix and ß-sheet structures, and that the transition between these two may have a role in signal peptide function. Mutations in the hydrophobic region of a signal sequence, in which one or two in a string of Ala residues is replaced by Val, Ile or Leu, are observed to show an increased translocation rate (82). It has been suggested the signal peptide is in a ß-sheet conformation when associating with the surface of the lipid phase of phospholipid monolayers and in an α-helical conformation when located within the lipid phase (83). Because translocation is now known to proceed through an aqueous pore, the increased tendency of two of these mutant signal peptides to form ß-sheets may alter the kinetics of translocation. On the other hand, these same mutations increase the overall hydrophobicity of the region which one might predict would increase the propensity of the peptides to reside in the membrane.

Stop-transfer sequences, also quite divergent in their exact sequences, have a similar structural organization to signal sequences (Fig. 2B). The amino and carboxyl domains of ST sequences contain polar and charged residues, while the core of the sequence consists of an unbroken stretch of 19 to 30 hydrophobic amino acids (84). Studies in which negatively charged residues were inserted into the hydrophobic core of a

Figure 2: Structural Domains of Common Topogenic Sequences

A. Signal Sequence Structure

size	2-8	8-15	4-7
overall charge	+	0	+/-
functions	trans-location/ (cleavage)	SRP binding	cleavage/ (trans-location)

B. Stop Transfer Sequence Structure

	?	19-30	?
	+/-	0	+
	integration?	trans-location arrest??	orientation?

Schematic representation of structural and functional domains of a signal sequence (A) and a stop transfer sequence (B). Ns, Nst, Hs, Hst, Cs and Cst have been defined on the basis of sequence analysis and mutagenesis of known signal and stop transfer sequences.

model stop transfer sequence resulted in membrane spanning proteins that did not integrate into the membrane (85), i.e. these mutant proteins failed to carry out the final step of ST action. Although most transmembrane regions are found in an α-helical conformation, helix destabilizing residues such as Gly, Val, Ile and Ser often constitute as much as 50% of transmembrane domains (86). Such a composition may contribute to conformational flexibility or susceptibility to helix destabilization of membrane proteins whose function is regulated by trans-acting factors. Clearly, the influence of the surrounding environment as well as the intrinsic properties of a given protein sequence are important in determining the conformation of the protein as it translocates across the ER membrane and, subsequently, the final topology of the protein.

IV. Methods by which ER translocation can be viewed experimentally

The development of cell-free translation systems has greatly aided our ability to analyze the events occurring during translocation of a nascent polypeptide across the ER membrane. Such systems are capable of reconstituting protein translation and, when supplemented with microsomal membranes derived from rough ER, of translocation across the lipid bilayer as well (87-90). Cytosolic extracts derived from rabbit reticulocyte lysate (RRL) and wheat germ embryo (WG) are two of the most commonly used cell-free systems, and canine pancreas is the most common source of translocation competent microsomal membranes. The cytosolic extracts contain ribosomes and an assortment of soluble factors necessary for protein synthesis and protein folding; when primed with cell-free transcription products of cDNA engineered into an expression vector, cell-free systems are capable not only of translation but of post-translational folding and assembly processes (91). Newly synthesized proteins can be detected by including at least one radiolabeled amino acid followed by analysis on denaturing polyacrylamide gels (SDS-PAGE). When ER derived microsomal membranes are included in the translation reaction, a number of criteria can be used to ascertain whether or not the protein has engaged the translocation

machinery and successfully translocated across the ER membrane. First, in the case of a protein with a cleavable, N-terminal signal sequence, comparison of translation products synthesized in the presence and absence of membranes often reveals a size shift due to the removal of the signal peptide that is reflected in a shift in electrophoretic mobility on SDS-PAGE. In addition, because translocation competent membranes typically retain 50-90% of their glycosylation activity, an upward size shift is often seen when glycoproteins are synthesized in the presence of membranes; this modification can be removed either by parallel synthesis in the presence of a tripeptide inhibitor of glycosylation (AP) or by treatment with Endo H. Although the processes of translocation and glycosylation are less efficient in the cell-free system than in an intact system, such as transfected cells or microinjected *Xenopus* Oocytes (95), this potential drawback can be used to an experimental advantage. Because translocation is now thought to be a receptor mediated process, presumably some nascent chain/translocon interactions are rate limiting. Thus, the reduced efficiency of the cell-free system can result in transient accumulation of nascent chains at various stages of translocation; such steps often occur too quickly to be observed in an intact cell. In addition, the use of recombinant DNA technology to create chimeric proteins has allowed detailed analysis of topogenic sequences (90). By constructing plasmids that express fusion proteins with different topogenic sequences between defined passenger domains, the functions of different topogenic sequences can be directly compared.

The simplest assay for assessing the function of a given topogenic sequence is one that determines the topology of the full-length protein, i.e. transmembrane or fully translocated, into the ER lumen. A commonly used method of determining topology is a proteolytic assay (Fig. 3) (17). When translation products synthesized in the presence of microsomes are treated with proteases, secretory proteins, which will be residing entirely in the ER lumen, will be protected from protease digestion by the surrounding membrane. On the other hand, cytosolically exposed domains of transmembrane proteins will be digested

by the protease and lumenally located domains will be protected from digestion . Upon SDS-PAGE analysis, proteolytic digestion of transmembrane proteins will be reflected in the appearance of membrane protected fragments and a corresponding loss of full-length protein. In addition, the use of anti-peptide antibodies to immunoprecipitate specific proteolytic fragments can aid in identifying those regions of the protein that resided in the lumen.

Proteins which are determined to be transmembrane by a proteolytic assay, or by their fractionation with membranes in nondenaturing detergents, are of two classes: integral membrane proteins (IMP), that are associated directly with the lipid bilayer and peripheral membrane proteins (PMP), that associate with the membrane solely through interactions with IMPs. A second assay for the topology of a protein distinguishes between integral and peripheral membrane proteins. When membranes are subjected to alkaline perturbants such as Sodium Carbonate, membranous sheets are formed releasing lumenal proteins and disrupting protein/protein interactions (92). Thus any proteins that are membrane associated due to interaction with other integral membrane proteins are released into the supernatant upon high speed centrifugation, while proteins that are directly interacting with the lipid bilayer are recovered in the pellet (Fig. 3B). In conjunction with proteolytic analysis, this method provides a powerful tool for characterizing the topology of a given protein or chimeric construct.

It is possible to experimentally prolong certain steps in translocation, either by perpetuating nascent chain/ribosome interaction by removing the stop codon, or by making use of mutations with altered translocational kinetics. By priming translation reactions with transcript prepared from DNA that had been digested at a restriction site upstream of the stop codon, ribosomal release of the nascent chain is prevented and translocational events that require ribosome attachment can be detected. Subsequent release of the ribosome can be achieved by treatment with puromycin or with puromycin followed by EDTA. Puromycin is a structural analog of a tRNA molecule which is incorporated into the

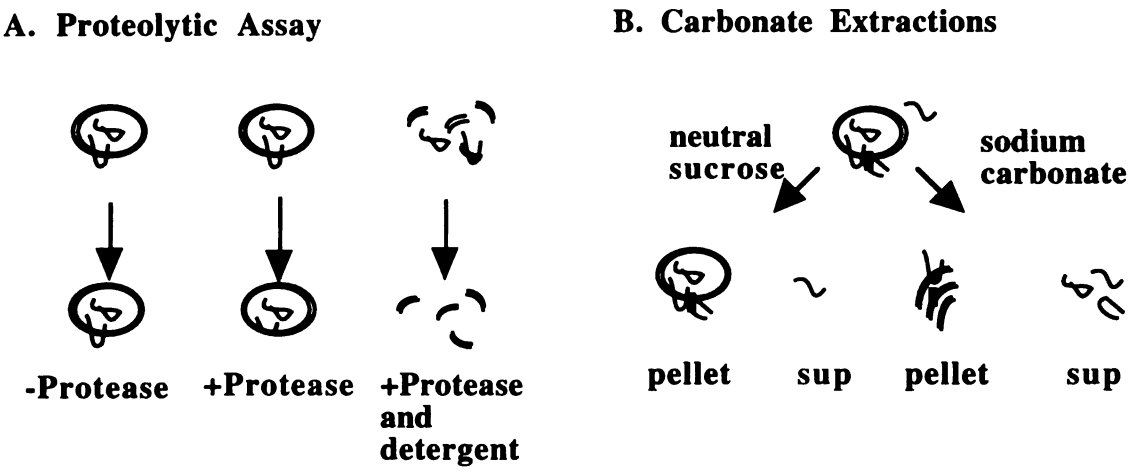
growing end of the nascent chain and triggers its premature release. EDTA causes premature detachment of the ribosome from the ER membrane by chelating Mg2+ ions, upon which maintenance of the ribosome/membrane junction is dependent. These techniques enabled the initial characterization of the pause-transfer element of Apo B as well as the similar activity directed by a related region of PrP (24, 94). In addition, translation can be uncoupled from translocation in the cell-free system, allowing characterization of translocational events that occur independent of translation (93). By presenting microsomal membranes with ribosomally attached nascent chains, translocation across the ER membranes occurs and, following treatment with puromycin, correct final topology is attained. In this manner, factors involved in post-translational regulation of protein topology can be identified by addition of extracts to cell-free systems, without sacrificing overall levels of protein synthesis (93).

While the cell-free system is highly manipulable, and thus a convenient way to identify translocation intermediates, an intact system is sometimes necessary to establish the physiological significance of observations made in the cell free system, or to examine steps in translocation that do not proceed with significant efficiency in fractionated systems. A commonly used expression system for this purpose is the *Xenopus* Oocyte, which has been found to reflect cellular topology and conformation of most proteins with relative accuracy (95). The large size comparatively slow kinetics of the oocyte make it more conducive to translocation studies than transfected cell lines. Oocytes are transcriptionally quiescent, and by microinjecting mRNA in the presence of radiolabeled amino acids, synthesis and subsequent processing of the resulting protein can be easily monitored.

Through use of a combination of the experimental systems discussed here, significant information regarding topogenic sequences and the mechanisms of translocation they direct can be ascertained. In the case of the prion protein (PrP), cited earlier as an example of a protein whose translocation into the ER lumen proceeds in a somewhat unconventional manner, these techniques have allowed the elucidation of a unique topogenic element,

regulation of which may have a consequential impact on the final conformation and function of the protein.

Figure 3: Assays for determining protein topology



A. Proteolytic assay. Translation products (either produced in the cell-free systems in the presence of dog pancreas membranes or expressed in microinjected *Xenopus* oocytes, homogenized in a neutral sucrose buffer to preserve microsomal integrity) are treated either with (center) or without (left) proteases or with proteases in the presence of non-denaturing detergents (right). Cytosolically localized domains are digested by protease to generate proteolytic fragments corresponding to the size of the lumenally localized domains. In the presence of detergent, membranes are solubilized and all protein is digested. B. Carbonate extraction. Translation products as described in A are diluted in either a neutral sucrose buffer (left) or sodium carbonate at pH11.5 (right) which forms membranous sheets and disrupts protein protein interactions. After high speed centrifugation only integral membrane proteins are recovered in the carbonate supernatant.

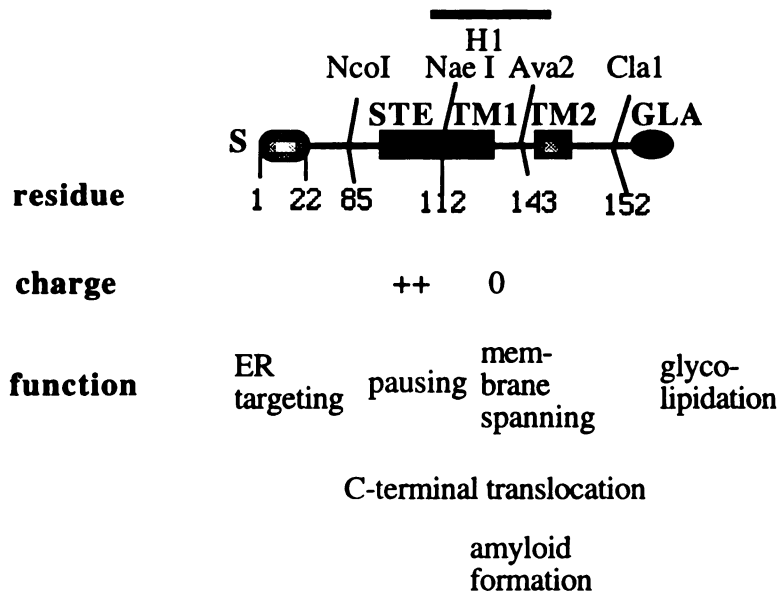
Chapter 2: Molecular determinants of Carboxyl-terminal domain translocation in the Prion Protein

I. Introduction

The prion protein (PrP) is a glycoprotein that has been implicated in the pathogenesis of several neurodegenerative diseases, such as scrapie in rodents and sheep, and Creutzfeldt-Jakob disease and Gerstmann-Straussler-Scheinker syndrome in humans (95). These diseases are marked by the accumulation of an abnormal PrP isoform (PrP^{Sc}) that is thought to be identical in amino acid sequence to the normal cellular isoform (PrP^{C}) (96). PrP^{C} is not associated with neurodegeneration and its normal function is unknown (97-99). PrP^{C} has been shown to be a fully translocated polypeptide, tethered to the external surface of the plasma membrane by a cleavable glycolipid anchor at its carboxy (C-) terminus (98). In addition to its association with neurodegenerative diseases, PrP^{Sc} is also distinguished from PrP^{C} by its relative protease resistance and insolubility, and by the fact that it cannot be released from the membrane by treatment with phosphatidylinositol specific phospholipase C (PIPLC) (100, 101). The precise transmembrane orientation of PrP^{Sc} has been difficult to determine because its protease resistance renders standard assays for topology (e.g. protease digestion in the presence and absence of detergent) difficult to interpret.

The diseases associated with PrP^{Sc} can result from germ line mutations in the PrP gene, but they are also found to be transmissible, the major protein component of the infectious prion being PrP^{Sc} (96). PrP is encoded in a single copy gene with the entire open reading frame contained in a single exon (102), suggesting that neither activation of a second gene nor an alternative splicing event gives rise to PrP^{Sc}. Structural analysis of PrP^{C} and PrP^{Sc} reveals no covalent differences between these isoforms (103), yet kinetic studies contend that PrP^{Sc} is derived from PrP^{C}, by a post-translational modification

Figure 4: Topogenic Information in PrP



A schematic representation of the coding region of hamster PrP is shown with the topogenic sequences, predicted amyloidigenic regions and relevant restriction sites depicted.

(104). Recent investigations with synthetic peptides corresponding to regions of PrP predicted to be amyloidigenic, suggest that the two isoforms differ in their secondary and tertiary structure (105). However it is still unclear how these structural differences between PrP^C and PrP^{Sc} are attained in the cell.

To explore the structural differences between the two PrP isoforms, the biogenesis of PrP has been studied in cell-free systems as well as in living cells. These studies have revealed some unconventional features. When synthesized in wheat germ extracts (WG) in the presence of dog pancreas microsomal membranes, PrP chains are found predominantly integrated in the bilayer, spanning the membrane twice (106). In this topology, roughly equivalent-sized domains corresponding to the amino (N-) and C- termini are translocated into the ER lumen . This transmembrane orientation is consistent with predictions of hydropathy analyses (107). However, it differs from the topology demonstrated for PrP^C in normal brain (97), transfected cell lines (108), *Xenopus* oocytes (XO) or in the rabbit reticulocyte lysate cell-free translation system (RRL) supplemented with dog pancreas microsomal membranes (75), in all of which most chains of PrP are fully translocated.

The determinants of these unusual features of PrP topogenesis have been investigated previously (26) (Fig. 4). Transmembrane chains of PrP span the membrane first at a hydrophobic region, termed TM1, located at amino acids 112-143 from the initial methionine (106). A 24 amino acid residue hydrophilic sequence just N-terminal to TM1, termed the Stop Transfer Effector (STE), was shown to direct nascent PrP chains to stop during translocation, hence placing the adjacent TM1 sequence in a membrane-spanning orientation (26). In addition, a cytosolic factor present in RRL was implicated in the generation of predominantly fully translocated chains in RRL, as opposed to the integrated transmembrane chains that predominate in WG (76).

Although these studies provide some insight into the molecular mechanism of PrP biogenesis, important questions remain unanswered. First, the mechanism by which the C-terminus is translocated into the ER lumen, resulting in a transmembrane form that spans

the membrane twice, has not previously been investigated. Based on earlier studies of multi-spanning membrane proteins, an internal signal (S) or a signal-anchor (SA) sequence, located downstream of TM1, would be expected to direct the C-terminal translocation event (18, 20, 109, 110). Alternatively, PrP might use a previously unrecognized mechanism to achieve this second translocation event. Second, the relationship between PrP chains with transmembrane versus fully-translocated topology remains to be clarified, as does the question of whether transmembrane PrP ever exists *in vivo*. One possibility is that only the fully translocated form is physiologically relevant and that the transmembrane form is an artifact generated only in cell-free translation systems (111). Alternatively, both transmembrane and secretory forms may be relevant to PrP biogenesis *in vivo*. Although a transmembrane form has not been observed as the final product of PrP biogenesis under normal conditions in cells, it may represent a transient intermediate in the pathway generating secretory PrP or it may be the final topology of PrP *in vivo*, under special circumstances. Such hypotheses are plausible, given that specialized transient transmembrane intermediates have been shown to occur during the biogenesis of other complex secretory proteins, such as Apolipoprotein B (24, 25) and Glucose Regulated Protein 94 (Hegde et al, submitted).

Here we have investigated these unanswered questions in PrP biogenesis by the use of protein chimeras in cell-free systems and *Xenopus* oocytes (XO). We find that PrP is distinct from other previously studied membrane proteins in that no information downstream of TM1 is necessary for translocation of the C-terminal domain. The sequence comprising STE and TM1 is sufficient not only to stop translocation of the N-terminus at TM1, but also to direct translocation of the subsequent C-terminal domain, thereby generating a second transmembrane region (TM2). This conclusion is reached on the basis of expression studies presented here in which the codons for STE-TM1 were engineered into the coding regions for heterologous protein chimeras. Upon expression of these chimeras, the second membrane-spanning region and translocated C-terminal domain are

derived from either of two unrelated soluble proteins, globin and prolactin. Furthermore, we show that STE-TM1 does not appear to direct C- terminal domain translocation by acting as a simple internal S or SA sequence. Instead, the data suggest that this sequence stops chain translocation in a novel way that allows subsequent translocation of the C-terminal domain with generation of TM2. Finally, our studies demonstrate a precursor-product relationship between transmembrane but carbonate extractable and fully-translocated STE-TM1-containing protein chimera. These findings suggest a mechanism for STE-TM1 action and a model for PrP biogenesis (64).

II. RESULTS

Sequences beyond TM1 are not necessary for C-terminal domain translocation

By analogy to previous studies on multispanning membrane proteins (18, 20, 109, 110), we hypothesized that the C-terminal translocation of PrP, observed in WG, was directed by an internal S or SA sequence located downstream of TM1. The two most likely candidate sequences were TM2 and the hydrophobic signal for glycolipid attachment located at the extreme C-terminus (69, 98). To test this hypothesis, we replaced the entire C-terminus of PrP just beyond TM1 with the coding region for codons 1-143 of chimpanzee alpha globin containing an 8 codon N-linked glycosylation acceptor site. The resulting SP6 expression plasmid was called PrP-G and encodes a chimera of PrP (from initial methionine through amino acid residue number 152) and globin. Previous studies have demonstrated that this globin coding region contains no topogenic sequences to direct translocation across the ER membrane (20). Thus, if a second S or SA sequence downstream of TM1 was responsible for directing C-terminal domain translocation in PrP, PrP-G should span the membrane only once, with the globin domain residing in the cytosol where the engineered N-linked glycosylation site cannot be used. PrP-G was expressed by transcription-linked translation in WG in the presence of dog pancreas microsomal membranes (Fig. 5A). We analyzed topology of the encoded protein chimera by treatment of translation products with proteinase K (PK), in the presence or absence of non-denaturing detergents, followed by immunoprecipitation with domain specific antisera, polyacrylamide gel electrophoresis in sodium dodecyl sulfate (SDS-PAGE) and fluorography (F). In some cases, samples were treated with endoglycosidase H (Endo H) to determine if the glycosylation site engineered within the globin coding domain had been used. The presence of N-linked carbohydrate would provide independent evidence of domain translocation, since N-linked glycosylation is believed to occur only in the ER lumen (67).

To our surprise we observed that PrP-G spanned the membrane twice with translocation of the globin domain which was glycosylated. Translation products of PrP-G corresponding to the full-length protein, were identified by size and immunoreactivity to antisera directed against PrP N-terminus (anti-P1), or antisera directed against the region of globin containing the glycosylation site (anti-G) (Fig. 5A lanes 1, 3 and 10, 11). In the absence of membranes, a band at approximately 30kD is observed (Fig 1A lanes 2 and 10) corresponding to full length untranslocated protein. In the presence of membranes, two signal sequence cleaved, glycosylated bands at approximately 32 and 33 kD as well as a signal cleaved, unglycosylated band of about 27 kD are observed. In addition, there is some translation product corresponding to the 30kD unprocessed band (Fig 5A lanes 3 and 11). Unprocessed protein is observed to some extent with most proteins translated in the WG cell-free system in the presence of dog pancreas microsomes, due to the relative inefficiency of translocation in this system. Glycosylation was demonstrated by size shift upon treatment with Endo H (Fig. 5A lanes 3 vs. 6 and 11 vs. 14). PK digestion resulted in both N- (Fig. 5A lane 4) and C-terminal (Fig. 5A lane 13) proteolytic fragments identified by immunoreactivity to PrP and globin-specific antisera, respectively. The globin reactive fragments were larger than would be expected if only the globin region was protected by the membrane, suggesting the region comprising the cytosolic loop was very small and the C-terminal fragment contained nearly all of the chimera after TM1. To test this hypothesis, we did parallel immunoprecipitations with a peptide specific antibody directed against codons 127-135 of PrP, the domain immediately following TM1. The globin immunoreactive fragments show immunoreactivity to anti-P5 (data not shown), indicating that several residues of PrP thought to comprise the cytosolic loop, located at the fusion point of PrP and globin, are either spanning the membrane or inaccessible to proteolysis. The globin antibody does not recognize the anti-P1 immunoreactive fragments, nor does the anti-P1 antibody recognize the globin reactive fragments, demonstrating that only specific fragments representing either the N-terminus or C-terminus of the molecule,

and not random fragments containing both epitopes, are generated by proteolysis. To ensure that the bands generated by treatment of translation product to Proteinase K are not derived from intrinsically protease resistant digestive products of the untranslocated precursor, we show that protein translated in the absence of membranes is completely digested upon treatment with PK (Lanes 1 and 9). Fragments protected from PK were largely digested when the protecting microsomal membranes were solubilized with non-denaturing detergents (Fig. 5A lanes 5 and 14). Furthermore, the only N-linked glycosylation site in this molecule resides in the C-terminus and a portion of proteolytically generated C-terminal fragments were found to be glycosylated, as demonstrated by the specific downward shift of globin reactive proteolytic fragments when treated with Endo H (Fig. 5A lane 12 versus 17, but not 3 vs. 8). Because N-linked glycosylation is known to occur exclusively in the ER lumen, the presence of glycosylation on these larger globin-immunoreactive proteolytic fragments confirmed that the C-terminal globin domain had been translocated into the ER lumen.

To investigate whether or not these results were an artifact of translation in the WG cell-free system, the topology of PrP-G was analyzed in *Xenopus* Oocytes (XO) microinjected with PrP-G transcript and ^{35}S methionine and pulsed for 3 hours (Fig 1B). As was shown in WG, full-length product, reactive to both PrP and globin antisera (Fig 5B lanes 1 and 4) were converted upon proteolysis into separate fragments with reactivity to either PrP (Fig 1B lane 2) or globin (Fig 5B lane 5) antisera. These protected fragments were digested when proteolysis was carried out in the presence of non-denaturing detergents (Fig. 5B lane 3 and 6). Taken together, these results indicate that information sufficient to direct C-terminal domain translocation lies upstream of the end of TM1, and that this information can be recognized *in vivo*.

