

# Combinatorial Control of Prion Protein Biogenesis by the Signal Sequence and Transmembrane Domain\*

Received for publication, February 21, 2001, and in revised form, May 16, 2001  
Published, JBC Papers in Press, May 18, 2001, DOI 10.1074/jbc.M101638200

Soo Jung Kim, Reza Rahbar, and Ramanujan S. Hegde‡

From the Laboratory of Cellular Oncology, NCI, National Institutes of Health, Bethesda, Maryland 20892

**The prion protein (PrP) is synthesized in three topologic forms at the endoplasmic reticulum.  $^{sec}$ PrP is fully translocated into the endoplasmic reticulum lumen, whereas  $^{Ntm}$ PrP and  $^{Ctm}$ PrP are single-spanning membrane proteins of opposite orientation. Increased generation of  $^{Ctm}$ PrP in either transgenic mice or humans is associated with the development of neurodegenerative disease. To study the mechanisms by which PrP can achieve three topologic outcomes, we analyzed the translocation of proteins containing mutations introduced into either the N-terminal signal sequence or potential transmembrane domain (TMD) of PrP. Although mutations in either domain were found to affect PrP topogenesis, they did so in qualitatively different ways. In addition to its traditional role in mediating protein targeting, the signal was found to play a surprising role in determining orientation of the PrP N terminus. By contrast, the TMD was found to influence membrane integration. Analysis of various signal and TMD double mutants demonstrated that the topologic consequence of TMD action was directly dependent on the previous, signal-mediated step. Together, these results reveal that PrP topogenesis is controlled at two discrete steps during its translocation and provide a framework for understanding how these steps act coordinately to determine the final topology achieved by PrP.**

The prion protein (PrP)<sup>1</sup> is a 35-kDa brain glycoprotein involved in the transmission and/or pathogenesis of several neurodegenerative diseases, including scrapie, bovine spongiform encephalopathy, and Creutzfeldt-Jakob disease (1–3). Previous studies examining PrP biogenesis revealed that the normal protein is synthesized in three topologic forms at the endoplasmic reticulum (ER) (4–6). The predominant form (termed  $^{sec}$ PrP) is fully translocated into the ER lumen, whereas the other two forms (termed  $^{Ntm}$ PrP and  $^{Ctm}$ PrP) are single-spanning membrane proteins.  $^{Ctm}$ PrP spans the membrane with its C terminus in the lumen, whereas  $^{Ntm}$ PrP is in the reverse orientation, with its N terminus in the lumen (see Fig. 1B for a diagram).

Mutations that result in increased generation of  $^{Ctm}$ PrP were shown to result in dose-dependent development of neurodegen-

erative disease (4, 5). Additionally, a human disease causing mutation in PrP (A117V, resulting in Gerstmann-Sträussler-Scheinker disease (7–10)) was shown to result in increased generation of  $^{Ctm}$ PrP *in vitro*, in mice (which also developed neurodegenerative disease), and in humans (4, 5). Finally, recent studies have suggested that the ability to generate  $^{Ctm}$ PrP may also play a role in the neurodegeneration seen in transmissible forms of prion disease (4). Thus, elevated levels of  $^{Ctm}$ PrP appear to be one mechanism by which PrP is able to mediate neurodegeneration.

The role of  $^{Ctm}$ PrP in the pathogenesis of at least a subset of prion diseases highlights the importance of understanding the mechanisms by which PrP topology is determined and controlled. Generally, a protein's topology is thought to be unique and determined by "topogenic elements" encoded within the primary sequence (11–13). PrP contains three such topogenic sequences: an N-terminal signal sequence, generally used to target proteins to the ER (14); a hydrophobic stretch of amino acids that can serve as a transmembrane domain (TMD); and a C-terminal sequence for glycolipid anchor addition (15). However, not only do these elements fail to specify a homogeneous population of chains in a single topology, but the three topologic forms of PrP differ in *two* fundamental ways: localization of the N terminus ( $^{sec}$ PrP and  $^{Ntm}$ PrP have their N terminus in the ER lumen, whereas  $^{Ctm}$ PrP has it in the cytosol) and integration of the potential TMD into the lipid bilayer ( $^{sec}$ PrP is not integrated, whereas  $^{Ntm}$ PrP and  $^{Ctm}$ PrP are).