Figure 5: Topology of PrP-G in WG and *Xenopus* Oocytes

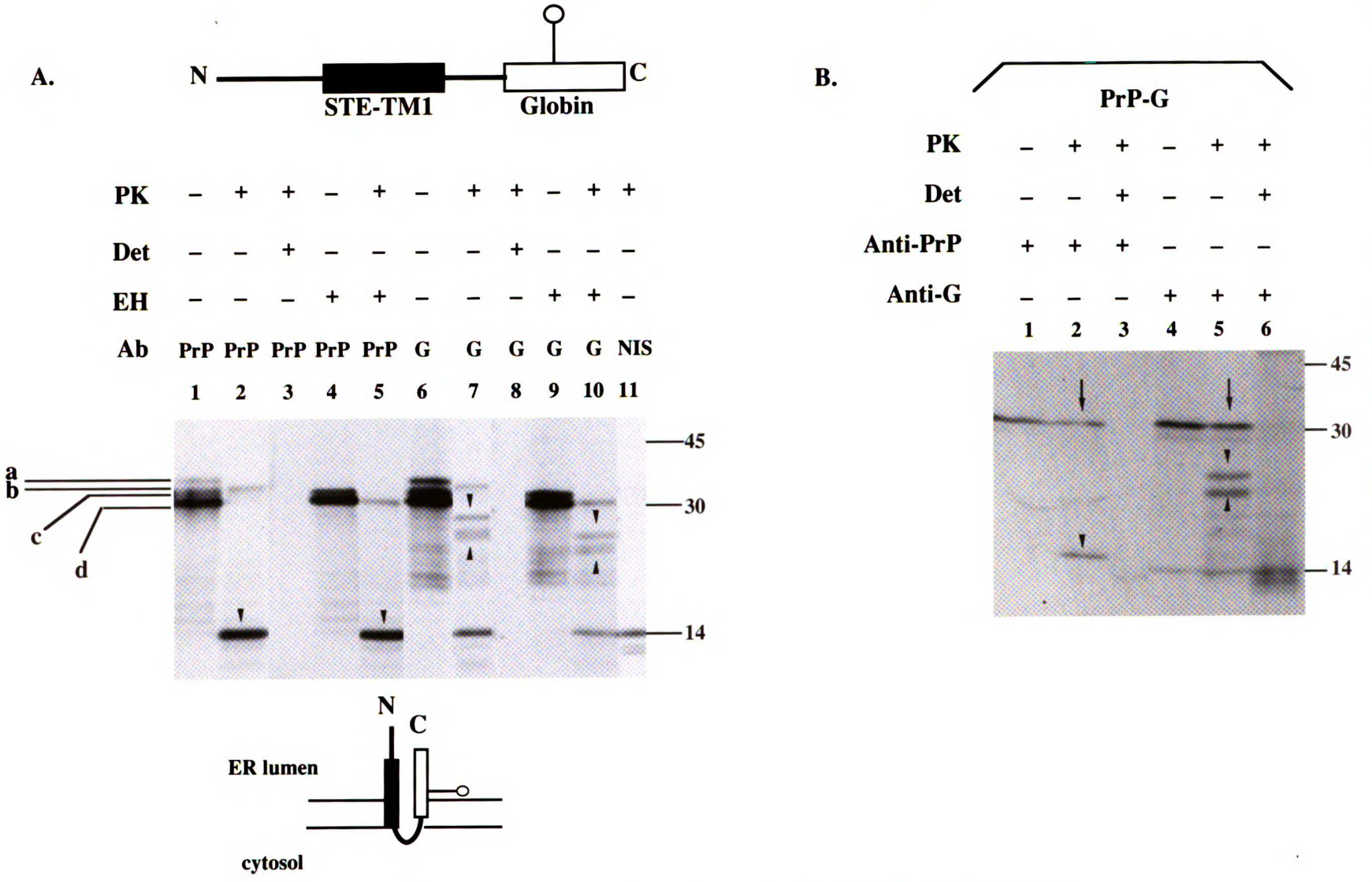


Figure 5 (A) Codons for the entire C-terminus of PrP, from amino acid residue 152 of the cytosolic loop to the termination codon, were deleted and replaced with codons for the globin domain as described, into which a glycosylation site had been engineered (PrP-G). This chimera is depicted in the line diagram at the top of panel A, with the sugar moiety added at the glycosylation site drawn as a ball and stick. N and C refer to amino and carboxyl termini, respectively. *In vitro* synthesized mRNAs were translated in a WG cell-free system in the presence of dog pancreas microsomal membranes for 1 hour at 28^0C. Samples were digested with or without 0.25mg/ml proteinase K in the presence or absence of 1.0% Triton X-100 for 1 hour at 0^0C. Proteolysis was terminated by addition of 10mM PMSF and heating at 100^0C in 10 volumes of 1%SDS for 5 minutes. Samples were then split into two equal aliquots, diluted 20 fold in TXSWB and immunoprecipitated with antisera to the N-terminus of PrP (anti-PrP, lanes 1-5) or globin (anti-G, lanes 6-10), to identify N and C-terminal domains, respectively. These samples were subsequently divided into two equal aliquots and treated with (lanes 4, 5, 9, 10) or without Endo H (lanes 1, 2, 3, 6, 7, 8, 11), to identify glycosylated products. Lanes 3 and 8 represent samples digested with PK in the presence of detergent. Lane 11 shows the products recognized by pre-immune serum. Note the presence of a considerably less intense non-specific band, coincidentally comigrating with the N-terminal PrP-immunoreactive fragment. Arrowheads in lanes 2 and 5 indicate N-terminal proteolytic fragments. Arrowheads in lane 7 indicate larger glycosylated C-terminal proteolytic fragments, which migrate as a doublet. Arrowheads in lane 10 indicate the same doublet after treatment with Endo H and the unaltered unglycosylated fragment. The cartoon below panel A depicts the transmembrane topology of PrP-G with the N- and C-termini located in the ER lumen. (B) *In vitro* transcribed PrP-G mRNA was injected into XO in the presence of radiolabeled methionine. After a 1 hour incubation at 18^0C, XO were homogenized and analyzed by proteolysis, SDS-PAGE and F as described. Lanes 1-3 show N-terminal anti-PrP immunoprecipitated products and lanes 4-6 show C-terminal anti-G immunoprecipitated products. Arrowheads indicate transmembrane fragments, while arrows indicate fully protected chains.

C-terminal domain translocation is directed by the STE-TM1 sequence

In view of the unusual features described for STE-TM1 (26, 76), it seemed plausible that this region itself, might be responsible for directing C-terminal domain translocation. To test this hypothesis directly, the STE-TM1 coding region was engineered into a chimeric protein which had been used previously for studies of protein topology (ref 19 and Fig. 3). This chimeric protein (S.G.X.P) consists of the cleavable bovine prolactin signal sequence (S) at the extreme amino terminus, followed by an N-terminal passenger domain (encoding amino acid residues 1-117 from chimpanzee alpha globin including the 8-codon N-linked glycosylation acceptor site (G), followed by a C-terminal passenger domain encoding residues 57-199 of bovine prolactin (P). Between the G and P coding regions (X) we engineered either STE-TM1 of PrP, a conventional stop transfer sequence (ST) from μ heavy chain of IgM (17, 112) or signal sequence (S) of prolactin. The ST-containing construct has been shown previously to encode an integral membrane protein, spanning the membrane a single time with the N-terminal G domain translocated into the ER lumen and the C-terminal P domain localized exclusively in the cytosol (20); the S-containing construct has been shown to be fully translocated into the ER lumen.

These protein chimeras were translated in WG supplemented with dog pancreas microsomal membranes, and aliquots of the translation products were subjected to PK digestion (Fig. 6A and C). The full-length translation products of the STE-TM1 containing chimeras are reactive to both globin (lanes 2, and 3) and prolactin (lanes 7 and 8) antisera. Both glycosylated and unglycosylated signal cleaved products were observed when translated in the presence of membranes, corresponding to the two bands observed in lanes 3 and 8. Full length products were cleaved upon treatment with .15mg/ml proteinase K, to generate separate N- and C-terminal fragments, identified by immunoreactivity to globin or to prolactin antisera, respectively. Immunoprecipitation with anti-globin reveals 2 bands (lane 4), corresponding to the glycosylated and unglycosylated N-terminal fragments, respectively; immunoprecipitation with anti-prolactin reveals a single band of 14kD (lane

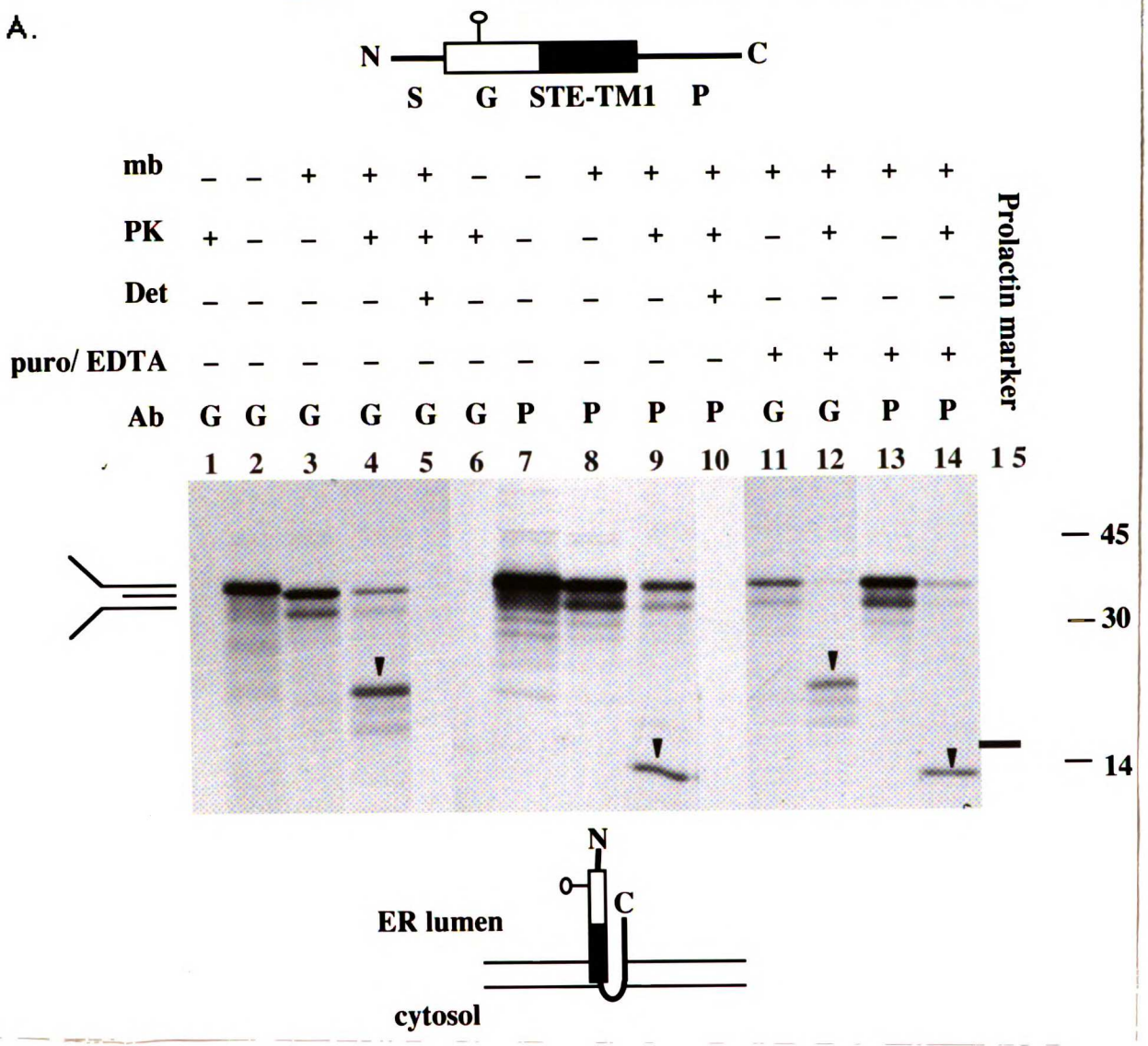
9). Because these bands are slightly larger than would be predicted if the protease were cleaving at the end of TM1, we suspected that a portion of the chimera though to comprise the cytosolic loop was either spanning the membrane or was folded in a manner rendering it less accessible to protease. To address this possibility, parallel digestions were carried out at .25mg/ml or .5mg/ml. When increasing concentrations of PK, the globin immunoreactive bands were digested to generate smaller bands (data not shown), suggesting that the region of the chimera comprising the cytosolic loop is relatively less accessible to protease digestion. Addition of puromycin/EDTA to samples prior to protease digestion to release any translationally paused chains from the ribosome had no effect on the generation of N and C-terminal immunoreactive fragments (Fig. 6A lanes 11-14). As a control for microsome integrity, a secretory protein consisting of the signal sequence of prolactin and codons 87-145 of beta-lactamase (S-L) was mixed with S.G.STE-TM1.P translation products prior to protease digestion and immunoprecipitated with anti-lactamase antibody (data not shown). Lactamase was fully protected in all cases.

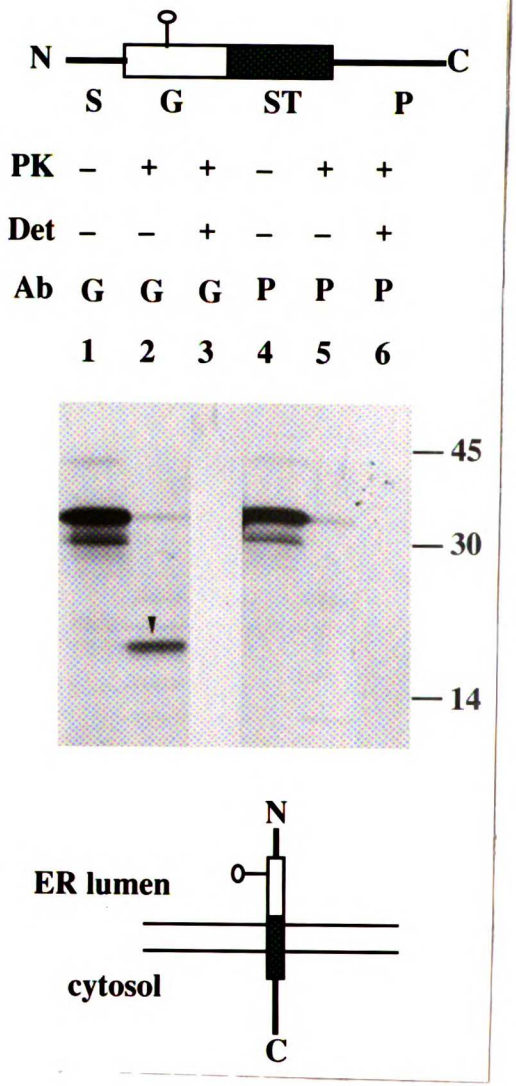
In contrast, when the chimera containing IgM ST was expressed in WG and subjected to PK digestion, only an N-terminal immunoreactive fragment was observed (Fig 6B lanes 2 vs.). Thus, upon expression in WG, STE-TM1 is sufficient to direct a chimeric chain to an orientation spanning the membrane twice with translocated N and C terminal domains similar to the orientation of transmembrane PrP observed in WG previously (106). An interesting feature of these results is that, in the chimeras described, sequences of two unrelated protein domains (globin in PrP-G and prolactin in S.G.STE-TM1.P), appear to be directed to span the membrane simply as a result of their position as the C-terminal domain downstream of STE-TM1, and despite their relative hydrophilicity.

STE and TM1 act together to direct C-terminal translocation

In light of both previous results implicating STE alone as a topogenic sequence (26, 75, 94) and studies on pause-transfer sequences such as found in Apolipoprotein B (24,

Figure 6: Topology of S.G.STE-TM1.P, S.G.ST.P, S.G.STE-ST.P and S.G.TM1.P in WG





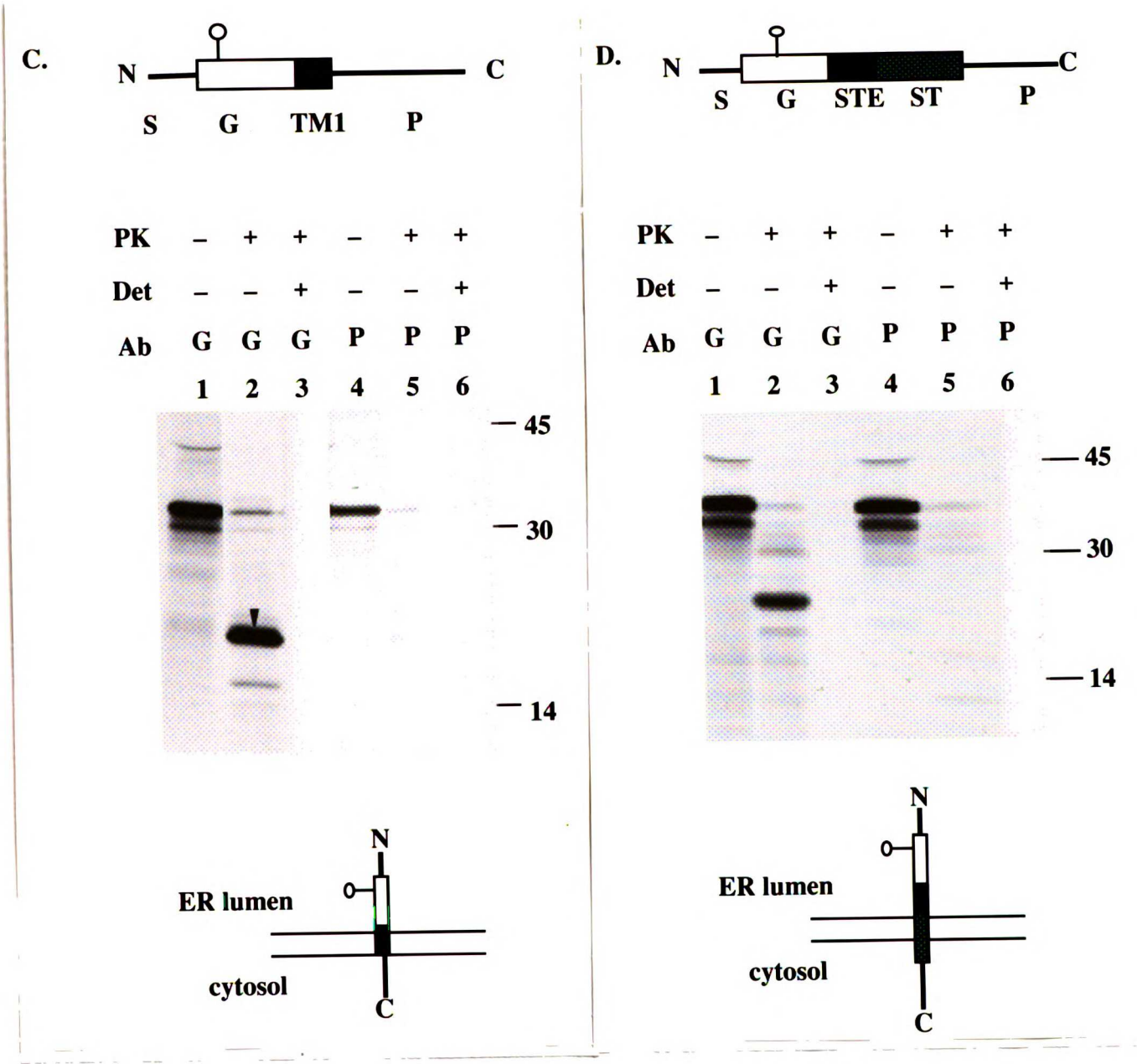


Figure 6: PrP STE-TM1 or IgM ST was engineered into chimeric proteins containing a bovine prolactin signal sequence (S) and chimpanzee globin domain containing an N-linked glycosylation acceptor site (G) and a bovine prolactin domain (P), as shown in the line diagram above each panel, labeled as described for Figure 1. (A) In vitro produced transcript of S.G.STE-TM1.P was translated in WG at 28^{0}C for 1 hour, CaCl2 was added to a final concentration of 10mM, and samples were digested with 0.15 mg/ml PK in the absence (lanes 4 and 9) or presence (lanes 12, 14) of 1mM puromycin and 10mM EDTA, or in the presence of 1% Triton X100 (lanes 5 and 10) for 1 hour at 0^{0}C. Proteolysis was terminated by addition of 10mM PMSF and boiling in 10 volumes of 1%SDS for 5 minutes. Samples were then divided in two aliquots and immunoprecipitated with globin and prolactin specific antisera to detect N- and C-terminal proteolytic fragments. Arrowheads indicate major N and C terminal transmembrane fragments reactive to anti-globin and anti-prolactin sera, respectively. (B-D). In vitro produced transcript S.G.ST.P(B), S.G.TM1.P (C) or S.G.STE-ST.P (D) was translated in WG at 28^{0}C for 1 hour and samples were digested with PK as above in the absence of puromycin/EDTA.

25), we thought that STE alone might be directing the C-terminal translocation event, providing it is followed by a domain as hydrophobic as TM1 that allows the nascent chain to stop translocation and span the membrane the first time. If this were the case, one would expect that if the domain X of the previously described chimera were replaced with TM1 alone, no C-terminal translocation would be observed; while placement of STE in front of the hydrophobic ST would now allow the C-terminus to translocate into the ER lumen. We expressed both of these chimeras, S.G.TM1.P and S.G.STE-ST.P in WG and both were found to generate only a single, N-terminal fragment upon digestion with PK (Fig. 6C and D). These data suggest that neither STE nor TM1 alone is sufficient to direct C-terminal translocation, but must both be present to generate the observed transmembrane topology of PrP.

STE-TM1 appears not to be an internal signal or signal anchor sequence

To explain these findings, we investigated the possibility that STE-TM1 serves both as an ST sequence (thereby terminating translocation of the N-terminal domain) and also as an internal S sequence (to reinitiate translocation of the subsequent C-terminal domain). Previously we have used a family of related protein chimeras to study the activity of internal S or SA sequences in the biogenesis of multispanning membrane proteins (20). This family of protein chimeras is of the structure S.L.ST.G.X.P, where S is a signal sequence that is cleaved upon chain translocation, L is codons 83 to 181 of beta lactamase, G and P are the codons for the globin and prolactin domains as described respectively in figure 2, and X represents the coding region for either an S or an ST sequence or no inserted codons. Expression and post-translational proteolysis of these chimera allowed us to assess the ability of sequence X to reinitiate translocation, by scoring the percent of chains in which the P domain was translocated into the ER lumen as determined by protection from protease digestion (20). Previously we have shown that the G and P domains are quantitatively localized to the cytosol when no coding sequence is present in

location X such that the encoded protein chimera is S.L.ST.G.P (20). When the coding sequence for S is placed between those of the G and P domains, creating the coding region S.L.ST.G.S.P, translocation can be reinitiated subsequent to the ST sequence. As a result, the protein spans the membrane twice, with both G and P domains localized to the ER lumen, consistent with the previously described activity of an internal S sequence (67, 113). This conclusion was based on the demonstration of glycosylation of the G domain and the presence of separate protected fragments generated by protease digestion that were reactive to anti-lactamase or to both anti-globin and anti-prolactin serum, for those chains in which the internal S sequence was not cleaved by signal peptidase (20). When codons for an ST sequence are placed between those of the G and P domains in S.L.ST.G.X.P, creating the coding region S.L.ST.G.ST.P, a much lower level of translocation of the P domain is seen. For the small fraction of chains in which P domain translocation is observed, the second ST behaves like an internal SA sequence, translocating the P but not the G domain, consistent with the conclusions from other studies (20). Thus, translocation of the P domain in these chimeras has previously been used as a reporter of both internal S and SA sequence activity (19). By engineering STE-TM1 in the position of X and expressing the resulting chimera S.L.ST.G.STE-TM1.P in XO as previously described (see Fig 5), we could determine if STE-TM1 displayed S or SA activity by comparison to the chimeras in which X represents S, ST or no sequence.

The chimera S.L.ST.G.STE-TM1.P was engineered and expressed in XO and its topology assessed in parallel with the matched constructions described above. The extent of P translocation, scored by quantitation of the protected P immunoreactive fragment generated after PK digestion, and normalized for methionine content, served as a reporter of internal S/SA activity, as described previously (19). As shown in Fig. 7, no significant translocation reinitiation activity was observed for the chimera containing STE-TM1, comparable to the chimera with no codons inserted between those of G and P, and in striking contrast to constructions in which X represented S or ST (20). Taken together, the

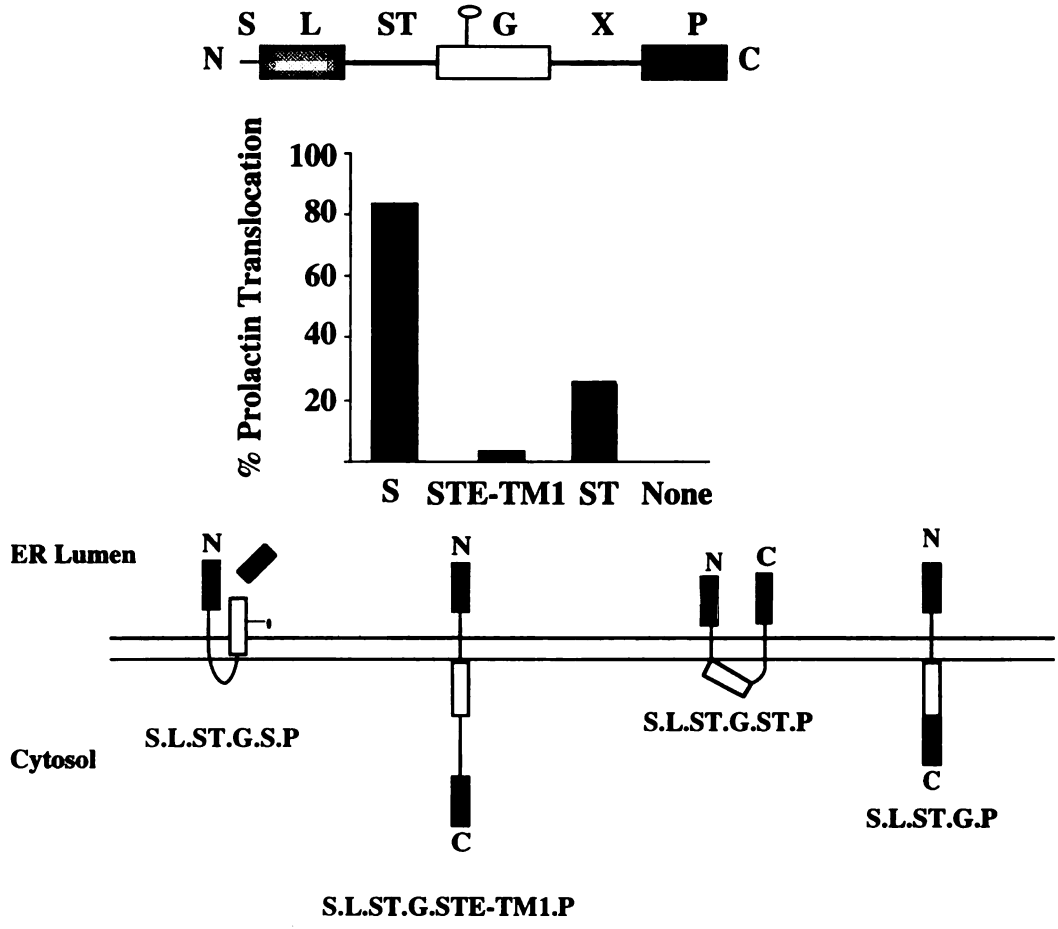
behavior of these chimeras argues that STE-TM1 is not an internal S or SA sequence capable of reinitiating translocation once it has stopped in the membrane. However, if STE-TM1 does not contain an internal S/SA activity, how do we account for its ability to direct C terminal domain translocation (see figure 6)? Perhaps the mechanism used by STE-TM1 to stop translocation is different from that of a conventional ST sequence. For example, C-terminal domain translocation in chimera such as PrPG and S.G.STE-TM1.P may be able to occur without a subsequent S or SA sequence because STE-TM1 stopped the chain without fully closing the initial translocation channel, in contrast to the action of a conventional ST. Since the relative inefficiency of targeting and translocation in the cell-free system might be a confounding variable in attempts to test this hypothesis, we decided to carry out all further studies in microinjected XO.

C-terminal domain translocation is directed by the STE-TM1 sequence in XO

We first confirmed that C-terminal domain translocation in the chimera S.G.STE-TM1.P as shown in Fig. 6 in WG, also occurs *in vivo* . XO were microinjected with S.G.STE-TM1.P transcript and ^{35}S methionine as previously described and after a 1 hour pulse in the presence of radiolabeled methionine, were homogenized and digested with PK in the absence or presence or detergent. Aliquots were immunoprecipitated with antisera directed to globin (Fig. 8A lanes 1-3) or prolactin (Fig 8A lanes 4-6). Quantitative densitometry, corrected for methionine distribution, revealed that 60% of the full-length chains spanned the membrane with separate N and C terminal translocated domains.

In view of these results, it was puzzling that we had not previously observed transmembrane chains of native PrP in either brain or heterologous cells (2, 13). The suggestion (see above) that STE-TM1 stops translocation in a manner different from that of a conventional ST sequence is one way to account for both C-terminal domain translocation in the absence of an internal S/SA sequence (see Fig 7, above), and also for the previous failure to detect transmembrane PrP *in vivo* (97, 108). If stop transfer

Figure 7: Internal Signal Sequence Activity of S.L.ST.G.X.P



Chimera of the structure S.L.ST.G.X.P as described in text, were transcribed *in vitro*, microinjected into XO and subjected to proteolysis and with domain specific antisera as described for Fig. 1C. Results were analyzed by quantitative densitometry, using translocation of prolactin domain, demonstrated by the generation of a prolactin -reactive proteolytic fragment, normalized to the amount of lactamase-reactive fragment, as a measure of internal signal sequence activity (19). This data is displayed as a bar graph accompanied by cartoons depicting the respective topology of each chimera.

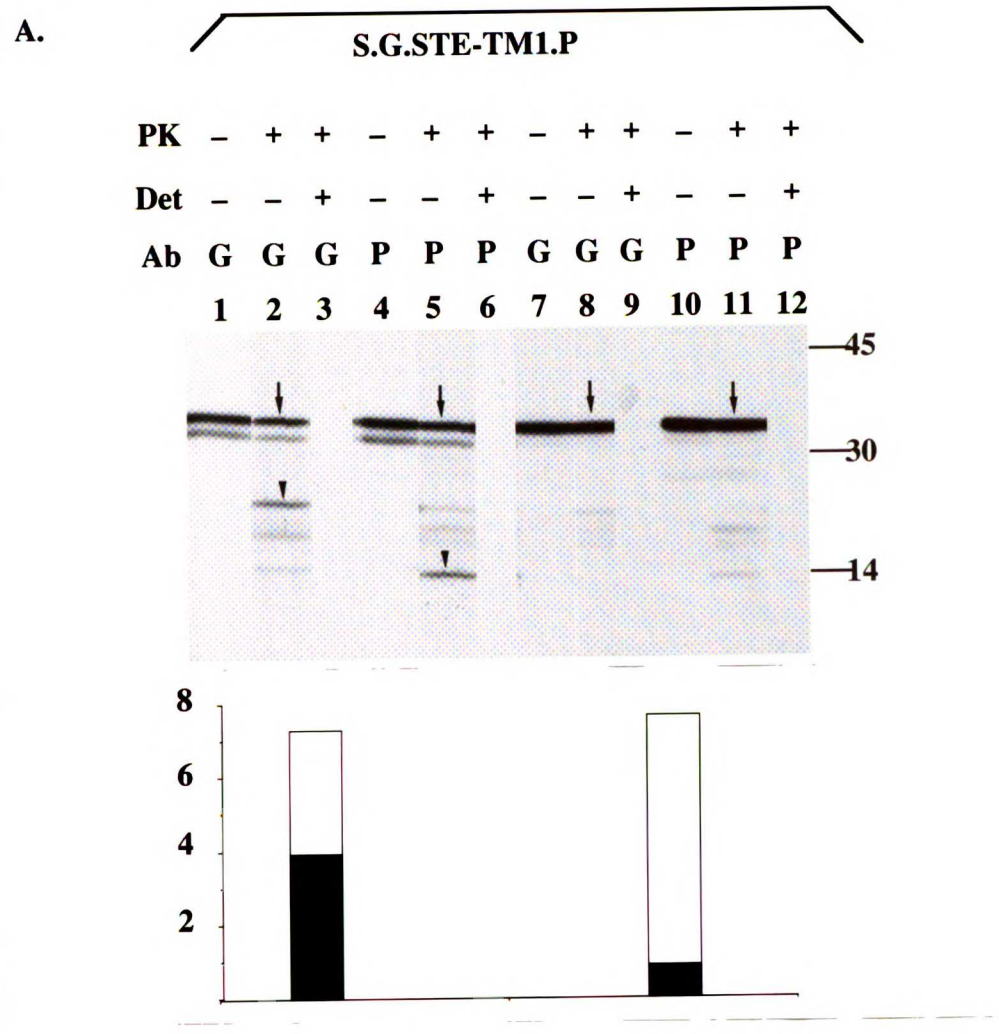
directed by STE-TM1 differs from that directed by a conventional ST in that closure of the translocation channel is delayed, then C-terminal domain translocation could proceed through the open channel without requiring the action of a subsequent S or SA sequence. According to this model, upon chain completion, the polypeptide would span the membrane twice but reside entirely within a single aqueous translocation channel. In PrP, where TM2 is an amphipathic helix, such an isoform could integrate into the bilayer, provided that channel disassembly occurred prior to release of the STE sequence from putative receptor proteins and movement of the chain from the channel into the ER lumen. If instead, channel disassembly were further delayed, only the fully translocated form would be observed. This hypothesis makes two predictions as to the behavior of STE-TM1-containing chimeras. A first prediction is that those chains which span the membrane twice should, under certain circumstances, chase into fully translocated chains. A second prediction is that chains of the STE-TM1-containing chimera which span the membrane twice should be extractable by sodium carbonate at alkaline pH, because they are proposed to reside in an aqueous, protein-lined channel rather than being integrated into the membrane (3, 64).

An STE-TM1-containing chimera transiently spans the membrane twice in XO

To test the first of these predictions, XO microinjected with S.G.STE-TM1.P and pulsed as described above, were chased for 4 hours following a second injection of unlabeled methionine. XO were then homogenized and digested with PK as above. The total synthesis of full-length PrP chains was comparable between both pulse and chase time points (compare intensity of total chains in Fig 8A lanes 1 vs. 7 or 4 vs. 10). The doublet observed at the pulse time point (Fig 8A lanes 1-6) represents glycosylated and unglycosylated chains Upon chase, nearly all chains become glycosylated (Fig 8A lanes 7-10). Consistent with the above hypothesis, almost all of the chains which had scored as transmembrane at pulse were found, upon chase, to be fully translocated and protected

from protease (Fig. 8A lanes 8 and 11 compared to lanes 2 and 5). When samples are treated with PK, 4 globin immunoreactive and 3 prolactin immunoreactive fragments are observed. As was observed with the WG translation products of this same chimera, the globin immunoreactive fragments are slightly larger than would be predicted; conversely, the prolactin immunoreactive fragment is slightly smaller than the full prolactin domain contained in the chimera. The small amount of globin immunoreactive fragments that are immunoprecipitated with the anti-prolactin antibody can be accounted for by the presence of the first few amino acids of the prolactin domain in the cytosolic loop, which is relatively inaccessible to protease. Treatment of these same samples with Endo H, which trims off carbohydrate modifications, causes this upper band to shift down in both the anti-globin and anti-prolactin immunoprecipitated samples; whereas the predominant lower anti-prolactin reactive band is unaffected by Endo H treatment (data not shown). Since the only glycosylation site is in the very N-terminus of the molecule, sensitivity to Endo H of the upper weakly anti-prolactin reactive fragment suggests that the proteolytic cleavage event occurs a few residues into the prolactin domain. No anti-globin immunoreactivity is observed in the 14kD fragment corresponding to the C-terminally translocated domain, although the unglycosylated fully digested N-terminal fragment appears to comigrate on some gels. To control against a subpopulation of leaky vesicles, the secretory control, S-L, used in Fig. 6 was homogenized an proteolyzed with S.G.STE-TM1.P; S-L was fully protected under all conditions (Fig. 8D). To confirm that this behavior reflected differences between the action of STE-TM1 and ST, the matched chimera containing ST instead of STE-TM1 was analyzed in parallel pulse-chase experiments in XO. As shown in Fig 8B, only globin-reactive but not prolactin-reactive, fragments were generated by PK digestion (Fig. 8B lanes 2 vs. 5), and all chains, at both pulse and chase, were transmembrane in the identical orientation (Fig.8B lane 2 vs. 8). To demonstrate that this membrane-spanning intermediate is not a product of expression of any secretory form of this chimera in XO, a corresponding chimera containing a S sequence in place of STE-TM1 was found to be fully

Figure 8: Pulse-chase analysis of S.G.STE-TM1.P and S.G.ST.P in XO.

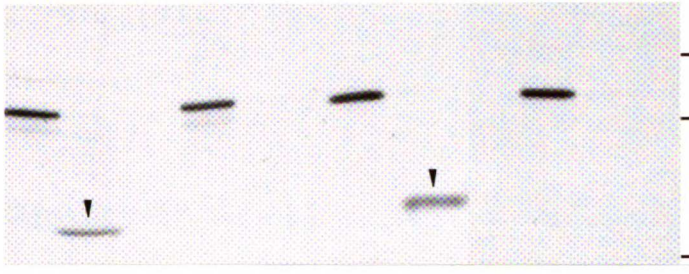


In vitro transcribed S.G.STE-TM1.P (A), S.G.ST.P (B), and S.G.S.P (C) were microinjected into XO in the presence of ^{35}S methionine and incubated for 1 hour at 18^0C. Parallel samples of A and B were then processed immediately (pulse, lanes 1-6) or injected with excess unlabeled methionine and incubated for 4 hours at 18^0C (chase, lanes 7-12). XO were homogenized and aliquots digested with PK at 0.5mg/ml for 1 hour. Proteolysis was terminated as described (113). Samples were divided into equal aliquots and immunoprecipitated with anti-G (N-terminal) and anti-P (C-terminal) sera and analyzed by SDS-PAGE and F. Arrowheads indicate N- and C-terminal protected fragments and arrows indicate full-length chains after PK digestion. Data was analyzed by densitometry and shown in the graph below panel A. Total chains are indicated on the y-axis, expressed in arbitrary units reflecting band density. Total density was calculated as (mean density -background) X (total pixels). Band density was calculated as total density normalized for methionine content. At pulse the total band density of full-length chains (glycoslylated + unglycosylated /12 methionines) was 7.7. At chase this value was 7.75. Densitometry of proteolyzed lanes showed that at pulse, density of both globin-reactive and prolactin -reactive full length products was 2.95, and density of globin-reactive and prolactin-reactive transmembrane fragments was 4.1 and 3.85, respectively. At chase, the density of both globin and prolactin full-length protected band was 7.2 and density of transmembrane fragments was 0.6. Secretory chains are shown as the open portion of the bar, and transmembrane chains as the filled region. (D) S.L was microinjected into XO as described above and homogenized with S.G.STE-TM1.P at both pulse and chase to control for vesicle integrity. Immunoprecipitations with anti-lactamase antibody are shown in panel D.

B.

S.G.ST.P

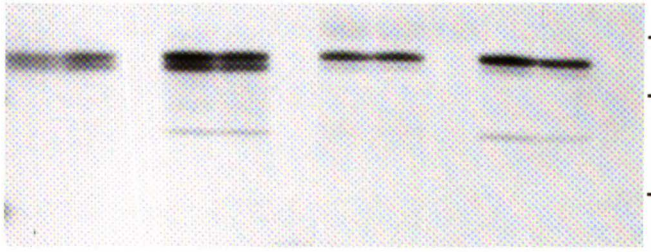
PK	–	+	+	–	+	+	–	+	+	–	+	+
Det	–	–	+	–	–	+	–	–	+	–	–	+
Ab	G	G	G	P	P	P	G	G	G	P	P	P
	1	2	3	4	5	6	7	8	9	10	11	12



C.

S.G.S.P

PK	–	+	+	–	+	+	–	+	+	–	+	+
Det	–	–	+	–	–	+	–	–	+	–	–	+
Ab	G	G	G	P	P	P	G	G	G	P	P	P
	1	2	3	4	5	6	7	8	9	10	11	12



D.

Lactamase

PK	–	+	+	–	+	+
Det	–	–	+	–	–	+

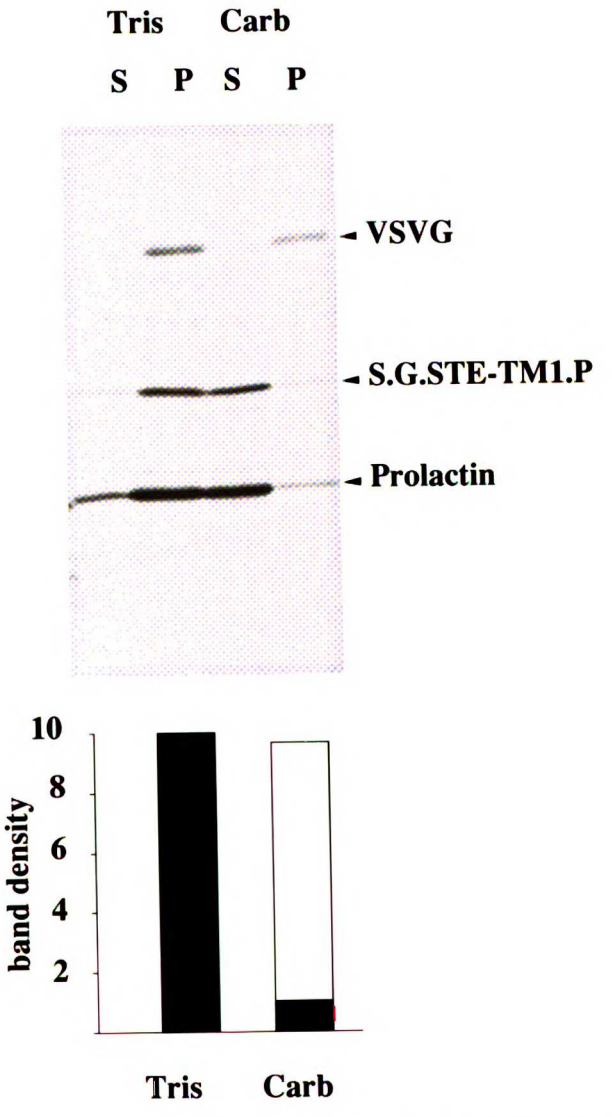
translocated, as demonstrated by complete protection upon proteolysis at both pulse and chase points (Fig.8C). The results of figure 8B and 8C suggest that our findings are specific to STE-TM1 and not due to the sequence context of the chimera itself. These findings serve not only to clarify the paradox of the existence of two topologically distinct forms of PrP, but also to distinguish the action of STE-TM1 from that of a conventional ST sequence.

STE-TM1 chimera spanning the membrane twice are not integrated

We then tested the second prediction of the above model, namely that the transmembrane chains observed to span the membrane twice at pulse, would be found to be accessible to extraction by aqueous perturbants such as sodium carbonate at alkaline pH. Chains of S.G.STE-TM1.P expressed in XO under pulse conditions as described in Figure 8 (i.e. in which 60% of chains score by proteolysis to be spanning the membrane twice), were subjected to extraction with either Tris, pH 7.5, or sodium carbonate, pH 11.5. This procedure has been used to distinguish proteins integrated in the lipid bilayer from those that are accessible to aqueous perturbants, as would be expected for a protein residing within the translocation channel (24, 25, 92). Additional XO were microinjected in parallel with transcript for known secretory (bovine prolactin, see ref 20) and integral transmembrane (Vesicular Stomatitis Virus G, see ref.76) proteins and ^{35}S methionine. After a one hour pulse incubation, XO separately expressing S.G.STE-TM1.P and control proteins were homogenized, aliquots were mixed, and subjected to extraction. Figure 9 demonstrates that essentially no S.G.STE-TM1.P chains were found in the carbonate pellet, comparable to that of the secretory control. In contrast, VSVG (the integral membrane control) was found exclusively in the pellet. These results demonstrate that the transmembrane chains of S.G.STE-TM1.P, shown in figure 8A to span the membrane twice, are not integrated into the membrane. Thus, while the action of an ST results in the termination of chain translocation and immediate integration into the membrane, the action

of STE-TM1 directs termination of the N-terminal domain translocation without concomitant integration. As a consequence of these differences in action, protein domains C-terminal to STE-TM1 are able to enter the translocation channel, thereby generating both a second transmembrane region and a translocated C-terminal domain, upon completion of protein synthesis. Because these chains reside entirely within the translocated channel they are not only subject to carbonate extraction (figure 9) but also can be converted into fully translocated chains upon chase (figure 8A).

Figure 9:
Carbonate Extraction of S.G.STE-TM1.P



Homogenate of XO expressing either S.G.STE-TM1.P or prolactin and VSVG pulse-labelled as described for figure 4A, were mixed together and diluted 330-fold in Tris Cl pH 7.5 and 0.25M sucrose or 0.1M sodium carbonate, pH 11.5. Samples were centrifuged at 70,000rpm for 30 minutes, supernatants were removed and proteins precipitated with TCA, and both TCA pellets and Tris and carbonate pellets were solubilized in 0.1M Tris pH 8.0 and 1.0 %SDS, diluted with 20 volumes of TXSWB, immunoprecipitated with VSVG and prolactin antisera as described and analyzed by SDS-PAGE. Results were analyzed by densitometry and graphed as shown. Band density of S.G.STE-TM1.P in Tris pellet was 9.75 and density of the same band in the carbonate pellet was 1.35. The band density of the carbonate supernatant was 8.2. Filled region represents the density of radiolabeled band recovered in Tris or carbonate pellets and open region represents the density of radiolabel recovered in Tris or Carbonate supernatants.

III. DISCUSSION

Downstream domain translocation without additional topogenic sequences

Studies on other multi-spanning membrane proteins have elucidated a mechanism by which a series of S and ST or SA sequences allow a nascent polypeptide chain to stop translocation at one transmembrane domain, integrate the chain into the bilayer, then reengage the translocation apparatus and resume translocation until another transmembrane domain is reached (13, 20, 110). In these proteins the chain is integrated even before the internal S or SA sequence has been translated (18, 20). Thus it seems likely that the action of a conventional ST sequence is to both stop further translocation of the chain and also close the translocation channel. Under these circumstances, subsequent translocation of a C-terminal domain requires that the channel be reopened or reassembled, hence the requirement for an internal S or SA sequence. The data presented here suggest a different way by which multispanning topology can be achieved, as may occur in the case of PrP. The STE-TM1 sequence of PrP stops translocation of the N-terminal domain in a way that allows subsequent translocation of the C-terminal domain without an additional S or SA sequence. We cannot formally eliminate the possibility that an additional S or SA sequence exists in the C-terminus of PrP. However, if such a sequence exists, it would appear to be redundant and unnecessary to account for the orientation of transmembrane PrP, since information capable of directing C-terminal domain translocation has been localized entirely within STE-TM1. Moreover even if an additional S or SA sequence does exist in the C-terminal domain of PrP, its action would not be sufficient to explain the transmembrane orientation of chimeras such as PrP-G and S.G.STE-TM1.P. In both of these proteins, unrelated C -terminal passenger domains (derived from globin and prolactin, respectively, which have been shown to lack internal S or SA sequences, see ref. 20), are demonstrated to be directed across the ER membrane when engineered downstream of STE-TM1. The region of PrP just C-terminal to TM1, which encompasses the P-5 epitope, is approximately 13 amino acids longer in the chimera PrP-G. As is

mentioned in Figures 1, 2 and 4, this region is relatively inaccessible to proteolysis, most probably due its predicted orientation as a hairpin loop at the cytosolic end of the translocation channel (see Fig 10e). In the case of PrP-G, where more of the P-5 epitope is present, the transmembrane form is oriented such that proteolytic cleavage occurs closer to TM1 and anti-P5 immunoreactivity is found on the C-terminal fragment. When a smaller portion of this loop region is present, as in S.G.STE-TM1.P, the orientation is such that cleavage occurs a few residues into the prolactin domain and anti-P5 as well as weak anti-Prolactin immunoreactivity are found on the N-terminal fragment.

Studies by Nakahara, et al. suggest that STE alone can direct a nascent chain to pause in the membrane (94); however, subsequent events in the translocation of PrP appear to require both STE and TM1. Furthermore, the inability of STE to direct C-terminal translocation when followed by a random membrane domain suggests that TM1 serves a function beyond merely providing sufficient hydrophobicity to span the membrane. On the other hand the inability of TM1, when its hydrophobicity is extended to allow for an STE-independent membrane spanning topology, to direct C-terminal translocation suggests that this information is not located entirely within the membrane domain. The preponderance of helix breaking glycine residues within TM1 may render the membrane domain flexible enough to undergo a conformational change directed by STE, or a distinct topogenic sequence, comprised of a region overlapping STE and TM1, may direct C-terminal domain translocation. Although our data suggest that STE alone cannot direct this second translocation event, even when followed by a known transmembrane domain such as ST, we cannot rule out the possibility that there may exist a subset of hydrophobic domains which, in combination with STE, are capable of doing so. One possible explanation for the inability of STE-ST to direct C-terminal translocation is that ST somehow directs the translocation channel to close before C-terminal translocation can occur.

STE-TM1: A variation on the theme of conventional stop transfer action

One interpretation of these findings is that STE-TM1 translocates its downstream domains because it fails to carry out all the actions of an ST sequence. A two-step model has been proposed by which ST sequences allow transmembrane domains to become integrated into the membrane (64). According to this model, conventional ST sequence action involves stopping chain translocation within the translocation channel (e.g. through engagement of an ST sequence by a receptor), followed by disassembly of the channel (accompanied by receptor disengagement), thereby allowing the chain to integrate into the lipid bilayer. A conventional ST sequence directs both of these steps in such rapid succession that it is difficult to distinguish them except by the most sensitive experimental procedures (50). In contrast, the action of STE-TM1 appears to be analogous to that of ST with respect to the first step, i.e. stopping the nascent chain in the channel, but not with respect to the second step. Instead of disassembly followed by membrane integration, the channel that has engaged STE-TM1 appears to remain open, even while translocation of the chain within it has stopped. Thus, the polypeptide domain C-terminal to STE-TM1, emerging as part of the nascent chain from the ribosome, can spontaneously translocate via the open channel into the lumen of the ER. This model assumes that the dimensions of the channel are large compared to those of the translocating chain that is, that the presence of the stopped N-terminal limb of the nascent chain is not an absolute steric hindrance to entry to a C-terminal loop into the same channel. This assumption seems reasonable since i) there is no direct evidence as to the size of the translocation channel, ii) the upper limit of channel size, as defined by the size of a ribosome (20 nm) is enormous compared to an individual unfolded polypeptide chain, and iii) indirect evidence from ion conductance suggests the channel has a large dimension (3). Consistent with this hypothesis, the translocation of domains downstream of STE-TM1 in the chimeras described here, result in

domains that would not have been predicted to be transmembrane by hydropathy, being found to span the membrane (see figures 5, 6 and 8A).