The region(s) of PrP that encode the key determinants for each of the topologic forms have not been clearly elucidated. The N-terminal signal sequence is likely to be necessary for at least targeting PrP to the secretory pathway. This is supported by the observation that deletion or replacement of this domain results in PrP being made as a cytosolic protein (16).<sup>2</sup> Whether the signal plays any role in topogenesis beyond its targeting function has not been studied. By contrast, previous studies demonstrating that mutations within or immediately preceding the TMD can alter the topologic forms of PrP generated have implicated this domain in PrP topogenesis (4–6). However, the role of this domain in the generation of each of the topologic forms remains obscure. Finally, results from Stewart and Harris (6) demonstrating that various mutations in the C terminus of PrP do not significantly affect topology suggest that regions C-terminal to the TMD may not play a significant role in topogenesis. This is consistent with previous observations that replacement of the entire C-terminal domain of PrP with a protein domain from globin does not significantly affect the generation of any of the topologic forms (17). Thus, other than a poorly defined role for the TMD, relatively little is currently understood about either the domain(s) involved or their respective role(s) in directing PrP topogenesis. This study

\* This work was supported by NCI, National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Laboratory of Cellular Oncology, NCI, NIH, Bldg. 36, Rm. 1D32, 9000 Rockville Pike, Bethesda, MD 20892. Tel.: 301-496-4855; Fax: 301-480-5322; E-mail: hegder@mail.nih.gov.

<sup>1</sup> The abbreviations used are: PrP, prion protein; ER, endoplasmic reticulum; TMD, transmembrane domain; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

<sup>2</sup> S. J. Kim, R. Rahbar, and R. S. Hegde, unpublished data.

was undertaken to elucidate a conceptual framework for understanding PrP topogenesis that would both provide tools and serve as a starting point for future mechanistic studies.

#### EXPERIMENTAL PROCEDURES

**Materials**—Rabbit reticulocyte lysate and dog pancreatic rough microsomes were prepared and used as described (Ref. 18 and references therein). Anti-PrP monoclonal antibody 3F4 was a gift from the laboratory of S. B. Prusiner. Restriction enzymes, other DNA-modifying enzymes, and SP6 RNA polymerase were from New England Biolabs Inc. RNase inhibitor was from Promega. Routine laboratory chemicals were of the highest quality available commercially from Sigma, Mallinckrodt Chemical Works, or ICN.

**Plasmid Constructions**—Standard techniques were used in the creation of all plasmid constructs (19). All constructs were made in the pSP64 vector (Promega). Signal sequence mutants were all derived from a modified wild-type Syrian hamster PrP construct that contained silent *NheI* and *AatII* sites introduced at codons 8 and 19, respectively. This plasmid was digested with the desired combination of *BglII* (immediately preceding the start codon), *NheI*, *AatII*, or *PflMI* (codon 26) and ligated to synthetic oligonucleotides encoding the desired mutations. Most of the TMD mutations were made by site-directed mutagenesis. The remaining TMD mutations were made by first introducing silent restriction sites (*BstBI* and *NdeI*) at codons 103 and 111 to facilitate replacement of selected regions surrounding the TMD with synthetic oligonucleotides encoding the desired mutations. Combinations of signal and TMD mutants were made by replacing the wild-type TMD region of the wild-type mature region (excised with either *KpnI* and *XbaI* or *Bsu36I* and *EcoRI*) with the relevant mutant TMD. All constructs were verified by automated sequencing.