An intermediate spanning the membrane twice gives rise to fully translocated chains

A second outcome of these studies is the demonstration of a precursor-product relationship between carbonate-extractable chains that span the membrane twice, and fully translocated chains, *in vivo*. This feature poses a plausible explanation for the discrepancy between the transmembrane topology predicted for PrP by hydropathy analysis, and experimentally demonstrated in the wheat germ cell-free system (106, 107), and the fully translocated form observed in normal brain, XO and RRL (108, 75). We propose that some, if not all, chains of PrP^C achieve their mature fully translocated orientation with respect to the membrane, via conversion from this transmembrane topologic intermediate. These and previous findings (76) suggest that the machinery of chain translocation may be used to influence the pathway of folding utilized by a given nascent chain. As has been suggested for other forms of novel post-translational regulation of gene expression , this level of control may occur in response to cues from the environment (115).

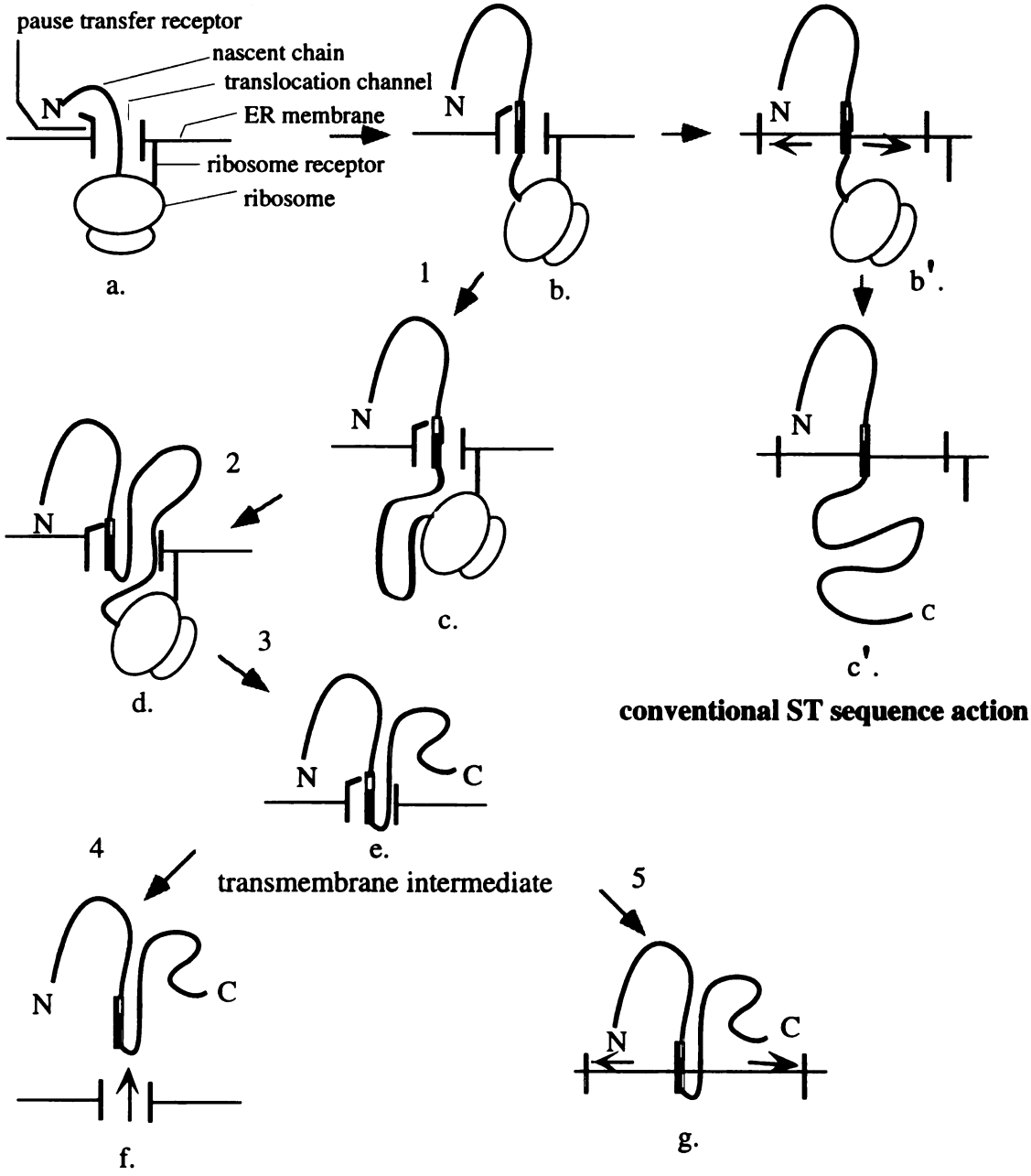
A model for PrP topogenesis

Our studies allow us to propose a model for PrP chain translocation that reconciles the results of studies *in vitro* and *in vivo* (see figure 10). According to this view, nascent PrP initiates translocation in a manner indistinguishable from a conventional secretory protein (Fig. 10 a). Upon engagement of STE by receptors in the ER membrane, translocation stops (Fig. 10b). However, unlike a conventional ST sequence (Fig. 10b'), receptor engagement is not accompanied by channel disassembly. Perhaps this is because some aspects of the ribosome-membrane junction are maintained upon action of STE-TM1, but not ST. Alternatively, perhaps STE-TM1 directs only a subset of the receptor-mediated

actions that occur upon binding of receptors by ST. A third possibility is simply that the kinetics of STE action are slower than that of a conventional ST sequence, thereby allowing a longer window of time during which the stopped chain resides in the translocation channel, and during which alternate pathways of biogenesis are possible. Regardless of the precise mechanism, we propose that receptor engagement by STE-TM1 is not accompanied by immediate channel disassembly (Fig.10b). Thus as the C-terminal domain of the nascent polypeptide emerges from the ribosome, (Fig. 10c), it can enter the open channel (Fig. 10d). Upon termination of protein synthesis, STE-TM1-containing chains reside within the translocation channel in an orientation with two membrane-spanning domains (Fig. 10e). In contrast, chain growth subsequent to receptor engagement by a conventional ST sequence remains permanently in the cytosol because the channel is disassembled too rapidly for a C-terminal domain to be synthesized and translocated (Fig.10c'). The final topologic outcome for STE-TM1-containing chains may be determined by a number of factors. Either channel disassembly and chain integration or receptor binding of STE may serve as a sort of cellular clock, defining the amount of time allowed for the chain to translocate fully. Thus, a membrane-integrated topology with two membrane-spanning regions might occur (as is observed for PrP in WG) if channel disassembly occurs before completion of translocation (see Fig. 10g). Another determinant of topology may be the presence of cytosolic and/or membrane factors which affect the relative rates of STE release from a receptor versus channel disassembly. Perhaps under normal circumstances fully translocated chains are the only form of PrP observed, because such factors (e.g. which likely exists in the cytosol of RRL, see ref 76) delay channel disassembly or speed release of STE from a receptor, allowing chain translocation to complete (see Fig. 10f). It is tempting to suggest that, under certain physiological or pathological circumstances, channel disassembly occurs with sufficient rapidity (or conversely receptor disengagement of the chain is sufficiently slowed) *in vivo* to generate integrated transmembrane chains, as are observed in WG (Fig. 10g). The existence of

mutants which appear to alter the kinetics of conversion from transmembrane to fully translocated PrP and the recently constructed PrP null mouse (124), should allow testing of hypotheses relating these observations to events which occur in the pathogenesis of prion diseases.

Figure 10: STEPS IN PrP BIOGENESIS



Events 1-6 occuring during PrP translocation from cytosol to ER lumen across the ER membrane are as described in text for steps 1-6 and products a-g. Events for a conventional integral membrane protein are as described in events 1, 2' and 2" and products a, b, b' and c'. Black arrows indicate movement of components (e.g. receptors in the plane of the membrane in panels b' and g; the completed, partially folded PrP chain out of the translocation channel in panel f). Gray arrows (labelled 1-5) indicate the order of events in biogenesis of PrP versus a conventional integral membrane protein.

CHAPTER 3: MUTATIONS OF STE-TM1 AFFECT PrP TOPOLOGY

I. Introduction

In Chapter 2, it was demonstrated that, under normal cellular conditions, PrP achieves its final fully translocated topology via a transient transmembrane intermediate that spans the membrane twice. The information directing this translocational mechanism lies within a 58 amino acid sequence, consisting of a hydrophobic region that serves as the first membrane domain (TM1) and the relatively hydrophilic sequence just N-terminal to it known as Stop Transfer Effector (STE). The second membrane spanning domain appears to do so merely as a result of its location downstream of STE-TM1, as evidenced by the ability of STE-TM1 to direct two unrelated passenger domains to span the membrane when placed in heterologous chimeras . The mechanism of translocation utilized by PrP to achieve a fully translocated topology is reminiscent of that observed with proteins containing "pause-transfer" sequences such as Apolipoprotein B(Apo B), and, in fact, appears to be a variation on this theme. While a hydrophilic sequence containing some degree of homology directs both proteins to pause in the Endoplasmic Reticulum (ER) membrane during translocation (25, 94), PrP STE is followed by a hydrophobic region capable of spanning the membrane (TM1), whereas the corresponding element in Apo B (B') is not. Thus, under certain experimental conditions, PrP can be observed in a transmembrane form. It has been demonstrated that PrP topology in cell-free systems depends on the cytosolic lysate used. When translated in Rabbit Reticulocyte Lysate (RRL) supplemented with dog pancreas microsomes, PrP is predominantly secretory; while, when translated in Wheat Germ Lysate (WG) with the same membranes, it is predominantly transmembrane (106). From this standpoint, the action of STE-TM1 in PrP is also a variation on the theme of a classic stop transfer (ST) sequence responsible for directing topology of integral membrane proteins. While both consist in part of hydrophobic

domains that span the membrane, ST directs only a singly membrane spanning topology and is able to trigger disassembly of the translocation channel thereby resulting in membrane integration (20). On the other hand, STE-TM1 directs the initial membrane spanning topology, as well as the translocation of the C-terminus of the molecule, and does not trigger channel disassembly, thus allowing the chain to continue translocation. An integrated transmembrane form of PrP may never be observed under normal cellular conditions (108), however regulation of the kinetics of translocation at the level of a transmembrane intermediate, either resulting in a transmembrane molecule or affecting the folding of the molecule in some other manner, may be of significance in the generation of the altered form of PrP (PrP sc) associated with prion diseases.

A number of germ line mutations of the PrP gene have been found in animals with prion diseases. One such mutation found in some humans who contract GSS, a Proline to Leucine mutation at codon 105, has been shown to result in *de novo* development of Scrapie when expressed in transgenic mice (116). A number of other mutations associated with GSS appear to occur within the region defined as STE-TM1 in Chapter 2 (117). Given this ability of a genetically altered prion protein to cause Scrapie, and in light of data suggesting the prion protein is handled by the cell in a somewhat unconventional manner, we wished to further define the amino acid residues responsible for directing the transient transmembrane intermediate. We have created deletions as well as site specific mutants of STE-TM1. The resulting proteins have been analyzed for their ability to show a system dependent difference in topology, i.e. transmembrane topology when translated in wheat germ lysate and secretory topology when translated in rabbit reticulocyte lysate and translocation of the C-terminus resulting in the second membrane spanning domain. In fact, the region defined as STE has been reduced to a ten amino acid sequence just adjacent to TM1 and several point mutations have been identified within STE-TM1 that alter the topology of PrP in both cell-free and *in vivo* systems, one of which corresponds to a mutation observed in some patients who develop GSS (117).

II. Results:

To further characterize the topogenic element STE-TM1 and the mechanism of translocation it directs, we wished to narrow down the amino acid residues necessary for its function and identify point mutations that alter its activity. The rationale for such experiments is severalfold. First, from the perspective of general cell biology, the identification of several pause-transfer proteins, such as Apo B and GRP 94, that appear to possess a similar mechanism of translocation and show some degree of sequence similarity in their respective topogenic elements, points to the possibility of identifying common residues required for pause-transfer and initial STE-TM1 directed events (118). Second, studies by other groups have identified point mutations within this region in patients with the prion disease, Gerstmann Straussler Shenkner Syndrome (GSS), suggesting a possible role for this particular region of the PrP protein in the pathogenesis of associated neurodegenerative disorders (117). Third, spectroscopic studies on PrP have identified a region denoted H1 (consisting of codons 109-122), that has a propensity to shift from α-helical to β-sheet conformations under conditions associated with Scrapie (105). This same region encompasses the region defined as TM1 and part of STE, further suggesting an important role for conformational events directed by this region in the pathogenesis of Scrapie. Finally, narrowing down the region necessary for STE-TM1 action will provide a tool with which to observe the effect of loss of STE-TM1 activity, by making the small deletions and point mutations in the PrP gene.

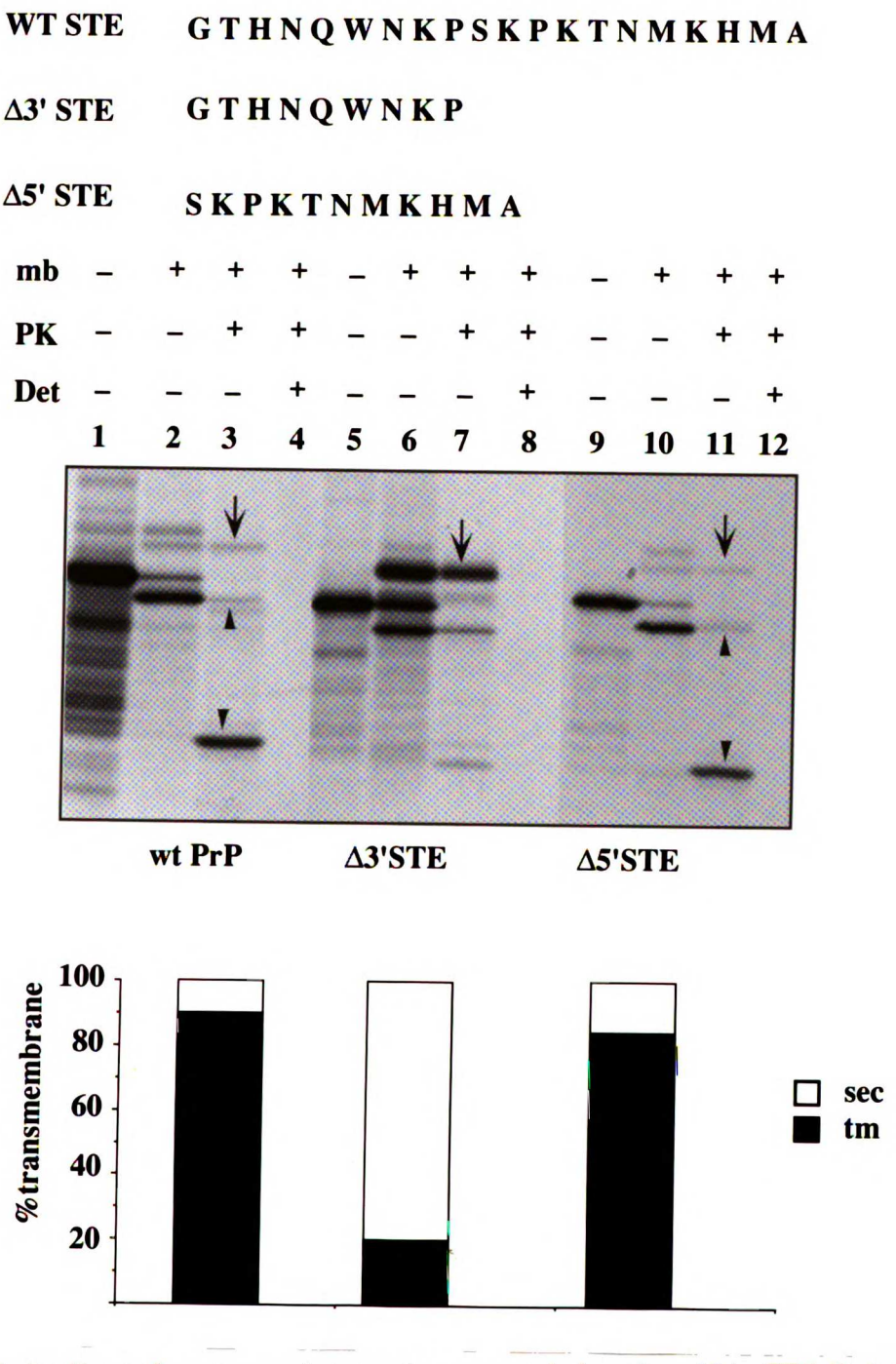
Deletion of codons 102-113 of hamster PrP result in a lack of transmembrane chains.

Unique restriction sites, KpnI and Csp45I, were engineered at nucleotides 280 and 310, respectively, of plasmid PrP, using site directed mutagenesis. Digestion of the resulting plasmid, PrP30, with Csp45I cleaves PrP between codons 102 and 103 located

within the region defined as STE-TM1, and digestion with KpnI cleaves between codons 94 and 95 located at the beginning of STE-TM1. A pre-existing NaeI site, which cleaves between codons 113 and 114 had previously been used to define the end of STE and the beginning of TM1. A Kpn I to Csp45I fragment, consisting of codons 94-102, or a Csp45I to NaeI, consisting of codons 103 to 113, was deleted to generate either PrPΔ5' or PrP Δ3', respectively (Figure 11A). The resulting plasmids were transcribed *in vitro*, and translated in either wheat germ (WG) or Rabbit Reticulocyte Lysate (RRL) in the presence and absence of dog pancreas microsomes for 1 hour. Samples of translated product were subjected to treatment with .25mg/ml Proteinase K for 1 hour at 0°C before termination of reaction with .01MPMSF followed by 10 fold dilution in boiling 1%SDS. Total products of this reaction are shown in Figure 11B. Full-length product of the native PrP molecule translated in the absence of membranes migrates at approximately 30kD on SDS-PAGE, whereas the corresponding translation product of each of the two deletions migrates slightly lower (compare Fig. 11B lanes 1 vs. 5 and 9). In the presence of membranes, both glycosylated and unglycosylated signal cleaved products are observed as well as some untranslocated protein corresponding with chains that did not engage translocation apparatus (Fig. 11B lanes 2, 6 and 10). Note that with both native PrP as well as the deletion PrPΔ5', a doublet of glycosylated, signal cleaved product is observed.* Upon treatment with proteinase K, full-length chains are digested generating membrane protected fragments corresponding to 80-90% of total translocated chains, with native PrP as well as PrPΔ5' (Fig.11B lanes 3 and 11). In contrast, 75-80% of full length translocated chains observed for PrPΔ3' remain fully protected upon proteolysis (Fig. 11B lane 7). When membranes are permeabilized with non-denaturing detergent TritonX-100 prior to proteolysis, no full-length or protein fragments are observed, demonstrating that the

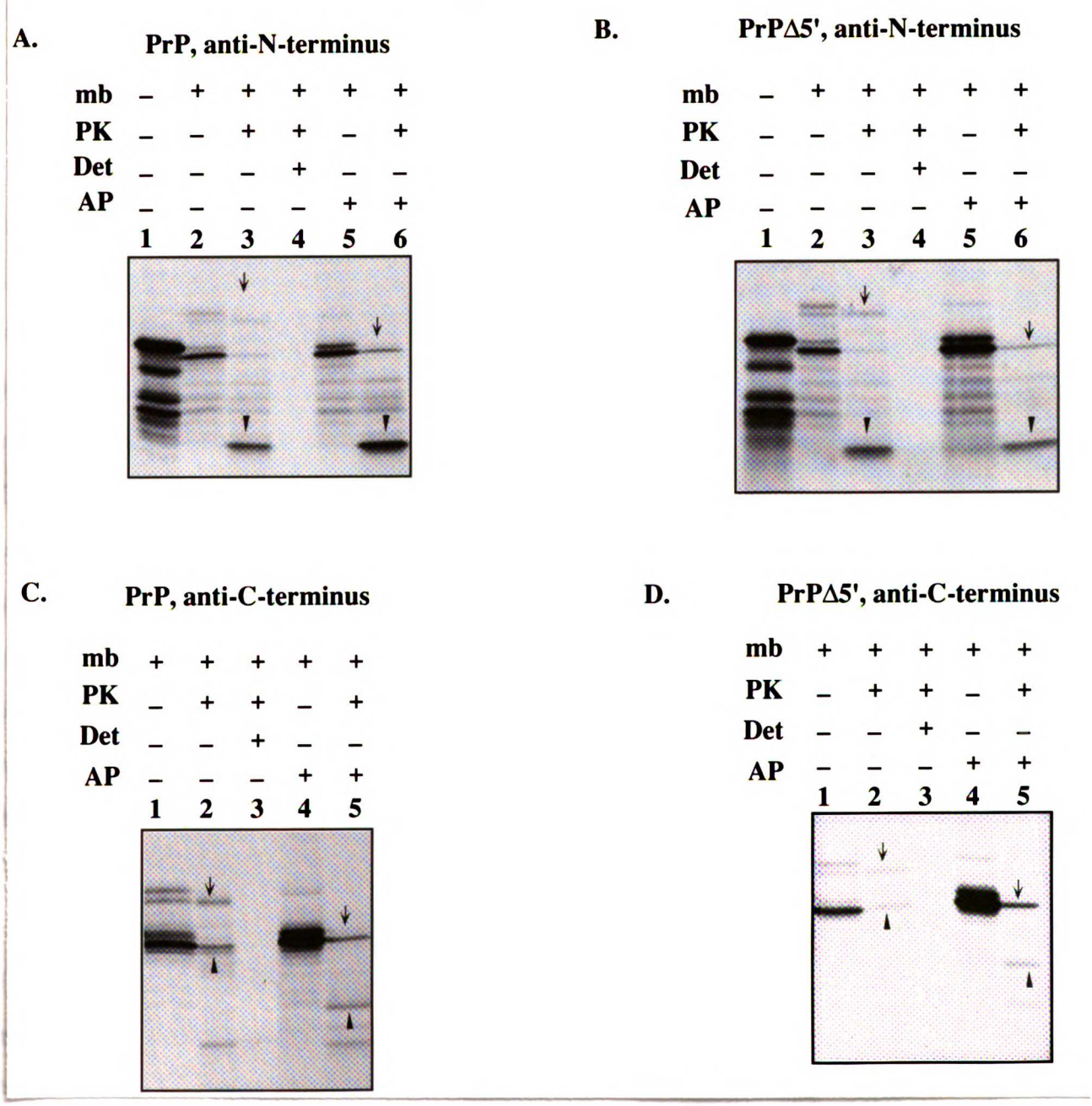
* This phenomenon is consistently observed when transmembrane PrP chains are present, and the upper band reproducibly corresponds to transmembrane chains while the lower band corresponds to secretory chains (see proteolysis data). Because the phenomenon is absent upon treatment with Endo H (data not shown) or translation in the presence of a tripeptide inhibitor of glycosylation (see Fig 12), we surmised that the migration differences might reflect different carbohydrate side chains.

Figure 11: Deletions in STE region



We used site directed mutagenesis to engineer a restriction site within STE. Deletions of either codons 97-104 (PrPΔ5') or 105-112 (PrPΔ3') were made and the resulting proteins were assayed for topology in a cell-free translation system. In vitro synthesized transcripts were translated in a WG cell-free system in the absence or presence of DPMs. Products were proteolyzed as described above. Percentages of transmembrane and secretory chains were calculated by densitometry of full length versus protected bands in the presence and absence of proteolysis and results are shown in the graph beneath the panel. Arrowheads indicate protected fragments generated upon proteolysis and arrows indicate protected full-length products.

Figure 12: C-terminal translocation of PrP and PrPΔ5'STE



In vitro produced transcripts of PrP (A and C) or PrPΔ5'STE (B anbd D) were translated in WG at 28^{0}C for 1 hour in the presence (C and D) and absence (A and B) of a tripeptide inhibitor of glycosylation. Samples were proteolyzed as described above and immunoprecipitated with either anti-P1 (against the PrP N-terminus) or anti-P3 (against the PrP C-terminus). Arrowheads indicate proteolytically protected fragments, while arrows indicate full length protected chains.

fragments observed are not due to intrinsic protease resistance or incomplete proteolysis. These results suggest that the 10 amino acid sequence consisting of codons 103-113 is necessary for STE function, while the 9 amino acid sequence consisting of codons 94-102 can be deleted without affecting topology. To ensure that PrPΔ5', indeed shows the same membrane orientation as the native molecule with two membrane spanning domains, we immunoprecipitated proteolyzed samples of PrP and PrPΔ5' translated in the presence or absence of a tripeptide inhibitor of glycosylation with N-terminal specific (anti-P1) and C-terminal specific (anti-P3) antisera (Fig.12). Both glycosylated and unglycosylated full-length chains reactive to both N- and C-terminal specific antisera are seen in the presence of membranes and all translocated products migrate as a single band when translated in the presence of acceptor peptide (Compare lanes 2 vs. 5 of Fig.12A and B and 1 vs. 4 of Fig. 12 C and D). Upon treatment of samples with proteinase K, N-terminal immunoreactive fragments are observed in the presence and absence of acceptor peptide (Fig. 12A and B lanes 3 and 6), and C-terminal immunoreactive fragments that shift down in the presence of acceptor peptide (Fig. 12 C and D lanes 2 and 5), consistent with the only glycosylation sites being located in the C-terminus of the molecule, are observed for both PrP and PrPΔ5'. When these PrP Δ5' (data not shown) and PrPΔ3'(Fig. 13A) are translated in RRL, a predominance of secretory chains is seen upon proteolysis, consistent with previous observations regarding the system dependent difference in PrP topology (75). Thus, it appears that deletion of codons 94-102 previously defining the N-terminal portion of STE has no effect on the generation of transmembrane chains and translocation of carboxyl-terminal domain observed for full-length PrP, whereas deletion of codons 103-113 drastically reduces the number of transmembrane chains observed; the few transmembrane chains that are seen (20-25%) with PrPΔ3' span the membrane only once. It is impossible to ascertain at this point whether removal of codons 103-113 prevents the molecule from passing through the carbonate extractable membrane spanning intermediate described previously and instead uses a more conventional means of translocating into the

ER lumen, or whether PrPΔ3' actually passes through the transmembrane intermediate with such greater efficiency that it is impossible to detect the intermediate stages. However, the fact these 10 amino acids followed by TM1 are sufficient to direct PrP to a topology indistinguishable from that of the wild type would support the former hypothesis.