**Cell-free Translocation Assays**—*In vitro* transcription with SP6 RNA polymerase, translation with rabbit reticulocyte lysate in the presence of [<sup>35</sup>S]methionine, and translocation into canine rough microsomal membranes have been described (Refs. 5 and 18 and references therein). Translations were carried out at 32 °C for 30 min. Proteolysis was with 0.5 mg/ml proteinase K for 60 min at 0 °C. Reactions were terminated with 5 mM phenylmethylsulfonyl fluoride and transferred into 10 volumes of 1% SDS and 0.1 M Tris (pH 8) preheated to 100 °C. Samples were either analyzed directly by SDS-polyacrylamide gel electrophoresis on 12% Tris/Tricine gels (20) or immunoprecipitated with anti-PrP monoclonal antibody 3F4 prior to SDS-polyacrylamide gel electrophoresis as previously described (5, 18). All of the translocation reactions shown were performed in the presence of a competitive peptide inhibitor of glycosylation (NH<sub>2</sub>-Asp-Tyr-Thr-COOH). Inhibition of glycosylation does not affect the ratios of topologic forms generated (5). It does, however, simplify the analysis since each topologic form is then represented by a single band, rather than the heterogeneous banding pattern seen with variable glycosylation.

**Quantitative Analysis of PrP Topology**—Quantitative ratios of topologic forms were determined by analysis of digitized autoradiographs on either Kodak X-Omat or BioMax films. Band size (in pixels) multiplied by mean band density (subtracted for film background) was used to assign a value to each band, followed by calculation of the appropriate ratios. It should be noted that the relative effect of each mutant was consistently and reproducibly observed in multiple experiments (with an effect as little as 10% difference in the formation of a topologic form being readily detectable). However, the absolute amount of the topologic forms generated for any given construct (including the wild type) varied from experiment to experiment, depending on temperature, time of translation, and batch of *in vitro* translation extract and microsomal membranes used. This is consistent with the observation that PrP topogenesis is dependent on multiple protein factors in both the cytosol (21) and ER membrane (18). For this reason, the data in Figs. 2 and 4 represent quantitative analysis of experiments in which all of the signal mutants (Fig. 2) or TMD mutants (Fig. 4) were analyzed simultaneously, in triplicate, to allow accurate and direct comparisons between the mutants. The raw data shown in Figs. 1 and 3 are from translation reactions performed at another time with different batches of reagents. Thus, although the relative differences between the mutants in these experiments and the quantitative analyses in Figs. 2 and 4 are similar, the absolute amounts of each topologic form are somewhat different.

#### RESULTS

**Signal Sequence Mutants and PrP Topology**—During the course of our ongoing studies of signal sequence function, we noticed that replacing the signal sequence of PrP with functional signal sequences from certain other secretory proteins

(for example, prolactin and angiotensinogen) resulted in a change in the ratio of topologic forms.<sup>2</sup> This observation raised the possibility that, in addition to facilitating targeting of nascent PrP to the ER, the signal sequence may play a role in PrP topogenesis. To explore this idea, we generated and analyzed the effect on topology of mutations introduced into the PrP signal sequence.

Signal sequences generally contain three domains. The h-region, a feature that is common among all signal sequences (22, 23), forms the hydrophobic core of at least 6 amino acids. Preceding the h-region in many signals is the n-region, a polar and often charged domain that is at the amino terminus of a signal sequence. The c-region is composed of the amino acids immediately preceding the signal sequence cleavage site. The polar n-region of the PrP signal (residues 1–7) was chosen for mutagenesis for three primary reasons. First, the n-region is highly divergent among signals of different proteins, varying in both length (from 1 to 17 amino acids) and net charge (ranging from –2 to +4) (22, 23). Second, mutations disrupting the h-region often severely impair the obligate targeting function of a signal sequence (24). And third, the signal sequences from other proteins that affected PrP topology most notably differed from each other in the n-region, particularly charged residues.<sup>3</sup> Thus, we focused our mutagenesis primarily on those changes that alter the net charge of the n-region of PrP (Fig. 1A).