Alteration of amino acids within STE-TM1 can result in accumulation of transmembrane intermediates

Site directed mutagenesis was used to create amino acid changes in STE-TM1. The resulting PrP mutations were transcribed *in vitro* , translated in RRL or WG, and topology assessed by proteolysis as described in Figure 11. The effect of each mutation on topology is shown in Table 1 as the percentage of transmembrane chains observed in either RRL or WG as compared to wild type PrP. Three mutations, an Asn to Ile at codon 108 (N(108)-I), alteration of Lys and His at codons 110 and 111 to Ile residues (KH(110)-II), and an Ala to Val change at codon 117 in TM1 (A(117)-V), result in accumulation of transmembrane intermediates, as assayed by topology in the predominantly secretory RRL system. When translated in WG, all three mutations show 85-95% transmembrane chains, similar to that observed for wild type PrP (data not shown). However, when translated in RRL, while the wild type molecule is approximately 80% secretory, N(108)-I and A(117)-V are only 50% secretory, corresponding to a 30% increase in transmembrane chains. The number of secretory chains observed with KH(110)-II is reduced to 25%, corresponding to a 55% increase in transmembrane chains (see graph below Fig. 13). Translation of PrP and mutations in RRL in the presence of microsomes reveals both glycosylated and unglycosylated signal cleaved products of approximately 27 and 32-33kD respectively; glycosylated products shift down to the size of unglycosylated product when translated in the presence of acceptor peptide (Fig. 13A-E lane 1 vs. 4). Upon treatment with .25mg/ml Proteinase K, 80% of full length, translocated PrP is fully protected (Fig. 13B lanes 2 and 5). When N(108)-I and A(117)-V are treated with PK, only 50% of the total translocated

Table 1: Mutations of STE-TM1 and their effect on Topology

		RRL	WG
Mutation	Sequence		%TM
w t	GTHNQWNKPS K P K T N M K H M A G A A A A G A V	<20	85
ΔN-I	GTHNQWNKPS K P K T **I** M K H M A G A A A A G A V	50	95
ΔK1,2T	GTHNQWNKPS **T** P **T** T N M K H M A G A A A A G A V	<20	75
ΔT-L	GTHNQWNKPS K P K **L** N M K H M A G A A A A G A V	<20	85
ΔTN-Nae	GTHNQWNKPS K P K **A G** M K H M A G A A A A G A V	<20	60
ΔN-G	GTHNQWNKPS K P K T **G** M K H M A G A A A A G A V	<20	80
ΔN-D	GTHNQWNKPS K P K T **D** M K H M A G A A A A G A V	<20	85
ΔA117-V	GTHNQWNKPS K P K T N M K H M A G A A **V** A G A V	55	95
ΔKH-II	GTHNQWNKPS K P K T N M **I I** M A G A A A A G A V	70	100
ΔSTE5'	G K P K T N M K H M A G A A A A G A V	<20	85
ΔSTE3'	GTHNQWNKPS A G A A A A G A V	<20	<25
ΔNae	GTHNQWNKPS K P K A A G A A A A G A V	<20	<25

Figure 13:

We used site directed mutagenesis to alter specific amino acids within STE-TM1. PrP mutations, (A) Δ3'STE, (B) PrP, (C) PrPN(108)-I, (D) A(117)-V, (E) KH(110)-II, were transcribed as described previously and translated in Rabbit Reticulocyte Lysate supplemented with DPMs for 1 hour at 32°C. Products were proteolyzed as described above and total products analyzed by SDS-PAGE followed by autoradiography. Percentages of transmembrane and secretory chains were calculated by densitometry as described above.

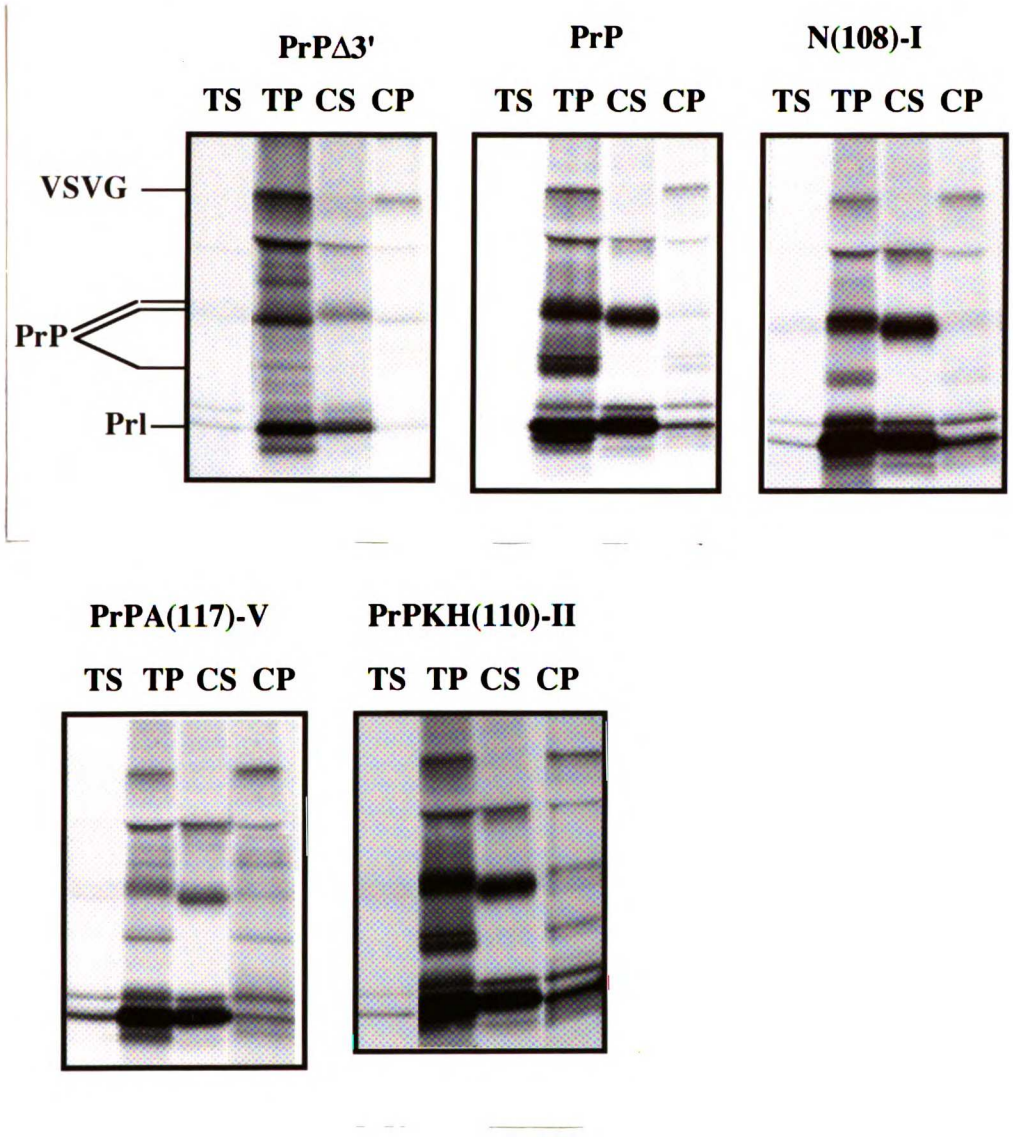
chains are protected and both N-terminal (Fig. 13C and D lanes 2 and 5) and C-terminal (data not shown) transmembrane bands are observed. When KH(110)-II is treated with PK, 70% of the total translocated chains are digested to generate N-terminal (Fig. 13E, lanes 2 and 5) and C-terminal (data not shown) membrane protected fragments. A small percentage of transmembrane chains is observed with the wild type PrP molecule and this may be due to very weak stop transfer activity of TM1 itself; a similar observation is made when the deletion PrP(Δ3'STE) is translated in RRL (Fig. 13A). Since PrP(Δ3'STE) does not show the predominantly transmembrane topology observed with PrP in WG, but a small percentage of transmembrane chains are observed in both RRL and WG systems, it is likely that the transmembrane chains observed with PrP in RRL are not due to STE activity. The upper glycosylated band observed for transmembrane PrP in WG can be seen when these mutations resulting in transmembrane chains are expressed in RRL (compare lane 1 of Fig. 13A, B, C, and D and see footnote p.42). Parallel proteolysis carried out in the presence of non-denaturing detergent results in complete digestion of translation products with all four proteins. Total products of these results were analyzed by densitometry, corrected for methionine distribution, and the percentages of secretory and transmembrane chains are shown in the graph below Fig. 13.

The mutations N(108)-I and KH(110)-II result in accumulation of transmembrane chains which are not integrated into the membrane

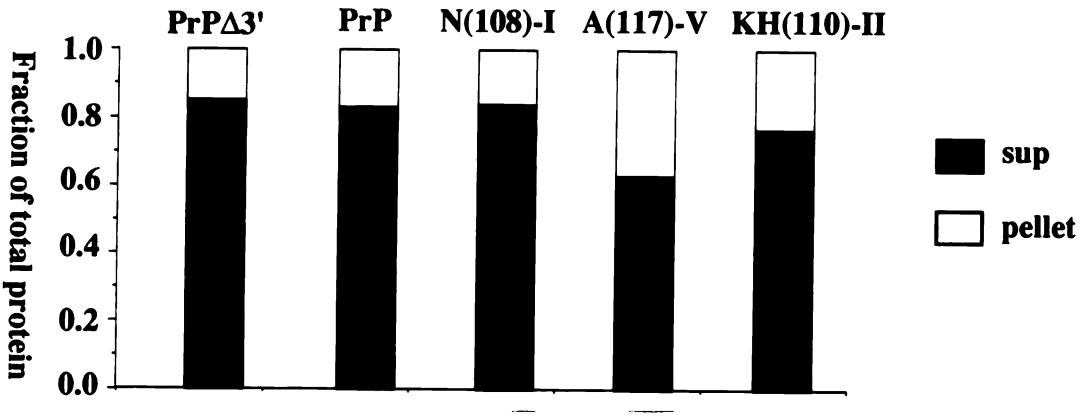
Previously we had demonstrated that the events directed by STE-TM1 result in a transmembrane intermediate that spans the membrane twice and that is, as evidenced by its extractability in .1M Sodium Carbonate, not integrated into the membrane. In light of this data, we wished to establish whether the mutations described in Fig. 13 resulted in accumulation of transmembrane intermediates that had integrated into the membrane or whether they remained in an aqueous channel, unable to proceed on to the ER lumen. Equal

aliquots of either PrP, N(108)-I, KH(110)-II, A(117)-V or PrP (Δ3'STE) were mixed with secretory and transmembrane controls, S-P and VSVG respectively, and diluted 125-fold in either Tris/sucrose at neutral pH (T) or .1M Na HCO3, pH11.5(C) and subjected to high speed centrifugation. While proteins spanning the membrane due to protein-protein interactions will be recovered in the membrane pellet under neutral Tris/sucrose conditions, the alkaline conditions of sodium carbonate will disrupt such interactions and only proteins that have integrated into the membrane will be recovered in the pellet, with peripherally associated membrane proteins being released into the supernatant. In the case of PrP, 80% of the total chains were recovered in the carbonate supernatant along with the prolactin control, with the 20% recovered in the pellet along with the transmembrane control, corresponding to the small number of transmembrane chains observed in RRL translation systems (Fig. 14 A). In the case N(108)-I, although 50% of the chains were in a membrane spanning orientation, only 20% of the total proteins chains were recovered in the carbonate pellet (Fig. 14C). Likewise, although KH(110)-II resulted 70% of the chains spanning the membrane, only 30% of the total chains were recovered in the pellet (Fig. 14E). In contrast, 55% of the total A(117)-V chains were spanning the membrane by proteolysis and 45% of the total chains were recovered in the carbonate pellet (Fig. 14D). Because some of the chains are found to be fully translocated by proteolysis and thus would not be expected to be recovered in the carbonate pellet, the percentage of transmembrane chains that are carbonate extractable or integrated into the membrane is shown in Fig. 14F. The 15-20% of total chains found spanning the membrane in the case of both PrP and PrP(Δ3'STE) are comprised entirely of integrated membrane. These chains also appear to span the membrane only once, suggesting that in the cell-free systems, a small percentage of the chains stop in the membrane at TM1, never complete the subsequent STE-TM1 directed events and become integrated when the channel closes. Mutations of Asn(108) or Lys(110) and His(111) to Ile residues resulted in accumulation of mostly carbonate extractable transmembrane chains, suggesting that these mutations somehow prevent efficient

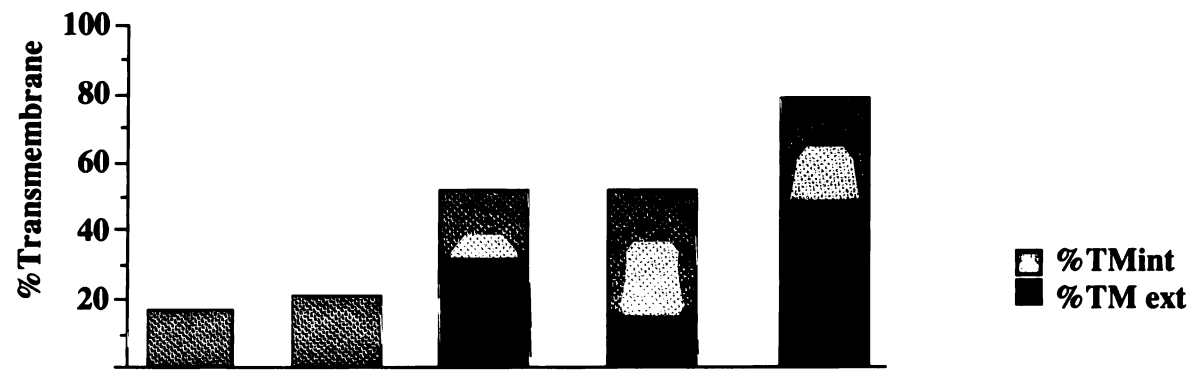
Figure 14: Membrane Integration of Mutants of STE-TM1



Translation products prepared as described in Figure 13 were mixed 5:1:2 with secretory and transmembrane controls, respectively, and diluted 120-fold in either Tris/sucrose, pH7.6 or Sodium Carbonate, pH11.5. Samples were centrifuged at 70,000rpm for 30 minutes after a 30 minute incubation on ice. Pellets were resuspended in 1%SDS, supernatants were precipitated with TCA, washed in acetone and resuspended in 1%SDS. Samples were dilted in 20 volumes of 1%Triton in PBS, spun in a microfuge for 10 minutes, supernatants transferred and immunoprecipated with antibodies to prolactin, PrP and VSVG before analysis on SDS-PAGE. Data was analyzed by densitometry and shown in the graph below as percentage of the total chains that are carbonate extractable. (A) PrPΔ3'STE, (B) PrP, (C) PrPN(108)-I, (D) PrPA(117)-V, (E) PrPKH(110)-II. (F) Data was normalized to the percentage of total transmembrane chains and shown in a graph as percentage of transmembrane chains that are carbonate resistant (gray bars) or integrated (black bars).



F. Integration of translocationally arrested transmembrane chains



translocation into the ER lumen either by prolonging interaction with translocation machinery or preventing interaction with necessary factors (76). If such were the case, one should be able to demonstrate interactions between these two mutations and translocation channel components that are not observed with the wild type PrP protein (119).

N(108)-I and KH(110)-II can be cross-linked to SSRα, a protein component of the translocation machinery

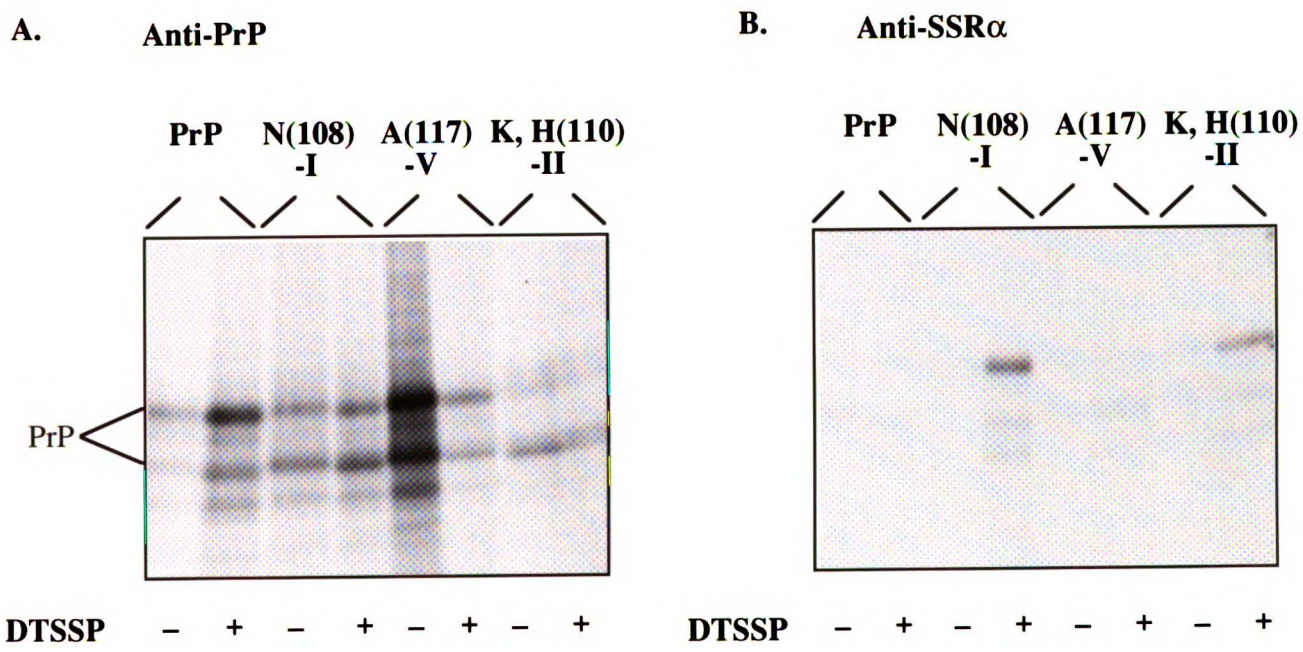
Equal aliquots of PrP, N(108)-I, KH(110)-II, A(117)-V or a secretory control (S-P) were spun over a .5M sucrose cushion in a Beckman airfuge at 27Psi for 10 minutes to pellet membranes and associated proteins. Pellets were resuspended in .1M sucrose and treated either with or without the reversible cross-linking agent DTSSP for 30 minutes at 24°C. After terminating reaction with Tris-Ac, the samples were boiled in SDS for 2 minutes to denature proteins, diluted in TritonX-100 and immunoprecipitated with antisera to either SSRα, PrP or non-immune serum (119). The crosslinking reaction was reversed by addition of 1M DTT and samples analyzed by SDS-PAGE. Antisera to PrP immunoprecipitates glycosylated and unglycosylated PrP species in the presence and absence of cross-linker in the case of PrP and the three mutations (Fig. 15A). Antisera to SSRα immunoprecipitates glycosylated and a smaller portion of unglycosylated PrP species of N(108)-I and KH(110)-II, but not of wild type PrP or A(117)-V (Compare Fig. 15B lanes 4 and 8 vs. 2 and 6). This immunoreactivity is seen only in the presence of cross-linker (Fig.15B lanes 1, 3, 5 and 7 vs. 2, 4, 6 and 8). These results strongly suggest that the membrane spanning but non-integrated chains that accumulate with the mutations N(108)-I and KH(110)-II are poised in the translocation channel, interacting with translocation machinery and unable either to efficiently proceed on into the ER lumen or to trigger channel disassembly. A(117)-V, on the other hand, which accumulates in the form of a predominantly integrated transmembrane form and is not found interacting with the

channel component SSRα, may have disengaged channel components, thus triggering disassembly of the translocation channel. That A(117)-V interacts with a different set of protein(s) involved specifically in channel disassembly is entirely possible and is not addressed in these studies.

Mutations of STE-TM1 result in decreased efficiency of translocation in XO

To determine whether the accumulation of transmembrane chains observed in RRL with mutations N(108)-I, KH(110)-II and A(117)-V was reflected in a similar alteration of translocation kinetics *in vivo* , PrP, N(108)-I, KH(110)-II, A(117)-V and PrPΔ3' transcripts were injected into XO in the presence of [35S]Met, incubated at 18^{0}C for 3 hours, after which XO were processed and subjected to proteolysis as described in Chapter 2. While at a 3 hour time point, PrP and PrPΔ3' both appear to be fully protected from protease, proteolysis reveals that N(108)-I is about 15% transmembrane, A(117)-V is about 20% transmembrane and KH(110)-II is about 30% transmembrane. These results suggest that although translocation into the ER lumen in the case of these three mutants is not prevented in an intact cell, it occurs with less efficiency than with the native molecule, thus providing physiological relevance for the findings described in the cell-free systems.

Figure 15: Mutations of STE-TM1 are found interacting with translocation channel proteins



Translation products prepared as described in Fig. 13 were spun over .5Msucrose at 27psi in a Beckman airfuge. Pellets were resuspended in an acetate free .1Msucrose buffer and treated either with or without .04MDTSSP at 24^{0}C for 30 minutes before terminating the reaction with 2M Tris-Acetate. Samples were boiled in 1%SDS, diluted 20-fold in 1%TX-100/PBS and immunoprecipitated with antibody to PrP (A), SSRα (B) or non-immune sera (data not shown). Due to aggregation, recovery was variable and so results were not quantitative.

III. Discussion:

Previously, it was demonstrated that PrP utilizes an unusual mechanism of translocation to achieve its final fully translocated topology, pausing first in the membrane at a hydrophobic region between codons 113-137 and subsequently translocating its C-terminus into the ER lumen to generate a transmembrane intermediate and proceeding on to a fully translocated form. The information directing these events was found to reside entirely in the region defined as STE-TM1, consisting of codons 90-143 and comprised of the first membrane spanning region and the hydrophilic sequence preceding it. We have identified a 10 amino acid deletion in the hydrophilic STE region that abolishes this activity as well as specific amino acid mutations that cause the accumulation of the transmembrane intermediate. The effects of these manipulations on PrP topology allow us to make some statements regarding the necessary amino acid composition of this topogenic sequence as well as elaborate on a previously proposed model for PrP translocation. In addition, the existence of these altered PrP molecules provide tools by which to investigate the physiological role of PrP topology in the pathogenesis of Scrapie and related diseases.

Amino acid Composition of STE-TM1

One of the aims of these studies was to further define the specific amino acid residues necessary for STE-TM1 action, perhaps identifying a general motif applicable to both STE and related pause-transfer sequences. The similarity between STE and the pause-transfer sequence of Apo B was noted previously. In other studies (94), STE alone was found capable of pausing translocation in the same manner as the pause-transfer element of Apo B (B'); however, these experiments only address the events occurring when the chain is still engaging the ribosome. Upon puromycin-induced release from the ribosome, no transmembrane chains are observed in the absence of TM1. Mechanistically, STE appears to be a variation on the pause-transfer theme, differing in that after directing an initial pausing event, STE and TM1 together direct subsequent C-terminal domain translocation.

This latter feature is suggested by the fact that neither STE placed in front of a heterologous membrane domain (ST) nor TM1 alone, in a context where it is allowed to span the membrane independent of STE, is capable of generating the twice membrane spanning intermediate described here (Chapter 2). It is quite possibly the juxtaposition of the pause-transfer sequence STE with a hydrophobic domain such as TM1 that allows PrP to remain transiently spanning the membrane, even after the ribosome has disengaged. The activities of these pause transfer sequences are not entirely exchangeable, however, as B' placed in front of TM1 does not direct C-terminal domain translocation. Despite notable sequence similarity between these two domains, we have not been able to identify specific mutations that alter the shared function, i.e. the initial pausing event, of both STE-TM1 and B' of Apo B (see Fig. 18). In fact, the mutations identified that affect PrP topology all share the common feature that they appear to prolong the membrane spanning step at the stage after ribosome release, while all of the mutations identified that affect B' pause/transfer action result in a loss of paused chains (118). Furthermore, numerous mutations of the residues that appear to occur in both sequences, have no effect on PrP topology. Rather, the observed similarity between STE and other pause transfer sequences may point to more subtle structural features that result in protein conformations lending themselves to efficient interactions with the required protein machinery.

In light of studies other groups (120, 121) suggesting a role for positively charged flanking regions in the action of signal anchor (SA) and stop transfer (ST) sequences, we thought that the 3 Lys and 1 His residues located in the essential portion of STE might be important in allowing TM1 to stop in the membrane. However, changing any one or two of these Lys residues to uncharged Thr residues had no effect on PrP topology; nor did changing an uncharged Asn residue to negatively charged Asp have an effect. Although these results do not rule out the possibility that sufficient charged residues exist upstream to compensate for these changes, this is an unlikely explanation as upstream sequences cannot compensate for removal of the region deleted in PrPΔ3'. What did appear to be a common

feature in generating mutations that accumulate in the transmembrane form, was the alteration of a charged or neutral residue to a hydrophobic residue with a ß-branched side chain. While alteration of Asn at codon 108 to Ile resulted in a 30% increase in transmembrane chains, all of which were extractable by sodium carbonate, changing this same Asn residue to either Asp, Gly or Ser had no effect on topology. Changing Lys (110) to Thr had no effect on topology, but changing that same Lys and the adjacent His residue to Ile residues resulted in a 55% increase in transmembrane chains, 90% of which were carbonate extractable. Comparison of Kyte-Doolittle hydrophobicity plots of both of these mutations to the wild type STE-TM1 region reveals that both N(108)-I and KH(110)-II extend the region of relative hydrophobicity, which may play a role in the increased tendency of these mutants to span the membrane. When Ala at codon 117 was changed to a Val residue, an alteration found in some patients contracting GSS, the resulting 30% increase in transmembrane chains was comprised predominantly of membrane integrated chains. When the Kyte-Doolittle hydrophobicity plot of this mutant STE-TM1 domain is compared to wild type STE-TM1 or to N(108)-I and KH(110)-II, an increase in the degree of hydrophobicity of TM1 is observed (Fig. 17). The relative inefficiency of translocation in cell-free system has been demonstrated to result in a small percentage of transmembrane chains that appear to become integrated by default when the channel closes (Fig. 14A and B, 14F). It is likely that this Ala to Val mutation may be so malfolded it is either unable to interact with necessary channel components, or does so with such inefficiency, that it becomes integrated in RRL by default when the channel closes. Although it is an α–helix that is most often found in membrane spanning regions (77), the structural disruption caused by the Ala to Val mutation might trigger the premature release of channel components and channel disassembly in the fractionated cell-free system (80, 122, 123). One possible interpretation of these results, with respect to the required amino acid composition of STE-TM1, is that TM1 has some of the features of a classic stop transfer sequence but lacks sufficient hydrophobicity to efficiently stop translocation by itself.

When preceded by a pause transfer sequence such as STE, TM1 spans the membrane, and due to its relative hydrophobicity does so transiently even after ribosome release, while a pause transfer sequence alone only pauses when the chain is nascent. In addition, features within STE-TM1 that have yet to be identified, allow the C-terminus of the molecule to be extruded into the lumen of the ER rather than remaining in the cytosol, before the second event in pause-transfer action, i.e. restarting, is achieved. One candidate feature of TM1 that may allow this C-terminal translocation event is the preponderance of Glycine residues, which are only marginally hydrophobic and are known to destabilize an α-helix (80). Studies on membrane proteins whose function mandates a conformational change, such as receptor mediated signalling, suggest that such proteins contain a high percentage of Val and Gly residues, probably to allow for greater flexibility. One would imagine, then, that substitution of Gly residues with Ala or Leu or with nonpolar Thr residues, might decrease efficiency of C-terminal translocation. Until recently, our research has focused on STE and further investigation of mutations within TM1 would be worth pursuing. It appears that while a pausing event directed by STE can be separated from TM1 (94), not all of the activities directed by STE-TM1 can be directed by one domain without the other.

Mutations of STE remain trapped in the translocation channel

The existence of mutations that render a translocationally paused protein trapped in the translocation channel allow us to elaborate on the distinction between STE-TM1 and a classic stop transfer sequence. A two step model has been posed previously to explain the events occurring when a stop transfer sequence interacts with translocation apparatus. The first event is the halting of translocation. The second event is the triggering of conformational changes that disrupt the ribosome membrane junction, disassemble the translocation channel and allow the chain to integrate into the membrane. We had previously demonstrated that STE-TM1 is similar to ST in that it directs the first of these

two events, but that the resulting transmembrane chain spans the membrane twice, does not integrate into the membrane and passes on into the ER lumen (Figs. 8, 9). Studies on transmembrane proteins with mutations in the hydrophobic core of their ST sequences (H^{ST}) have suggested that reducing the number of hydrophobic residues below 17 amino acids or inserting a charged residue into H^{ST} results in a transmembrane protein that is carbonate extractable, such as is directed by STE-TM1 (82). While the domain comprising TM1 is 22 amino acids long, it is not composed entirely of hydrophobic residues but contains a number of uncharged polar residues such as Glycine, Serine and Tyrosine, which may make it more like these mutant ST sequences. Presumably it is the interaction of STE with its receptor that allows the transmembrane intermediate to disengage the translocation machinery and proceed on into the lumen. This idea is supported by the evidence presented here that mutations preventing restarting of translocation result in prolonged interaction with components of the translocation channel. One question that remains unanswered is whether the ST sequence triggers channel disassembly, or merely disengages the translocation machinery rendering it unable to continue translocating, and becomes integrated into the membrane when the channel closes. In contrast, STE-TM1 might prolong interaction with translocation machinery to hold the channel open longer, thus allowing translocation to resume, or it may merely be incapable of triggering the channel disassembly that results in integration. That mutations resulting in accumulation of carbonate extractable transmembrane chains can be cross-linked to a component of the translocation machinery suggests that perhaps disengaging the protein machinery from the translocating chain allows the channel to close, and that prolonged interaction keeps the channel open. This is consistent with both the idea that a sufficiently hydrophobic sequence such as ST can speed the rate of channel disassembly by disengaging the translocation machinery and with the idea that the pause-transfer like sequence, STE-TM1, can delay channel closure by prolonging its interaction with translocation machinery.

Model of PrP Translocation and effect of point mutations

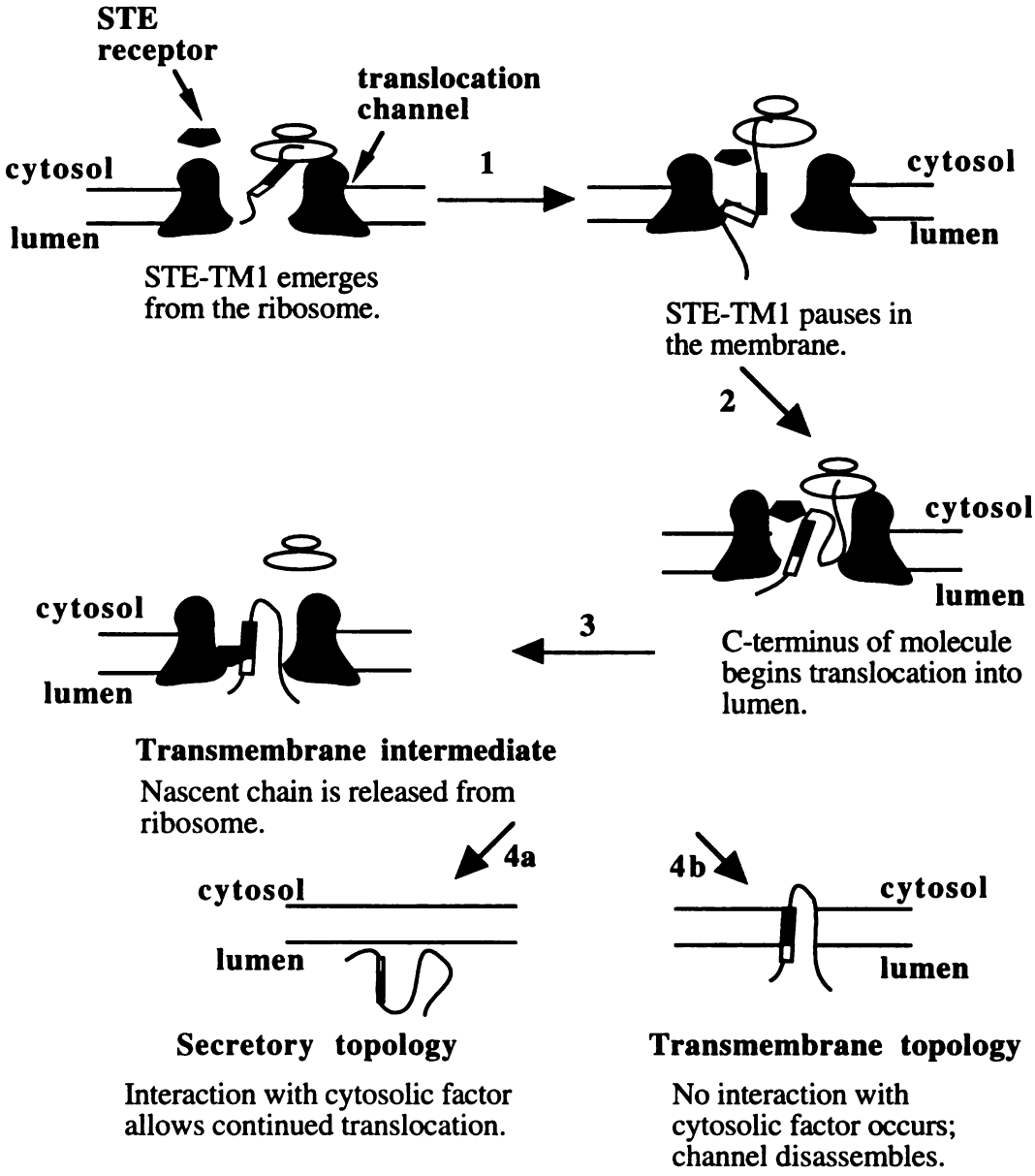
It had been shown previously by Lopez, et al., that achievement of a fully translocated PrP molecule required the presence of a cytosolic factor (75). Recognition of this factor appeared to require STE-TM1, but the factor is not necessary for the STE mediated pausing to occur (94). In light of our present data, we present here a model for PrP translocation and attempt to predict which steps are affected by the point mutations described (Fig. 16). In step 1, as STE-TM1 emerges from the ribosome, it engages the translocation machinery and the chain pauses in the membrane. In step 2, as chain elongation continues, a conformational change causes the C-terminus of the nascent chain to be extruded through the channel into the lumen of the ER as it emerges from the ribosome. In step 3, a termination codon is reached and the chain disengages the ribosome while STE-TM1 maintains its interaction with the translocation machinery. In step 4, STE binds to its receptor allowing release of channel components and in step 5, the chain proceeds on into the ER lumen. In the case of N(108)-I or KH(110)-II, the translocating chains accumulate at step 4, unable to disengage the translocation machinery, either before or after interaction with the STE receptor. In the case of A(117)-V, chains might disengage the translocation machinery at step 3, resulting in premature channel closure and integration into the membrane.

PrP mutations and topological regulation in the pathogenesis of prion related diseases

Although at this point, any connection between alteration of PrP biogenesis at the level of topology and the pathogenesis of Scrapie and related diseases is correlative at best, it is important to note that one of the mutations, an Ala to Val at codon 117, that in ours studies results in increased transmembrane chains, is also found in some patients contracting a human form of Scrapie, GSS (117). There is evidence to suggest that

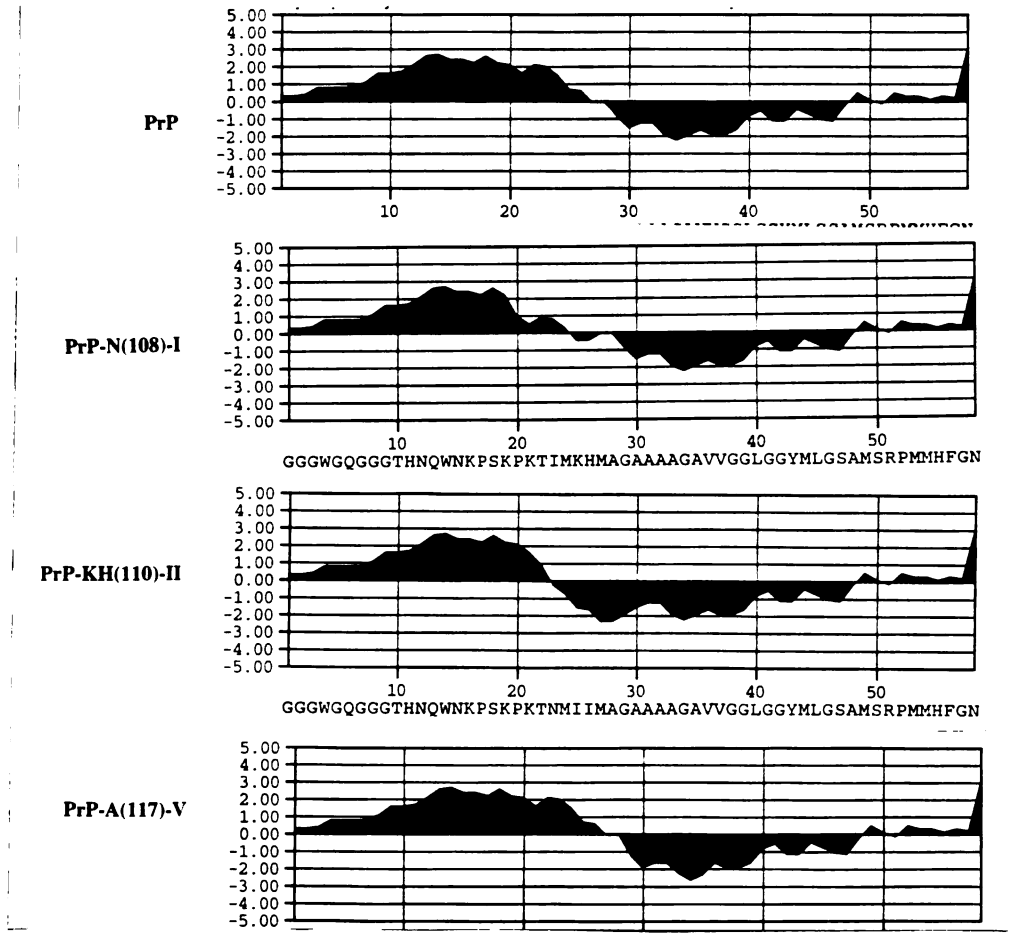
introducing a Val into a string of Ala residues disrupts the α-helix secondary structure, due to the introduction of a ß-branched side chain (80, 120). In the aqueous environment of the translocation channel, where the disruptive effects of ß-branched side chains is greater than in a hydrophobic environment, such a structural change may greatly impair chaperone mediated folding events (80). If one were to propose that the purpose of the STE-TM1 mediated mechanism of translocation is to promote certain folding events necessary to achieve a normal functional PrP molecule, alteration of events at this level of biosynthesis might have deleterious effects on the ultimate fate of the protein. It has been demonstrated in studies using synthetic peptides corresponding to various regions or PrP that the progression from the normal to diseased form of PrP is accompanied by a shift of the region corresponding to TM1 from an α-helical structure to a ß-pleated sheet (105) . In an intact system, this Ala to Val mutation may render PrP more susceptible to such an alteration, and this is reflected in the cell-free system in the abnormal kinetics of translocation and eventual integration into the membrane upon channel closure. In addition, the modification associated with the diseased isoform of PrP (PrPsc), has been shown to occur post-translationally and is most probably due to a conformational change as extensive structural analysis has revealed no physical modifications. By observing the effect of the accumulation of transmembrane chains seen with A(117)-V, N(108)-I and KH(110)-II or the absence of transmembrane chains seen with the 10 amino acid deletion in STE (PrPΔ3') on the susceptibility of transgenic mice expressing these mutations to Scrapie, it may be possible to establish the relevance of topologic regulation in the pathogenesis of this disease. In addition, mutations which are permanently paused at different stages in the translocation of PrP or which lack the ability to pause altogether, may provide useful tools to identify cytosolic, lumenal and membrane factors involved in the biogenesis of PrP under normal cellular conditions.

Figure 16: Model of PrP Translocation



Based on these and previous results with respect to PrP topology, we have proposed the following model of PrP translocation. (1) As the nascent PrP molecule is being translocated, the N-terminal signal sequence targets it to the ER membrane. STE emerges from the ribosome and the chain pauses in the translocation channel. (2) As translation continues, the C-terminus of the molecule is extruded into the ER lumen. (3) The termination codon is reached, the chain is released from the ribosome and the molecule achieves a bitopic transmembrane topology. (4a) In the presence of a cytosolic factor, such as is present in RRL, the molecule translocates into the ER lumen. (4b) In the absence of a cytosolic factor, such as is the case in WG, the channel disassembles an the chain remains in the membrane.

Figure 17: Hydrophobicity of PrP and Mutants



Kyte Doolittle hydropathy plots of codons 85-143, corresponding to STE-TM1, of hamster PrP(A) or mutations of STE-TM1, N(108)-I (B), A(117)-V (C) and KH(110)-II (D).

Figure 18:Comparison of Pause-Transfer Sequences

	11	10	9	8	7	6	5	4	3	2	1				
STE	K	P	K	T	N	M	K	H	M	A	G	A	A	A	A
B'	L	L	K	K	T	K	N	S	E	E	F	A	A	A	M

α-helix→

Sequences of a portion of STE-TM1(codons 104-118 of PrP) and B' (previously described from Apo B(26, 118). Sequences are aligned and numbered so as to compare the distance of a hydrophilic sequence (at positions 5-9) from a string of Alanines (solid line). Residues that showed decreased transmembrane or paused topology when mutated are indicated by rectangles, residues that showed increased transmembrane topology when mutated are indicated by circles.

CHAPTER 4: SUMMARY AND CONCLUSIONS

The data presented in Chapters 2 and 3 serves to reconcile some of the previously disparate findings with regards to PrP topology, between studies conducted in cell-free systems and observations made in neuroblastoma cells and brain (75, 98, 104, 106, 108). Previously, both transmembrane and fully translocated forms of PrP had been observed in cell-free systems, but only a fully translocated form was observed in the protein's native context or in intact cells. Furthermore, the predominantly transmembrane wheat germ system could be converted to a predominantly secretory system by the addition of reticulocyte lysate, but not the converse, an observation that had suggested that complete translocation of PrP was dependent on a cytosolic factor (75). By demonstrating a precursor/product relationship between the transmembrane and secretory forms, the results in Chapter 2 are consistent with both cell-free and *in vivo* findings.

The data in Chapter 2 suggests that codons 85 to 143 of hamster PrP, termed STE-TM1, comprise a unique topogenic element that directs three translocational events: first, the pausing of the nascent chain in the membrane, second, the translocation of the C-terminus of the molecule into the lumen as it emerges from the ribosome and third, the subsequent resuming of translocation into the lumen after release of the nascent chain from the ribosome. In light of similarity in both sequence structure and function between previously identified pause/transfer elements and PrP-STE, these findings point to the existence of a family of functionally diverse proteins utilizing variations on a common mechanism of translocation. Although no consensus sequence was identified by the mutagenesis studies described in Chapter 3, definition of a common pause/transfer sequence may depend on more subtle structural features rather than direct amino acid sequence homology (Fig. 18). It has been demonstrated, for example, that conversion of Lys and Thr in the Apo B pause-transfer sequence (two residues also appearing in PrP-STE) to Leu residues abolishes pausing, but similar mutations in STE do not eliminate transmembrane topology in WG (118, Chapter 5, Appendix i). From the data presented in

Chapter 3, i.e. that elimination of charges in or insertion of negative charges into the positively charged STE region of PrP, do not affect topology, one might postulate that the critical characteristic of both STE and B' is a specific secondary structure and not overall charge or a specific consensus sequence. Introduction of two hydrophobic residues with a high propensity to form alpha helices into the hydrophilic B' may alter the secondary structure such that recognition of translocation machinery involved in pausing is prevented. Introduction of equally hydrophobic residues with ß-branched side chains, such as Ile or Val, into the same position may have a different effect on pausing. Alteration of Asn to Ile in STE resulted in accumulation of transmembrane intermediates, whereas this same mutation in B' did not result in accumulation of paused chains. One explanation for these results would address the overall structure of the two domains. Both B' and STE have predominantly unbranched hydrophilic residues (Lys-Lys-Thr-Lys-Asn in the case of B' and Lys-Thr-Asn-Met-Lys in the case of STE) placed 5 residues upstream of a string of Alanine residues (see Fig. 18). Introducing a hydrophobic ß-branched side chain at a given position in this sequence may be a critical factor in generating mutations that accumulate a paused or transmembrane intermediate (80), as is observed with PrP-N(108)-I in Chapter 3. Thus, the Asn residue might occur in both sequences by coincidence, the important feature being its polarity and side chain structure. Similarly, alteration of the Thr residue in STE to Leu residues had no discernible effect on topology, but alteration of the Lys and Thr residues in B' resulted in ablation of paused chains. Leu and Ile are both extremely hydrophobic, but Ile has a helix disrupting potential where Leu does not. In fact, if one were to align the two sequences such that the Ala series in both sequences are matched (see Fig. 18), the corresponding residues in STE to the Lys-Thr in B' would be Thr-Asn, alteration of which to Ala-Gly, did reduce pausing by 40-50% (see Table 1). Clearly, more extensive mutagenesis of both regions would be necessary to establish a consensus motif for pause-transfer, but the studies described in Chapters 2 and 3, serve to characterize STE-

TM1 and lay the groundwork for investigation of the relationship between various pause-transfer elements.

An important outcome of the results described in Chapters 2 and 3 is the implication that alteration of protein conformation during the translocation of a secretory protein across the ER membrane may be another mechanism by which protein expression is subject to regulation by external factors. A role for chaperones in protein folding, and therefore proper function, has been observed for many proteins (72-74), and the existence of a protein lined channel through which proteins translocate into the ER is relatively well established (3, 6, 64). The concept of a class of proteins containing topogenic information that directs specific translocational events, subject to regulation by cellular factors, is consistent with these previous findings. Genetic mutations of such sequences may be involved in the progression of diseases associated with proteins of this class. In the case of PrP, mutations of the STE-TM1 sequence have already been found in some patients contracting GSS, and at least one of these mutations appears to result in alteration of an STE-TM1 directed translocational mechanism. Many of the studies on post-translational modifications of proteins, such as occurs with PrPsc, have concentrated on post-ER trafficking events and covalent modifications. Given the data presented in this work and studies on other pause-transfer proteins, it would appear that folding events occurring as a nascent chain traverses the ER membrane provide another possible step where protein expression can effectively be regulated.

The cellular function of PrP remains unknown but that it plays a major role in the pathogenesis of Scrapie, GSS and related neurodegenerative diseases has been clearly demonstrated in studies where transgenic mice expressing a mutant PrP gene, (Pro(102)-L) found in some patients with GSS (116), develop Scrapie *de novo* , while a PrP null mouse is immune to infection with Scrapie (124). While there is most likely more than one pathway by which PrPc can be converted to PrPsc, these and previous studies characterizing the biogenesis of PrP point to translocation of PrP into the ER as a possible

step for this conversion. If such were the case, transgenic mice expressing only a mutant PrP gene (PrPΔ3') that does not pass through a transmembrane intermediate might be immune to infection with Scrapie, and mice expressing mutant PrP molecules that accumulate transmembrane intermediates (PrP-N(108)-I, PrP-KH(110)-I, and PrP-A(117)-V) might show increased susceptibility to Scrapie infection or even development of the disease *de novo*. In addition, inactivation of factors involved in PrP translocation, such as the factor present in RRL, might have an affect on the development of Scrapie. Future studies on transgenic mice expressing the translocation mutants described in this work as well as identification of the factors involved could greatly enhance our understanding of the overall role of the STE-TM1 sequence and the activities it directs in the development of prion related diseases.

The studies described in Chapters 2 and 3 as well as the previous studies conducted on PrP biogenesis open the door for a host of new experimental directions. With respect to the role of STE-TM1 mediated events in the pathogenesis of Scrapie, there are the transgenic studies described above. In addition, structural studies using synthetic peptides containing mutations that affect PrP topology may help to elucidate the relationship between secondary structure and STE-TM1 action (105). Further mutagenesis of STE-TM1 and other pause-transfer sequences such as B' of Apo B may generate a structural motif for a common topogenic sequence; the studies presented here make a case for hydrophobicity and side chain structure as relevant features in the amino acid sequence of STE-TM1. Finally, the mutant PrP molecules described in Chapter 3 could provide useful tools for identifying factors involved in PrP translocation and even in the actions of other topogenic elements such as Stop Transfer sequences, which may use some of the same machinery.