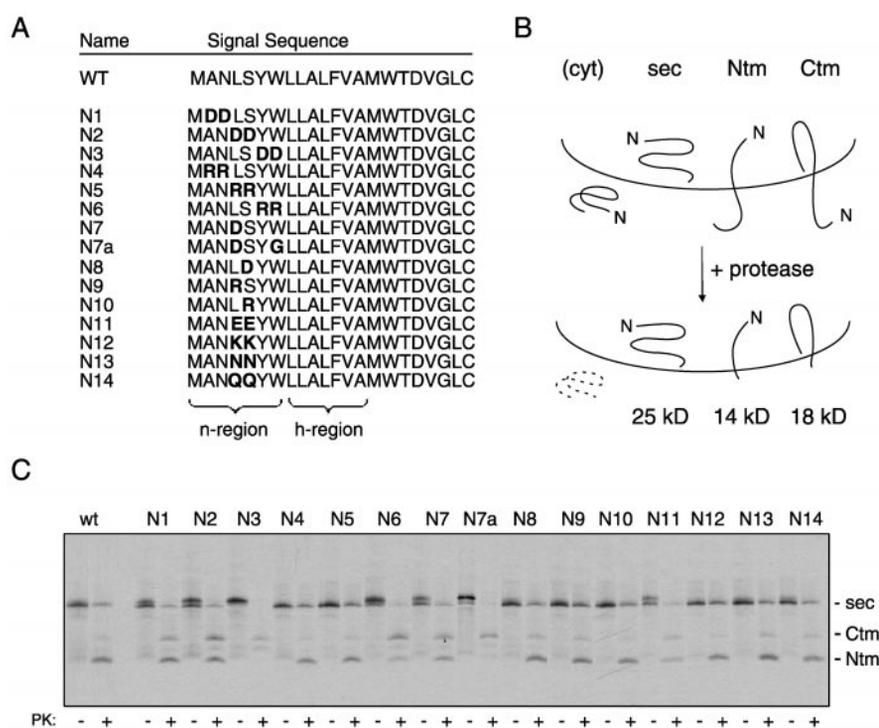
A protease protection assay was used to assess the topology of the n-region signal sequence mutants synthesized in an *in vitro* translation and translocation system (5) (Fig. 1B). In this assay, only PrP that is translocated across the microsomal membrane generates protease-protected species. The size of the fragment generated upon protease digestion indicates the topologic form from which it was derived (an ~18-kDa C-terminal fragment from C<sup>tm</sup>PrP and an ~14-kDa N-terminal fragment from N<sup>tm</sup>PrP), with protection of full-length PrP indicative of secPrP. As shown in Fig. 1C, many of the signal sequence mutations resulted in a significant change in comparison with the wild type in the relative amounts of the protease-protected PrP fragments. The N1, N2, N6, N7, N7a, and N11 mutants generated, to varying degrees, increased amounts of the 18-kDa C<sup>tm</sup>PrP fragment upon protease digestion. By contrast, we consistently observed that the N4, N5, N10, and N12 mutants generated slightly less C<sup>tm</sup>PrP than the wild type (see the quantitative analysis in Fig. 2 below).

The N3 and N7a mutants, unlike the other mutants, showed a clear discrepancy in the amounts of the protease-protected fragments relative to the amount of synthesized PrP. In both cases, less than half of the synthesized PrP could be accounted for after protease digestion of the sample (Fig. 1C, compare –PK and +PK lanes for these constructs). Presumably, these unaccounted chains are cytosolic and thus digested completely upon protease addition. This suggests that these two signal mutants, in addition to affecting the ratio of topologic forms generated, translocate less efficiently, resulting in some of the PrP remaining in the cytosol. Thus, although some mutations (e.g. N3) appear to reduce translocation efficiency, probably by affecting the targeting function of the signal sequence, other mutations (e.g. N2) appear to have a significant impact on topogenesis without an overall decrease in translocation efficiency. Together, these results suggest that, in addition to targeting, the signal may play a separate role in topogenesis.