One major question that arises from the above discussion is why the cell would choose such a costly, and potentially dangerous, mechanism of biogenesis for a protein such as PrP. It would seem that having a built in system, the regulation of which might lead to conversion of a normal cellular protein to a pathogenic isoform, would not be one

lead to conversion of a normal cellular protein to a pathogenic isoform, would not be one which is evolutionarily conserved, and yet STE-TM1 is a highly conserved region of the PrP sequence (96). There are several possible explanations to this conundrum. One explanation is that amyloidogenesis, under certain cellular circumstances and when tightly controlled, serves some necessary function. Studies on various regions of PrP have identified a segment (codons 109-122), which overlaps with STE-TM1, as the most amyloidigenic region of the protein. In addition, this region demonstrates a high propensity to assume a ß-pleated sheet structure under conditions that might be associated with Scrapie, such as the substitution of a Valine for an Alanine at residue 117. The ability to undergo a conformational change from an α-helical structure to a ß-pleated sheet, in a manner that is mediated during translocation into the ER by interaction with specific protein factors might contribute to the formation of amyloidogenic PrP. Perhaps, under normal cellular conditions, only a small portion of the total PrP synthesized is in this conformation, whereas, under diseased conditions, the majority of PrP is in the amyloidigenic conformation. The potential role of amyloid formation in normal brain is not yet understood, but the correlation between large amounts of amyloid deposits and neurodegeneration is well documented (1, 117). A second possible explanation for this seemingly complex method of PrP biogenesis is that the tradeoff in overall cellular costs, manifested in the generation of a properly folded, functional PrP molecule weighed against the rare instances of translocationally aberrant molecules, makes this mechanism physiologically expedient. The energy cost of unfolding or disposing of misfolded PrP molecules the cell might incur utilizing a mechanism of translocation that did not involve a transient transmembrane intermediate might be greater than the risk of generating PrPsc during the STE-TM1 mediated events described in these studies. A third possibility is that post-translational regulation of PrP expression, prior to its exit from the ER, permits the cell to respond rapidly to changes in the environment, perhaps targeting PrP to a different compartment under certain cellular conditions. PrPsc has, in fact, been found to accumulate

of a highly regulated folding event occurring in the protein rich environment of the translocation channel may be to ensure proper folding of PrP into a non-amyloidigenic conformation, while targeting misfolded molecules to a degradative pathway before they exit the ER. Mutations which render the protein more likely adopt the amyloidigenic conformation, or conditions which affect other factors involved in PrP translocation, may reduce the efficiency of these folding events and thus result in the accumulation of PrPsc. It is difficult at this point to determine the exact role of the STE-TM1 mediated mechanism of translocation in either the normal cellular function if PrP or in the pathogenesis of Scrapie, but these studies provide a springboard from which to investigate these questions.

CHAPTER 5: APPENDICES

i) Effects of Asn to Ile mutation on STE-like element of Apo B

Based on the similarity between the sequence of STE of PrP and B' of Apo B, it seemed plausible that a mutation of a shared residue (Asn, located at codon 108 in PrP, to Ile) that affected STE function, might have a similar effect on B' function. The chimera S.L.B'.G, used to establish the pause-transfer activity of B', was subjected to site-directed mutagenesis, altering the Asn residue of B' to Ile (S.L.B'(N-I).G). The chimeras S.L.B'.G, S.L.B'(N-I).G, S.L.ST.G and S.L.G were truncated at either BamHI (Fig. 19A) or BstEII sites (data not shown), located in the G domain prior to the stop codon. These truncated cDNAs were transcribed and translated in WG for 1 hour, at which point reactions were terminated with the addition of emetine to a final concentration of 2mM. Equal aliquots were treated with or without puromycin (final concentration 10mM) followed by EDTA (final concentration 2mM), proteolyzed as described previously and analyzed by SDS-PAGE (Fig.19A). Mutation of the Asn residue in B' did not affect pausing or the ability of B' to resume translocation after release from the ribosome.

To investigate the possibility that the effects of this mutation were masked by the absence of a hydrophobic domain long enough to span the membrane, such as is the case with STE-TM1, the mutant B' was engineered into a chimera in which B' was followed by TM1 from PrP (S.L.B'-TM1.G) to give S.L.B'(N-I)-TM1.G. The chimera S.L.B'-TM1.G has been shown previously to result in transmembrane chains in WG, in a manner similar to the parallel chimera containing STE-TM1 (S.L.STE-TM1.G). All three chimeras were translated in WG and proteolyzed as described above (FIg. 19B). No difference in topology was observed between them. Failure to observe any significant effect of this mutation may reflect a dependence of pause-transfer sequences on overall structural features rather than an exact sequence (see Chapter 4).

ii) Effects of temperature on the topology of PrP in cell-free systems

It has been demonstrated that the kinetics of membrane trafficking are affected by temperature. Incubation of cells at 4^0C will prevent vesicular budding from the ER membrane resulting of accumulation of proteins at this point in the secretory pathway. It has also been suggested that translocation rates are increased at higher temperatures. In addition, temperature has been shown to affect the folding of proteins into alpha-helical or beta-pleated secondary structures.

The topology of PrP in cell-free systems appear to be affected by the temperature of the translation reaction, at temperatures where the overall level of protein synthesis is not. When PrP is translated in WG at either 22^0C, 25^0C, and 30^0C, an increase secretory chains is seen at 22^0C (approximately 50%), whereas at 24^0C and 30^0C, 80-90% of the total chains are transmembrane (Fig. 20). A similar phenomenon is observed in RRL when temperature are varied between 24^0C and 35^0C, with transmembrane chains increasing from 20% to 60%. This temperature dependent effect is not observed with either a secretory control (Prl), a transmembrane control (S.L.ST.G) or the STE deletion shown previously to eliminate the STE-TM1 directed transmembrane topology, suggesting the existence of a temperature sensitive component of PrP translocation. To determine whether this temperature effect involved a membrane factor or a cytosolic factor, translations were carried out in WG at 22^0C with membranes that had been incubated for 10, 60 and 90 minutes at 30^0C. No temperature dependent effect on topology was seen between these membranes, suggesting increased temperature was not inactivating a membrane factor. Likewise, when translations were carried out at 22^0C with aliquots of WG that had preincubated at 30^0C for 10, 60 and 90 minutes, no effect on topology was observed. It is possible that the effect temperature is to increase the rate at which the translocation channel disassembles, thus trapping the transient transmembrane intermediates directed by STE-TM1 or by affecting the secondary structure of STE-TM1, impairing its ability to direct the restarting of translocation. In fact, the mutations described in Chapter 3 that result in

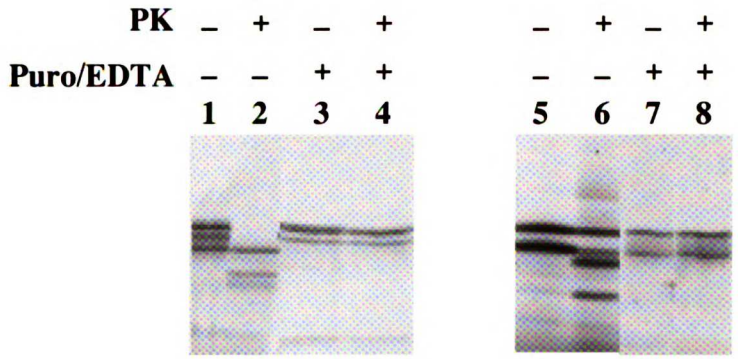
accumulation of transmembrane chains show an increased sensitivity to this temperature effect. This temperature sensitivity of PrP topology may serve as a useful tool to identify factors involved in STE-TM1 directed events, providing another means by which the system can be manipulated experimentally.

iii) Generation of transgenic mice expressing mutant PrP genes

A logical next step from the results presented in Chapters 2 and 3 is determining the phenotype of topological mutants of PrP with regards to the pathogenesis of prion related diseases. For this purpose, a number of DNA constructs, for expression in the previously established PrP null mouse have been made. Because the PrP null mouse is viable, but immune to infection with Scrapie (124), it is possible to observe the effect of expression of the mutations described in Chapter 3 on susceptibility to Scrapie without background expression of the endogenous gene. In addition, the possibility that a normal allele is necessary as well as the mutant alleles identified in GSS patients can be addressed by parallel expression of the mutant PrP molecules in mice expressing endogenous PrP. The strategy is diagrammed in Figure 21.

Figure 19: Effect of N-I mutation on the pause-transfer activity of B'

A. Pause/Transfer Activity of S.L.B'.G vs S.L.B'(N-).G



B. Topology of S.L.B'-TM1.G and S.L.B'(N-I)-TM1.G in WG

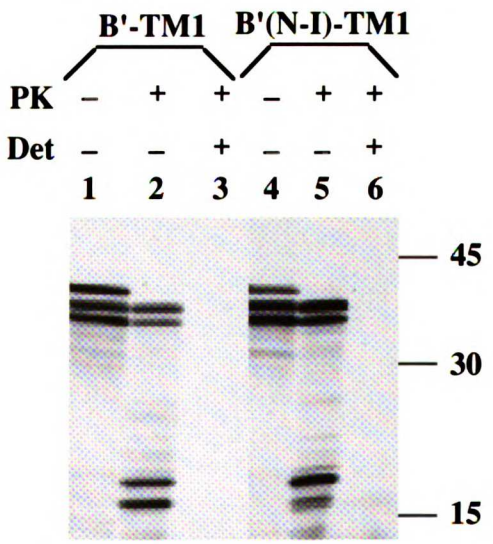
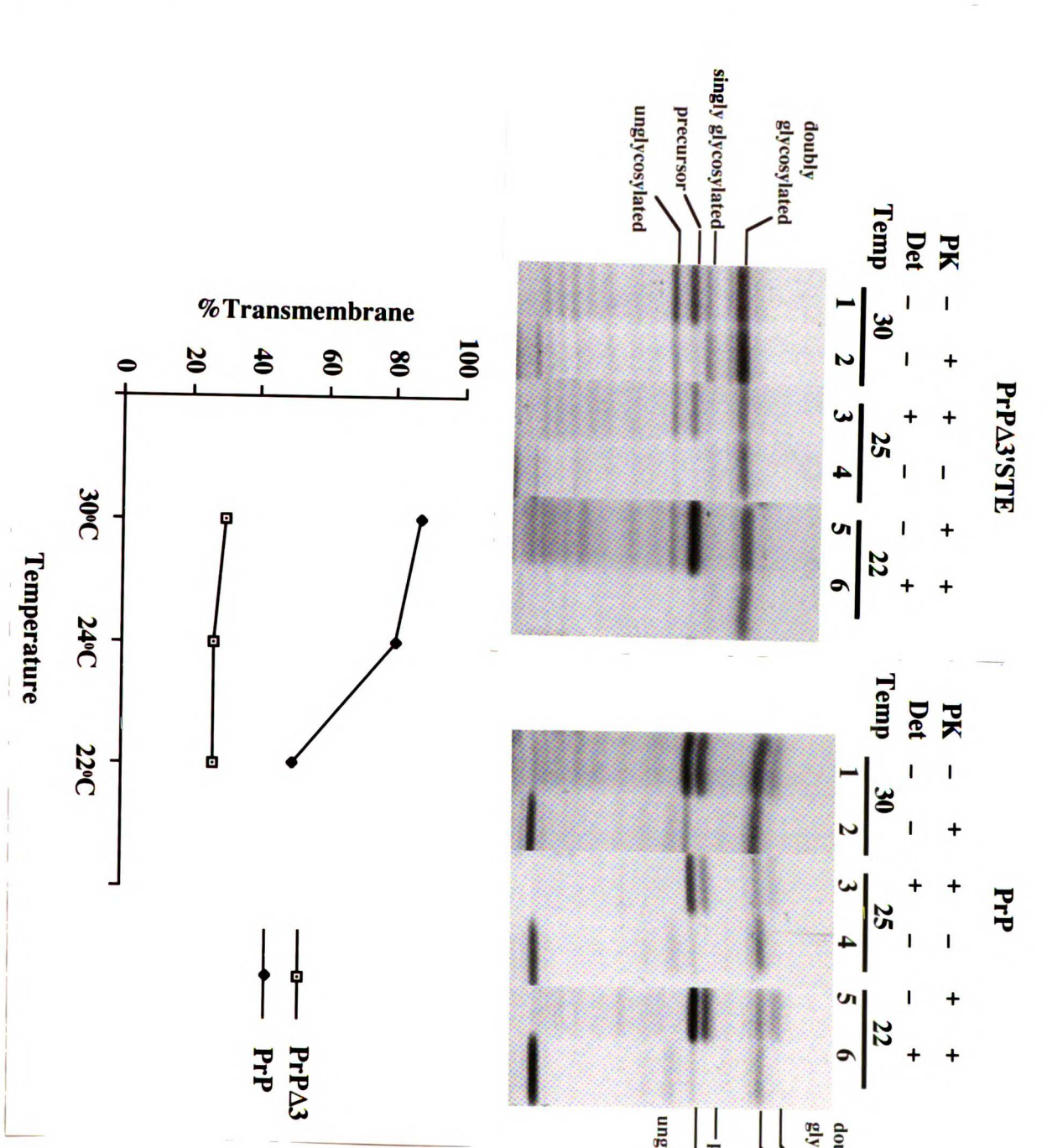


Figure 20: Temperature Dependence of Transmembrane PrP Topology



iv. Materials and Methods

Materials

Restriction endonucleases were obtained from New England BioLabs (Beverly, MA). Placental RNAse inhibitor was obtained from Promega Biotech (Madison, WI). Rabbit anti-sheep prolactin was obtained from United States Biochemical Corp (Cleveland, OH). Rabbit anti-human globin antibody was prepared by immunizing rabbits to the peptide C-K-T-N-V-K-A-A-W-G-K-V-G-A-H-N-G-S (including residues 8-21 of human globin) coupled to keyhole limpet hemocyanin. Sera obtained after serial boosts with immunogen were screened for reactivity to native and denatured chimpanzee globin transcription-linked translation products. Preparation and characterization of rabbit anti-hamster PrP antibodies, P-1 and P-3, have been described (125). Rabbit anti-hamster PrP antibody, P-5, was raised against a peptide comprising codons 140-178 of hamster PrP. Affigel Protein A-Agarose was obtained from BIORAD (Richmond, CA). Proteinase K was from E. Merck (Darmstadt, FRG), [^{35}S] methionine translabel from ICN Biochemicals (Indianapolis, IN). ENHANCE fluorography reagent was obtained from Dupont (Wilmington, DE). Plasmid GC1, used for site directed mutagenesis, was obtained courtesy of C. Spencer Yost (U.C.S.F.) and is described in reference 126. Restriction enzymes, T4 Ligase and T4 Polymerase, SP6 polymerase, RNAsin and Creatine Kinase were purchased from New England Biolabs (Beverly, MA), with the exception of Csp 45I which was purchased from Promega (). M13 helper phage for production of single stranded DNA was obtained from INVITROGEN (San Diego, CA). CJ236 (ung-dut-) were provided courtesy of John Forsayeth (U.C.S.F.). Purified oligonucleotides for site-directed mutagenesis were purchased from the U.C.S.F. Biomedical Resource Center. Crosslinking agents DTSSP and DSS were purchased from PIERCE (). Anti-SSRα antibody was a generous gift from Peter Walter (U.C.S.F.).

Construction of plasmids

All plasmids were derived from pSP64, into which the 5' untranslated region of Xenopus globin is inserted at the HindIII site to generate pSP64T (127).

Plasmid PrP-G: This plasmid encodes hamster PrP from the initial methionine of the S sequence through amino acid residue 152, followed by 143 residues of chimpanzee alpha globin engineered behind the SP6 promoter in pSP64T. The globin domain contains an 8 amino acid coding sequence engineered into the BssHII site that encodes an N-linked glycosylation acceptor site. Globin DNA was digested with NcoI, 5' overhangs were filled in with the Klenow fragment of DNA Polymerase I and digested with EcoRI, a recognition site for which was located downstream in the polylinker. PrP was digested with Cla I, previously created at nucleotide 462 by site-directed mutagenesis (126), 5' overhangs were filled in with Klenow Polymerase, and linearized DNA was digested with EcoRI. A 462 base pair fragment including globin and the PrP-encoding vector DNA were purified using low melting point agarose and ELUTIP columns. Fragment and vector were then ligated at 16°C for 12 hours in the presence of T4 DNA Ligase, ATP and nucleotides.

Plasmid S.G.X.P: Chimera contained a signal sequence from bovine prolactin followed by two passenger domains, denoted G and P, flanking a putative membrane domain, where G consists of amino acids 1-117 of chimpanzee alpha globin with a glycosylation site as described (20) and P consists of amino acids 57-199 of bovine prolactin. Putative membrane domains were denoted X, composed of either codons 85-143 from hamster PrP (STE-TM1), a stop transfer sequence (ST) from membrane IgM heavy chain (17, 20) or a signal sequence (S) from bovine prolactin. Nucleotides 255-440 of hamster PrP, with 3' EcoRI linkers added at a Klenow blunted Ava II site, and 5' Nco generated sticky ends were engineered into previously described S_L.L.ST.G (consisting of

a signal sequence from lactamase (S_L), amino acids 85-140 of b-lactamase, IgM stop transfer and the globin domain described above, ref. 20) that had been digested with Nco I and EcoR I. The resulting construct, S_L.L.STE-TM1.G, was digested at Nco I and Eco RI and ligated into previously described ST.P (20). The resulting intermediate STE-TM1.P was digested with Nco, treated with Klenow fragment of DNA Polymerase I, digested with Sal I and the purified fragment ligated into S_L.G.ST.P (20), which had been digested with BstE II, treated with mung bean nuclease and digested with Sal I. The resulting construct, S_L.G.STE-TM1.P was partially digested with Nco I in the presence of Ethidium Bromide, blunted with Klenow Polymerase and then digested with Nhe I located upstream of the SP6 promoter in the SP64T vector, to remove the region encoding S_L. which was replaced with signal of prolactin (S) by ligating an Nhe to Hinc II fragment from prolactin encoding the signal sequence into the above vector to generate S.G.STE-TM1-P. S. G.ST.P was generated from S_L.G.ST.P by replacing S_L with S as described for S.G.STE-TM1.P.

Plasmids S.L.ST.G.ST.P and S.L.ST.G.S.P: A BstE II to Sal I fragment from previously described S.L.ST.G was replaced with BstE II to Sal I fragments from G.ST.P or G.S.P (20), encoding ST.P and S.P respectively.

Plasmid S.L.ST.G.STE-TM1.P: Previously described S.G.STE-TM1.P and S.L.ST.G were digested with Bam HI and NheI; the fragment encoding S.L.ST was eluted and ligated into the S.G.STE-TM1.P vector, replacing the existing signal.

Site directed mutagenesis:

The entire coding region of hamster PrP had been placed into the plasmid GC1, which contains an M13 origin of replication , a PBr322 backbone, including the Ampicillin resistance gene and the origin of replication, a GC rich stretch of approximately 200 nucleotides and a polylinker (see ref.129). The PrP gene, with a KpnI site at nucleotide

280, an XbaI site at nucleotide 400 and a Cla I site at nucleotide 480, was inserted into the GC1 polylinker at Bgl 2 and Eco R1 sites, to give the plasmid CRP38. Single stranded DNA was then prepared by transforming CRP38 in CJ236 cells, an ung-dut- bacterial strain, followed by infection with M13 helper phage and preparation of single stranded DNA as described by Maniatas, et al. (130). Because of the orientation of the M13 origin opposite to that of the PrP gene, the single stranded DNA consisted of antisense DNA. Phosphorylated oligonucleotide GTGGAACAAGCCTTCGAAGCCAAAAACC which creates a Csp45I site at nucleotide 106 was annealed and DNA was filled in with T4 Polymerase and T4 Ligase. The resulting double stranded DNA was transformed into ung+dut+ DH1 cells and colonies screened for the presence of a Csp45I site. The resulting plasmid, CRP35 was digested with Bgl 2 and EcoR1, and the fragment, consisting of the PrP gene with new restriction sites, was purified from low melting point agarose using an ELUTIP column and ligated into the same sites in an SP64 plasmid, containing the SP6 promoter, to generate PrP35.

Construction of STE deletions:

PrP35 was either digested with KpnI and Csp45I and blunted with Klenow Polymerase before religating (PrPΔ5'), or digested with Csp45I, blunted with Mung Bean Nuclease and digested with NaeI before religating (PrP Δ3').

Construction of STE-TM1 mutations:

Plasmid GC1 (129) was digested with XbaI and Eco RI, located in the polylinker, and the vector purified as described above. Hamster PrP, previously engineered behind the SP6 promoter in the plasmid SP64, was digested with NheI, located in the SP64 backbone, and EcoRI, located in the polylinker downstream of the PrP gene, and the purified fragment was ligated into the GC1 vector. The resulting plasmid pGC-SP-PrP contains the M13 origin of replication for production of single stranded DNA as well as the SP6 promoter.

The following oligonucleotides were used for site directed mutagenesis as described above: AAAACCATCATGAAG (introducing a BspHI site as well as altering Asn to Ile at codon 107), AAAACCAACATGATCATCATGGCCGGCGCT (introducing a Bcl I site as well as altering Lys and His at codons 110 and 111 to Ile), ATGGCCGGCGCTGCAGTGGCAGGGGCCGTG (introducing a PstI site as well as altering Ala to Val at codon 117), AGTAAGCCAAAGCTTAAC (introducing a HindIII site as well as altering Thr at codon 107 to Leu), CCCAGTACTCCAACAACC (introducing a Sca Isite as well as altering Lys at codons 104 and 106 to Thr), AGTAAGCCAAAAGCCGGCATGAAGCACATG (introducing an NaeI site as well as altering Thr and Asn to Ala and Gly at codons 107 and 108), CCAAAAACCGGTATG (introducing an AGEI site as well as altering Asn to Gly at codon 108), CCGAAGACCGACATGAAG (introducing a BbsI site as well as altering Asn to Asp at codon 108), AAGCCAAAAACATCGATGAAGCACATG (introducing a ClaI site as well as altering Asn to Ser at codon 108). Purified, phosphorylated oligonucleotide AAAACCAAGATCTCTGAGTTT, which creates a PstI site and alters an Asn to Ile in Apo B B' element was annealed to single stranded DNA prepared from S.L.B'.G engineered into the same vector used for PrP mutagenesis.

Cell-free transcription-linked translation and translocation:

Plasmid DNA was linearized by restriction endonuclease digestion at sites located downstream of the coding region: PrP expression plasmids were linearized with EcoRI, and plasmids S.G.STE-TM1.P, S.G.ST.P and S.G.S.P, were linearized with Sal I. Secretory control proteins, full-length and truncated prolactin, were also linearized with SalI and a transmembrane control protein, VSVG, was linearized with EcoR I. Transcription by SP6 RNA Polymerase was carried out at 40°C for 1 hour as described (127), with modifications. Cell-free lysates of wheat germ or rabbit reticulocytes were prepared as described by Erickson and Blobel (87) and Perara and Lingappa (67),

respectively, with modifications. Cell-free systems were used to translate the synthesized mRNAs for 1 hour at 28°C, as previously described (75), with modifications. Microsomal membranes derived from dog pancreas were prepared as described (88), with modifications and used at a concentration of 5 absorbance units at 280 nm per 10µl in the translation reaction. Calcium chloride was added to a final concentration of 10mM to all translation samples after the 1 hour incubation and samples were either untreated or digested with 0.15 mg/ml proteinase K in the presence and absence of 0.5% (v/v) Triton X-100 for 1 hour at 0°C. All samples were then treated with 0.5µl of 0.2M phenylmethylsulfonylfluoride in dimethyl sulfoxide, transferred into either boiling 0.1M Tris Cl/1%SDS pH 8.9 or boiling 3% sarkosyl, and boiled for 5 minutes.

Immunoprecipitation of cell free products:

After completion of digestion with proteinase K as above, cell-free translation products were boiled in either 0.1M Tris Cl/1%SDS and diluted 20-fold with TX-SWB(1% (v/v) TritonX-100 in PBS, pH 7.5) or 3% sarkosyl and diluted 10-fold with PBS (), after which 1-2 µl of antiserum was added. After 8 hours at 4°C, 25 µl of protein A-agarose beads were added and tubes were rotated at 4°C for 2 hours. The beads were washed one time in TX-SWB and one time in PBS; proteins were then eluted by boiling in protein gel loading buffer containing 4% SDS and 1.0 M DTT. For samples which were subjected to treatment with Endoglycosidase H, proteins were eluted from Affigel beads after the final wash step by boiling in 50µl of 0.1M Na citrate, pH 5.5 for 1 minute. Samples were then divided into two equal sized aliquots and 0.25 ul of Endo H (1U/ml) added at 37°C for 6 hours before standard sample preparation (647). All immunoprecipitated samples were then analyzed by SDS-PAGE as previously described. Autoradiographs were visualized after treatment of gels with the fluorography reagent, ENHANCE.