To gain additional insight into which aspects of PrP topogenesis were most influenced by mutations in the signal sequence, we quantified (see “Experimental Procedures”) and plotted the relative ratios of the three topologic forms of PrP generated by

<sup>3</sup> R. S. Hegde, unpublished data.

**FIG. 1. Mutational analysis of the PrP signal sequence.** *A*, the 22-amino acid wild-type (WT) and mutant PrP signal sequences are shown, with mutated residues in *boldface type* and the n- and h- regions of the signal indicated below the sequences. *B*, shown is a schematic diagram of the topologic forms of PrP, including a cytosolically disposed form (*cyt*), and the protease protection assay used to discriminate between them. The N terminus of each chain is indicated to distinguish orientation. The approximate size of each fragment protected from protease digestion is also indicated. *C*, the signal sequence mutants shown in *A* were analyzed for their effect on PrP translocation and topology. Each of the mutants was translated in the rabbit reticulocyte lysate translation system in the presence of pancreatic microsomal membranes and a peptide inhibitor of glycosylation. Following translation, each sample was divided into two aliquots, one of which was digested with proteinase K (PK) as indicated. The positions of protease-protected full-length PrP (~25 kDa, indicative of <sup>sec</sup>PrP), the ~14-kDa NH<sub>2</sub>-terminal fragment (indicative of <sup>N<sup>tm</sup></sup>PrP), and the ~18-kDa COOH-terminal fragment (indicative of <sup>C<sup>tm</sup></sup>PrP) generated by proteinase K digestion are indicated to the right of the autoradiograph.



the signal mutants that did not significantly affect translocation efficiency (e.g. all mutants except N3 and N7a). As shown in Fig. 2, the <sup>C<sup>tm</sup></sup>PrP/<sup>sec</sup>PrP ratio varied dramatically between the different mutants, spanning an order of magnitude from ~0.17 (e.g. N12) to ~1.7 (e.g. N6) (Fig. 2A). The <sup>C<sup>tm</sup></sup>PrP/<sup>N<sup>tm</sup></sup>PrP ratio similarly ranged from ~0.28 to ~2.1 with these same mutants (Fig. 2B). By marked contrast, the <sup>sec</sup>PrP/<sup>N<sup>tm</sup></sup>PrP ratio was surprisingly invariant, being between ~1.1 and 1.8 regardless of the mutation analyzed (Fig. 2C).

These results suggest that the primary effect on topogenesis of signal sequence mutations is to increase or decrease the amount of <sup>C<sup>tm</sup></sup>PrP relative to both <sup>N<sup>tm</sup></sup>PrP and <sup>sec</sup>PrP. We have not observed a significant change in the ratio of <sup>N<sup>tm</sup></sup>PrP to <sup>sec</sup>PrP upon manipulation of the signal sequence with either mutations (Figs. 1 and 2) or replacement with other signal sequences.<sup>2</sup> Consequently, the amounts of <sup>N<sup>tm</sup></sup>PrP and <sup>sec</sup>PrP, relative to <sup>C<sup>tm</sup></sup>PrP, appear to either *both* increase (e.g. with N4) or *both* decrease (e.g. N2) with mutations in the signal sequence.

Two additional observations from the analysis of these mutants are noteworthy. First, introduction of charged residues into the n-region of the signal sequence generally alters the topogenesis of PrP. This does not appear to be due to disruption of an existent sequence motif, but rather an effect of the introduced charge. Thus, replacement of residues 4 and 5 with Asn or Gln (N13 and N14, respectively) does not alter topology, whereas Asp, Glu, Arg, or Lys (N2, N11, N5, and N12, respectively) significantly alters topology. Second, acidic residues in the n-region often result in increased <sup>C<sup>tm</sup></sup>PrP, and basic residues generally favor decreased <sup>C<sup>tm</sup></sup>PrP. However, this was not universally true. Introduction of Asp at position 4 (N7) resulted in a significant increase in <sup>C<sup>tm</sup></sup>PrP, whereas the same change at position 5 (N8) had a minimal effect on topology. Additionally, basic residues at positions 6 and 7 (N6) increased <sup>C<sup>tm</sup></sup>PrP significantly.