Carbonate Extractions:

Samples of PrP, N(108)-I, KH(110)-II or A(117)-V RRL translation product were mixed with transmembrane and secretory controls (Vesicular Stomatitus Virus Glycoprotein, VSVG, or codons 87-199 of bovine prolactin following a signal sequence, S-P, respectively) at ratio of 5:3:1 and diluted 125 fold in either .25M sucrose/.1MTris-Ac, pH7.6 or .1M NaHCO3, pH11.5. All samples were incubated for 30 minutes at 0°C before centrifugation at 70,000rpm for 30 minutes in a Beckman TL100 tabletop ultracentrifuge using a _ rotor. Supernatants were removed and protein recovered by precipitation with a final concentration of 20% Tricholoroacetic acid (TCA) followed by 3 washes in acetone. TCA precipitates were resuspended in 1%SDS/.1M Tris-Ac, pH7.6 and incubated for 12 hours at 40°C to allow for better recovery. Pellets were resuspended as

described for TCA precipitates. Specific translation products were identified by immunoprecipitation with specific antisera as described below.

Cross-linking:

30µl of PrP, N(108)-I, KH(110)-II, A(117)-V or the secretory control S-P translated in RRL in the presence of microsomes were spun over a 150µl sucrose cushion consisting of .5M sucrose/.1MKCl/2mMEDTA in a Beckman airfuge at 27psi for 10minutes to remove cytosolic proteins. Supernatants were aspirated and pellets resuspended in 100µl of .1Msucrose/.1MKCl/-MMgCl2. 50µl of each sample were treated with or without the reversible cross-linking reagent DTSSP at a final concentration of .4mM for 30 minutes at 24°C. Reactions were terminated by the addition of 1µl 2M Tris-Ac, pH7.6. Samples were then diluted 5 fold in 1%SDS and boiled for 2 minutes. Each sample was then split 3 ways and immunoprecipitated with either antisera to SSRα, PrP or Non-immune sera as described above. Prolactin control samples were immunoprecipitated with anti-prolactin antisera instead of anti-PrP.

Microinjection and processing of Xenopus Oocytes:

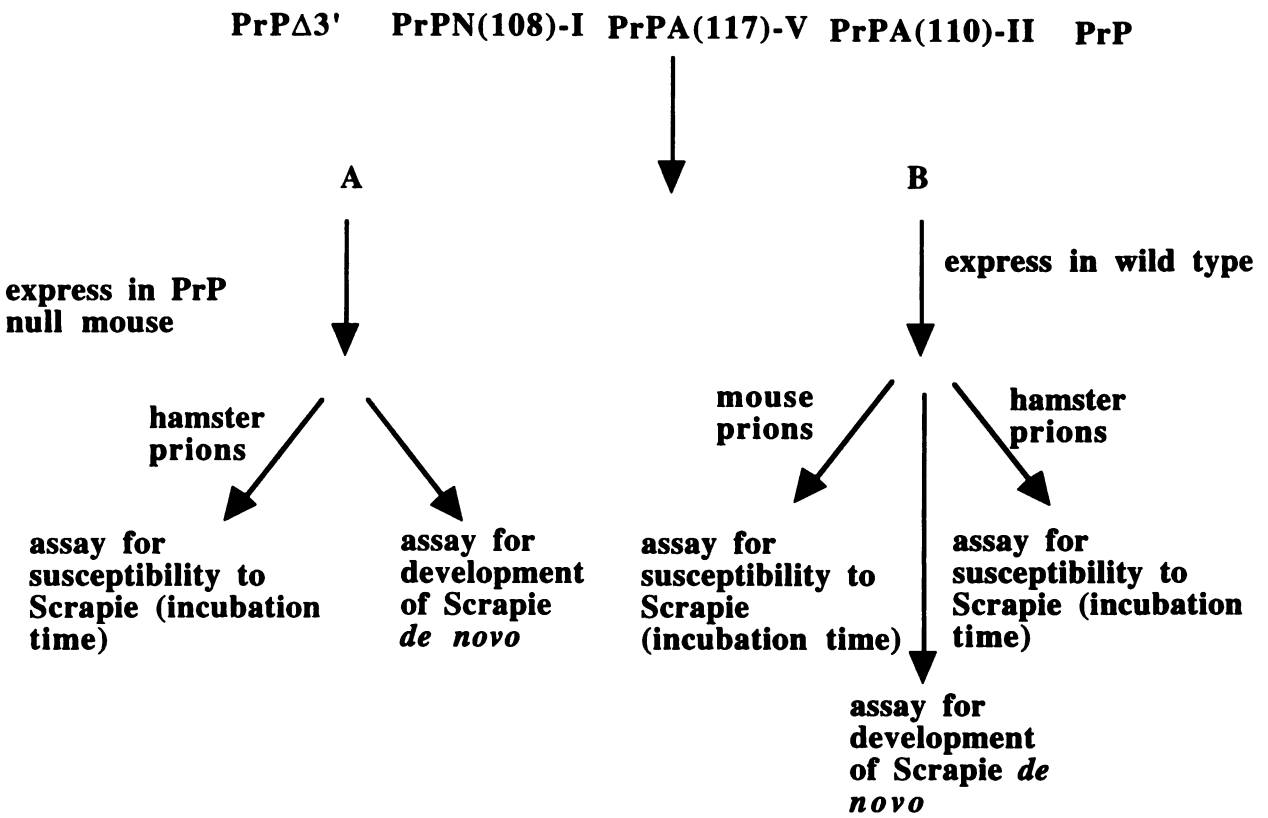
Stage 5 and 6 XO, stored at 18°C in MBSH for 1-4 days post surgery were injected with 50nl of total transcription products and 0.1 µCurie of a [^{35}S]methionine (Translabel). For pulse-chase studies on chimeras, oocytes were incubated for 1 hour post injection and equal numbers of oocytes were either homogenized immediately (pulse) or injected with 1000-fold excess unlabeled methionine and incubated at 18°C for 2 hours followed by homogenization (chase, see ref. 113). XO were homogenized in 90 mM KAc/40 mM Tris-Acetate, pH 7.6/10mM $CaCl_2$ using a microfuge tube-fitted teflon homogenizer. Three volumes of homogenization buffer per oocyte were used. XO injected with transcription products encoding VSVG, a known transmembrane protein (110) were included in each homogenized sample as a control for membrane integrity during proteinase K digestion. Samples were divided in three and either untreated, digested with .5mg/ml proteinase K, or

Proteinase K in the presence of 0.5% (v/v)Triton X-100, and incubated at 0°C for 1 hour. Protease digestion was terminated as described for cell-free products. Samples were then spun for 20 minutes in a microfuge, and supernatants transferred to fresh tubes. The cleared supernatants were then diluted 20-fold in TXSWB and immunoprecipitated as described for cell-free products.

Carbonate Extraction of XO products:

Homogenate of XO expressing S.G.X.P was mixed 2: 1: 1 with homogenate of XO expressing transmembrane and secretory controls (VSVG and bovine prolactin, respectively). Pooled samples were divided in half and diluted 330-fold with either Tris-HCl-buffered sucrose (pH 7.5) or 0.1M sodium carbonate (pH 11.8) (92), after which pH was checked and if necessary titrated to the indicated pH with 10N NaOH. After 30 minute incubation at 0°C, samples were centrifuged in a Beckman TL100 ultracentrifuge with a 100.2 rotor at 70,000 rpm in a TL100 centrifuge for 30 minutes at 2°C. The supernatants were removed carefully and those of carbonate samples were neutralized with glacial acetic acid. Proteins from all supernatants were precipitated with 20% TCA on ice for 15 minutes and spun in a microfuge for 10 minutes. Precipitates were washed twice with ice cold acetone and spun in a microfuge for 10 minutes each. TCA precipitates and carbonate pellets were resuspended in 0.1M Tris-HCl, pH 8.9/1%SDS and incubated at 37°C overnight before 20-fold dilution in TX-SWB. Samples were immunoprecipitated and SDS-PAGE analyzed as described for cell-free products.

Scheme for Expression of PrP mutants in Transgenic Mice



A1. PrP expressing mouse should develop Scrapie only when infected with hamster prions

A2. Will mouse expressing PrPΔ3' be immune to infection?

A3. Will mouse expressing A(117)-V, KH(119)-I or N(108)-I have a shorter incubation time, i.e., more susceptible to infection?

A4. Will mouse expressing A(117)-V, KH(119)-I or N(108)-I develop Scrapie *de novo* ?

B1. Mouse expressing PrP should develc Scrapie when infected with either mouse hamster prions.

B2. Will mouse expressing A(117)-V, N(108)-I or KH(110)-II show increased susceptibility to either mouse or hamster prions?

B3.Will mouse expressing A(117)-V, N(108)-I or KH(110)-II develop Scrapie *de novo* ?

REFERENCES

1. Walter, P and Lingappa, V.R. (1986) *Ann. Rev. Cell Biol.* **2**, 499

2. Perara, E. and Lingappa, V.R.(1988) *Protein Transfer and Organelle Biogenesis,* **Chapter 1**, Das, R.C. and Robbins, P.W., Eds., Academic Press, New York ·

3. Simon, S.M. and Blobel, G. (1991).*Cell* **65**, 371-380.

4. Gorlich, D., Hartmann, E., Prehn, E., Rapoport, T., (1992) *Nature,* **357**, 47-50

5a. Hartmann, E., Wiedmann, M., and Rappoport, T.A. (1989), *EMBO J.*, **82**, 2225-30.

5b. Miglaccio, G., Nicchita, C.V. and Blobel, G. (1992), *J. Cell Biol.*, **117**, 15-21.

6. Skach, W.R. and Lingappa, V.R. (1993), *Mechanisms of Intracellular Trafficking and Processing of Proproteins*, **Chapter 2,** Y. Peng Loh, CRC PRess, Boca Raton, Florida.

7. Nicchitta, C.V., Migliaccio, G., Blobel, G.(1991), *Cell*, **65,** 587-

8. Walter, P., *Nature*, **367**, 47

9. Evans, E.A, Gimore, R. and Blobel, G. (1986), *Proc. Natl. Acad. Scie U.S.A.*, **83**, 581

10. Cabelli, R. J., Chen, L., Tai, B.C. and Oliver, D.B. (1988), *Cell*, **55**, 683-

11. De Shaies, R.J., Sanders, S.L., Feldheim, D.A. and Scheckman, R. (1991), *Nature*, **349**, 806

12. Greenberg, G., Shellness, G.S., and Blobel, G. (1989), *J. Biol. Chem.*, **264**, 15762-

13. Blobel, G. (1980), *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 1496-

14. Northwehr, S.F. and Godon, J.I. (1990) *Bioessays*, **12**, 479

15. Lingappa, V.R., Katz, F.N., Lodish, H.F. and Blobel, G. (1978), *J. Biol. Chem.*, **253**, 8667-

16. Rothman, J.E. and Lodish, H.F. (1977), *Nature*, **269**, 775-

17. Yost, C.S., Hedgpeth, J., and Lingappa, V.R. (1983). *Cell* **34**, 759-766

18. Friedlander, M.and Blobel, G. (1985). *Nature* **318**, 338-340.

19. Spiess, M. and Lodish, H.D.(1986), *Cell*, **39**, 267

20. Rothman, R.E., Andrews, D.W., Calayag, M.C. and Lingappa,V.R. (1988). *J. Biol. Chem.* **263**, 10470 -10480.

21. von Heijne, G. and Gavel, Y. (1988), *Eur. J. Biochem.*, **174**, 671

22. Yamamoto, Y., Taniyama, Y., Kikuchi, M. and Ikehara, M. (1987), *Biochem. Biophys. Res. Comm.*, **149**, 431-

23. Davis, N.G., Boeke, J.D. and Model, P. J. (1985), *J. Mol. Biol.*, **181**, 111

24. Chuck, S.L., Yao, Z., Blackhart, B., McCarthy B.J. and Lingappa, V.R. (1990). *Nature* **346**, 382-385.

25. Chuck, S.L. and Lingappa,V. R. (1992) *Cell* **68**, 9-21

26. Yost, C.S., Lopez, C.D., Prusiner, S.B., Myers, R.M. and Lingappa, V.R. (1990). *Nature* **343**, 669-71

27. Walter, P. and Blobel, G. (1980), *Proc. Natl. Acad. Sci.U.S.A.*, **77**, 7112

28. Walter, Pl and Blobel, G. (1981), *J. Cell Biol.*, **91**, 557

29. Wolin, S.L. and Walter, P. (1989), *J. Cell Biol.*, **109**, 2617

30. Walter, P. and Blobel, G. (1982), *Nature*, **297**, 647-

31. Lingelbach, K., Zwieb, C., Webb, J.R., Marshallsay, C., Hoben, P.J., Walter, P. and Dobberstein, B. (1988), *Nucleic Acids Res.*, **16**, 9431-

32. Hann, B.C. and Walter, P.(1991), *Cell*, **67**, 131

33. Siegal, V. and Walter, P. (1988), *Cell*, **52**, 39

34. Wiedman, M., Kurzchalia, T.V., Bielka, H. and Rapoport, T.A. (1987), *J. Cell Bio.*, **113**, 229

35. Herz, J. Flint, N. Stanley, K. Frank, R. and Dobberstein, B.(1990), *FEBS Lett*, **276**, 103

36. High, S. and Dobberstein, B. (1989), *J. Cell Biol..*, **113,** 229

37. Rothman, J.E. (1989), *Nature*, **340**, 433

38. Römisch, K., Webb, J., Lingelbach, K., Gausepohl, H.,and Dobberstein, B. (1990), *J. Cell Biol..*, **111,** 1793

39. Gilmore, R., Walter, P. and Blobel, G. (1982), *J. Cell Biol.*, **85**, 470

40. Gilmore, R., Walter, P. and Blobel, G. (1982), *J. Cell Biol.*, **95**, 463

41. Gilmore, R., and Blobel, G. (1983), *Cell*, **35**, 677

42. Meyer, D.I. and Dobberstein, B.(1980), *J. Cell Biol.*, **87**, 498

43. Meyer, D.I. Krause, and Dobberstein, B. (1982), *Nature*, **297**, 647

44. Römisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M. and Dobberstein, B. (1989), *Nature*, **340**, 478

45. Lütcke, H., High, S., Römisch, K., Ashford, A. and Dobberstein, B. (1992), *EMBO J.*, **11**, 1543

46. Miller, J., Wilhelm, H., Gierasch, L., Gilmore, R. and Walter, P.(1993), *Nature*, **366**, 351

47. Connelly, T., and Gilmore, R. (1989), *Cell*, **57**, 599

48. Connolly, T., Rapiejko, P.J. and Gilmore, R. (1991), *Science*, **252**, 171

49. Lauffer, L., Garcia, P.D., Harkins, R.N., Loussens, L., Ulrich, M.A. and Walter, P. (1986), *J. Cell Biol.*, **103**, 1167

50. Thrift, R.N., Andrews, D.W., Walter, P. and Johnson, A.E.(1991). *J. Cell Biol.* **112**, 809-814.

51. Prehn, A., Herz, J., Hartmann, E., Kurzchalia, T.V., Frank, R., Römisch, K. Dobberstein, B. and Rapoport, T.A. (1990), *Eur. J. BIochem.*, **188**, 439

52. Wada, I., Rindress, D. Cameron, P.H., Ou, W.-J., Doherty, J.J, Louvard, D., Bell, A.W., Dignard, D., Thomas, D.Y. and Bergeron, J.J.M. (1991), *J. Biol. Chem.*, **266**, 19599

53. Margolese, L., Waneck, G.L., Suzuki, C.K., Degen, E., Flavell, R.A. and Williams, D.B. (1993) *J. Biol. Chem.*, **268**, 17959

54. David, V., Hochstenbach, F., Rajagopalan, S. and Brenner, M.B. (1993), *J. Biol. Chem.*, **268**, 9585

55. Ou, W., Cameron, P.H., Thomas, D.Y. and Bergeron, J.J.M.(1993), *Nature* , **364**, 771

56. Munro, S., and Pelham, H.R.B. (1986), *Cell* **46**, 291

57a. Kohno, K., Normington, K., Sambrook, J., Gething, M.J. and Mori, K. (1993), *Mol. and Cell. Biol.*, **13**, 877

57b. Brodsky, T.L., Hamamoto, S., Feldheim, D. and Scheckman, R. (1993), *J. Cell Biol.*, **120**, 93

58. Bonafacino, J.S., et al. (1990), *Science* **241**, 79

59. Musch, A., Wiedmann, M. and Rapoport, T.A. (1992), *Cell*, **69**, 343

60. Sanders, S.L., Whitfield, K.M., Vogel, J.P., Rose, M.D. and Schekman, R.W. (1992),*Cell*, **69**, 353

61. Görlich, D., Prehn, S., Hartmann, E., Kalies, K.-V. and Rapoport, T.A. (1992), *Cell*, **71**, 489

62. Bullied, N.J. and Freedman, R.B. (1988), *Nature*, **335**, 649

63. Hasel, K.W., Glass, J.R., Godbut, M. and Sutcliffe, J.G. (1991), *Mol. and Cell. Biol.*, **11**, 3484

64. Lingappa, V.R. (1991), *Cell*, **65**, 527

65. Hirschberg, C.B.and Snider, M.D. (1987), *Ann. Rev. Biochem.*, **56**, 63

66. Alberts, B.A., Bray, ,D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. (1983), *Molecular Bioogy of the Cell*, 3rd Edition, Garland Press, New York, NY

67. Perara, E. and Lingappa, V. R.(1985). *J. Cell Biology* **101**, 2292-2301.

68. Ferguson, M.A. J. and Williams, A.D.(1988), *Annu Rev. Biochem.*, **57**, 285

69. Caras, I. and Weddell, G. N. (1991). *J. Cell Biology* **113**, 77-85

70. Beckmann, R.P., Mizzen, L.A. and Welch, W.J. (1990), *Science*, **248**, 850

71. Hartl, F.U., Hlodan, R. and Langer, T. (1994), *TIBS*

72. Gething, M.J. and Sambrook, J. (1992), *Nature*, **355**, 33

73. De Shaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A. and Scheckman, R.(1988), *Nature*, **332**, 800

74. Chirico, W.J., Waters, M.G., and Blobel, G. (1988), *Nature*, **332**, 805

75. Hay, B., Prusiner, S.B. and V.R. Lingappa (1987). *Biochemistry* **26**, 8110-8115.

76. Lopez, C.D., Yost, C.S., Prusiner, S.B., Myers, R.M. and Lingappa, V.R. (1990). *Science* **248**, 226-229.

77. Branden, C., and Tooze, J.(1991), *Introduction to Protein Structure*, Garland Publishing, Inc. New York, NY

78. Davis, N.G., Boeke, J.D. and Model, P. (1985), *Cell*, **41**, 607

79. Kyte, J. and Doolittle, R.F. (1982), *J. Mol. Biol..*, **157**, 105

80. Li, S.-C. and Deber, C.M. (1992), *FEBS Lett.*, **311**, 217

81. von Heijne, G. (1985), *J. Mol. Biol.*, **184**, 99

82. Goldstein, J., Lehnhardt, S. and Inouye, M. (1990), *J. Bacteriol.*, **172**, 1225

83. Briggs, M.S., Cornell, D.G., Dluthy, R.A. and Gierasch, L.M. (1986), *Science*, **233**, 206

84. von Heijne, G. and Gavel, Y. (1988), *Eur. J. Biochem.*, **174**, 671

85. Kuroiwa, T., Sakaguvchi, M., Mihara, K and Omura, T., *J. Biol. Chem.*, **266**, 9251

86. Fasman, G.D. and Gilbert, W.A. (1990), *TIBS*, **15**, 89

87. Erickson, A. H. and Blobel, G. (1983). *Meth. Enzymol.* **96**, 38-50. New York, Academic Press, Inc.

88. Walter, P. and Blobel, G. (1983). *Meth. Enz.* **96**, 84-96. New York, Academic Press, Inc.

89. Merrick, W.C. (1983), *Meth. Enz.*, **101**, 606

90. Garoff, H. (1985), *Annu. Rev. Cell Biol.* **1**, 403

91. Lingappa, J.R., Martin, R., Welch, W.J., Ganem, D. and Lingappa, V.R., *J. Cell Biol.*, in press

92. Fujiki, Y., Hubbard, A.L., Fowler, S. and Lazarow, P.B. (1982). *J. Cell Biol.* **93**, 97-102.

93. Perara, E., Rothman, R.E. and Lingappa, V.R., (1986), *Science,* **232**, 348

94. Nakahara, D.H., Lingappa, V.R. and Chuck, S.L. (1994), *J. Biol. Chem.,* in press

95. Tigyi, G. and Parker, I. (1991), *J. Biochem. and Biophys. Methods*, **22**, 243

96. Prusiner, S.B. (1991). *Science* **252**, 1515-1522.

97. Oesch, B., Westaway, D., Walchli, M., McKinley, M.P., Kent, S.B.H., Aebersold, R., Barry, R.A., Tempst, P., Teplow, D.B., Hood, L.E., Prusiner, S.B. and Weissmann C. (1985). *Cell* **40**, 735-46.

98. Stahl, N., Borchelt, D.R., Hsiao, K. and Prusiner, S. B. (1987). *Cell* **51**, 229-40.

99. Bueler, H., Fischer, M., Lang, Y., Bluethmann, H., Lipp, H.P., DeArmond, S.J., Prusiner, S.B., Aguet, M. and Weissmann, C. (1992). *Nature* **356**, 111-118.

100. Meyers, R.K., Mc Kinley, M.P., Bowman, K.A., Braunfeld, M.B., Barry, R.A. and Prusiner, S. B. (1986). *Proc. Natl. Acad. Sci. USA* **85**, 2310-14.

101. Stahl, N., McKinley, M.P. and Prusiner, S.B. (1990). *Biochemistry* **29**, 5405-5412.

102. Basler, JK., Oesch, B., Scott, M., Westaway, D., Walchli, M., Groth, D.F., Mc Kinley, M.P., Prusiner, S. B. and Weissmann, C. (1986). *Cell* **46**, 417-28.

103. Stahl, N., Baldwin, M.A., Teplow, D.B., Hood, L., Gibson, B.W., Burlingame A.L. and Prusiner, S.B.(1993). *Biochemistry* **32**, 1991-2002.

104. Borchelt, D. R., Taraboulos, A. and Prusiner, S.B. (1992). *J. Biol. Chem.* **267**, 16188-16199.

105. Gasset, M., Baldwin, M.A., Lloyd, D.H., Gabriel, J.M., Holtzman, D.M., Cohen, F., Fletterick, R. andPrusiner, S.B. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10940-10944.

106. Hay, B., Barry, R.A., Lieberburg, I., Prusiner, S.B.and Lingappa V.R. (1987). *Mol. Cell Biol.* **7**, 914-919.

107. Bazan, J.F., Fletterick, R.J., McKinley, M.P. and S.B. Prusiner (1987). *Protein Engineering* **1**, 125-135.

108. Borchelt, D. R., Scott, M., Taraboulos, A., Stahl, N. and Prusiner, S. B. (1990). *J. Cell Biology* **110**, 743-752.

109. Wessels, H.P. and Spiess, M. (1988). *Cell* **55**, 61-69.

110. Lipp, J.F., Haeuptle, M.T. and Dobberstein, B. (1989). *J. Cell Biol.* **109**, 2013

111. Spiess, M., Handschin, C. and Baker, K.P. (1989) *J. Biol. Chem.* **264**, 19117-19124.

112. Kehry, M., Ewald, S., Douglas, R., Sibley, C., Raschke, W., Fambrough, D. and Hood, L. (1980). *Cell* **21**, 393-406.

113. Simon, K., Perara, E. and Lingappa, V.R. (1987). *J. Cell Biology* **104**, 1165-1173

114. Katz, F. and Lodish. H. (1979). *J. Cell. Biol.* **80**, 416-426

115. Bonifacino, J.S., McCarthy, S.A., Maguire, J.E., Nakayama, T., Singer, D.S., Klausner, R.D. and Singer, A.(1990). *Nature* **344**, 247-251.

116. Hsiao, K., Scott, M., Foster, D., Groth, D.F., De Armond, S.J. and Prusiner, S.B. (1990), *Science*, **250**, 1587

117. Prusiner, S.B. (1993), *Arch. Neurol.*, **50**, 1129

118. Chuck, S.L and Lingappa (1993), *J. Biol. Chem.*

119. Skach, W.R., et al. (1993), *J. Biol. Chem.*

120. von Heijne, G. (1986), *EMBO J.*, **5**, 3021

121. Szczesna-Skorupa, E. and Kemper, B., *J. Cell Biol.*, **108**, 1237

122. Flanagan, J.M., Kataoka, M., Fujisawa, T. and Engelman, D.M. (1993), *Biochemistry*, **32**, 10359

123. Popot, J.L. and Engelman, D.M. (1990), *Biochemistry*, **29**, 4031

124. Scott, M., Groth, D., Foster, D., Torchia, M., Yang, S.L., De Armond, S.J. and Prusiner, S.B., (1993), *Cell* , **73**, 979

125. Barry, R.A., Vincent, M.T., Kent, S.B.H., Hood, L. and Prusiner, S.B. (1988). *J. Immunol.* **140**, 1188-93

126. Hultman, T., Murby, M. Stahl, S., Hornes, E. and Uhlen, M. (1990), *Nucleic Acids Research*, **18**, 5107

127. Krieg, P.A. and Melton, D.A. (1984). *Nuc. Acids Res.* **12**, 7057-7070.

128. Taraboulos, A., Serban, D. and Prusiner, S.B. (1990), *J. Cell Biol.* , **110**, 2117

129. Myers, R.M., Tilly, K. and Maniatis, T.(1986), *Science*, **232**, 613