**Transmembrane Domain Mutants and PrP Topology**—We next performed a similar analysis of mutations either within or immediately preceding the hydrophobic stretch of amino acids (residues 113–135) thought to compose the membrane-spanning domain in both <sup>N<sup>tm</sup></sup>PrP and <sup>C<sup>tm</sup></sup>PrP (5, 6). Previous studies

have shown that mutations in this region of PrP can not only affect topology, but can also result in neurodegenerative disease (4, 5). However, the number of mutants analyzed was not sufficient to allow general conclusions to be drawn about the role of this domain in determining specific aspects of PrP topology. Over the past several years, numerous mutations have been made in this region for a variety of reasons, many unrelated to studies of PrP topogenesis. We took advantage of these existent mutants (Fig. 3A) to perform a careful analysis of their effects on PrP topogenesis. Just as with the signal sequence mutants, we reasoned that a systematic pattern may emerge that could provide insight into the role of the TMD in determining PrP topology.

The autoradiographs of representative TMD mutants show that, consistent with previous observations (4–6), mutations within or preceding the TMD can significantly affect PrP topology (Fig. 3B). However, it was also apparent that mutations in the TMD affect PrP topology in a qualitatively different manner than the signal mutants. First, unlike the signal mutants, several TMD mutants make essentially all of the translocated PrP in the <sup>sec</sup>PrP form (e.g. G123P, ΔSTE, and, to a lesser extent, A120G). And second, in contrast to some of the signal mutants (e.g. N3 and N7a), none of the 17 TMD mutants analyzed generated exclusively <sup>C<sup>tm</sup></sup>PrP or affected translocation efficiency.

To consolidate these observations, we quantitated and graphed the relative ratios of the topologic forms generated by each of the TMD mutants (Fig. 4). We found that the <sup>C<sup>tm</sup></sup>PrP/<sup>sec</sup>PrP ratio varied from <0.05 to ~2.4 among the various mutations (Fig. 4A). The <sup>sec</sup>PrP/<sup>N<sup>tm</sup></sup>PrP ratio also varied significantly, ranging from ~0.7 to >20 (Fig. 4C). By comparison, the <sup>C<sup>tm</sup></sup>PrP/<sup>N<sup>tm</sup></sup>PrP ratio was less variant, ranging from ~0.4 to ~2.3, with the majority of mutants showing a ratio of close to 1 (Fig. 4B). Therefore, it appears that the primary effect of mutations in the TMD region is to alter the amounts of the topologic forms that are membrane-integrated (<sup>N<sup>tm</sup></sup>PrP and <sup>C<sup>tm</sup></sup>PrP) relative to the fully translocated <sup>sec</sup>PrP. For example, some mutations, such as G123P and ΔSTE, make almost exclusively the <sup>sec</sup>PrP form. Others, such as KH-II, AV3, and A120L, make relatively little <sup>sec</sup>PrP in favor of increased <sup>N<sup>tm</sup></sup>PrP and <sup>C<sup>tm</sup></sup>PrP.

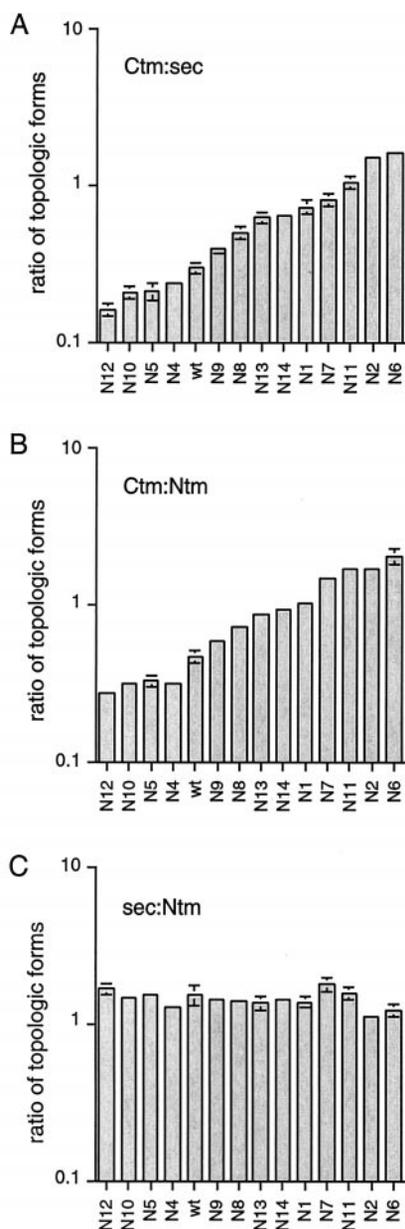


FIG. 2. The relative amounts of the topologic forms generated by each of the signal sequence mutants depicted in Fig. 1A (except N3 and N7a, both of which result in significantly decreased translocation efficiency) were quantitated (see “Experimental Procedures”), and the ratios between any two of the forms ( $C^{tm}PrP^{sec}PrP$  (A),  $C^{tm}PrP^{Ntm}PrP$  (B), and  $secPrP^{Ntm}PrP$  (C)) were plotted. The means and S.D. values (represented by the error bars) of each ratio for all of the mutants, calculated from triplicate determinations, are plotted. For some mutants, small S.D. values resulted in error bars too small to be displayed on the graphs. *wt*, wild type.

These results suggest that the major role for the TMD in PrP topogenesis is to specify the percent of PrP chains that integrate into the lipid bilayer.

**Combinatorial Control of PrP Topology**—The above analyses identified two domains within PrP that each contribute to the determination of topology. Although both of these domains affect topology, the quantitative analyses (Figs. 2 and 4) suggest that they affect it in qualitatively different ways. The general distinction is that the signal sequence increases or decreases  $C^{tm}PrP$  relative to the other two topologic forms, whereas the TMD increases or decreases  $secPrP$  relative to the other two topologic forms. Given that the signal and TMD affect PrP topogenesis in different ways, it was conceivable that

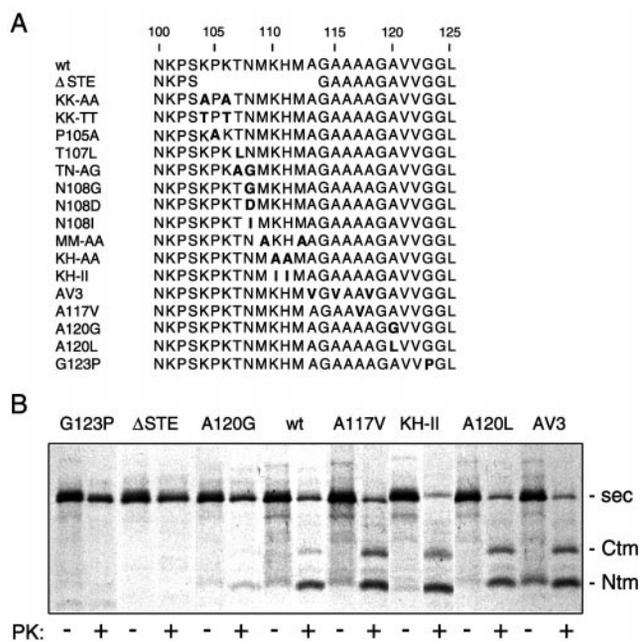


FIG. 3. **Mutational analysis of the PrP TMD.** A, the sequence of PrP surrounding the TMD is shown, with mutated residues in *boldface type*. B, representative TMD mutants shown in A were analyzed for their effect on PrP translocation and topology as described in the legend to Fig. 1. PK, proteinase K.

the effect of one domain is dominant to the other. Alternatively, the two domains may both contribute to PrP topogenesis by acting either sequentially or in concert. To determine the manner in which the signal and TMD mutants interact with each other, we generated and analyzed an array of PrP double mutants that contain various combinations of the signal sequence and TMD mutations (Fig. 5).

We analyzed constructs that combined various signal sequence mutations with either of two TMD mutants: PrP(A120L) and PrP( $\Delta$ STE) (Fig. 5, A and B). In addition, we also examined constructs in which various TMD mutations were introduced into PrP containing either the N4 signal sequence (Fig. 5C) or the N7a signal sequence (Fig. 5D). Together, the data in Fig. 5 reveal two important points. First, the double mutants do not display a phenotype that is identical to that of either of the individual single mutants. Thus, it appears that neither the signal sequence nor the TMD is dominant in its effect on PrP topology. Instead, each double mutant displays the combined characteristics of both the signal sequence and TMD. Second, the double mutants can be used to shift PrP topology primarily toward a single topologic form. For example, the N7a-A120L combination generates nearly all  $C^{tm}PrP$ , the N4-A120L combination generates primarily  $N^{tm}PrP$ , the N4- $\Delta$ STE combination makes all  $secPrP$ , and the N7a- $\Delta$ STE combination makes mostly cytosolic PrP. These observations suggest that the final topology achieved by PrP is the consequence of the combined action of the signal sequence and TMD. Thus, the signal and TMD, by serving distinct and separate roles during the topogenesis of PrP, cooperate to determine PrP topology.

**Evolutionary Conservation of Topogenic Determinants in PrP**—The above studies demonstrate that the topologic outcome of PrP can be readily and dramatically manipulated by any of numerous changes in either the signal sequence and/or the TMD. We found this quite surprising in light of two previous observations: (i) even very subtle changes in topology can influence the development of neurodegeneration (4, 5), and (ii) signal sequences and TMDs are generally considered to have

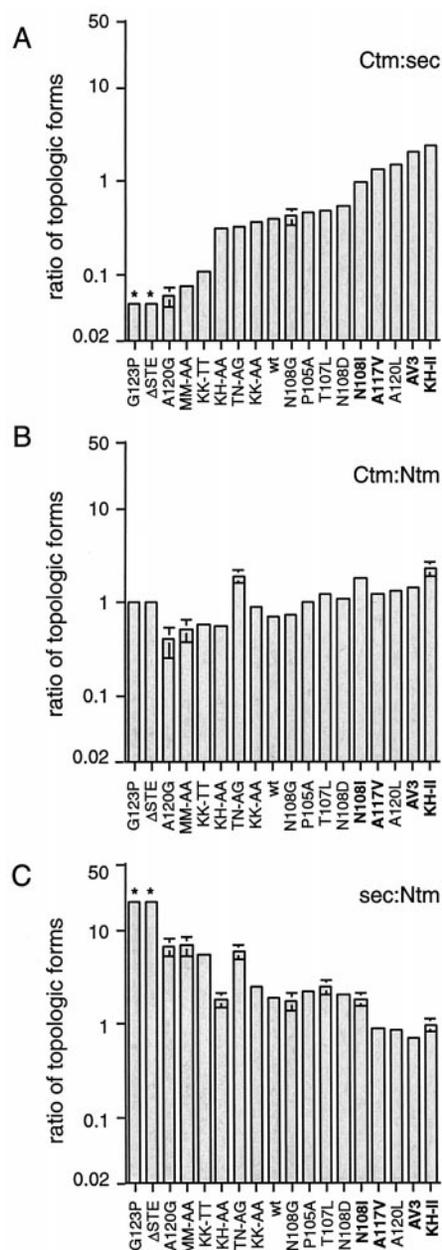


FIG. 4. The relative amounts of the topologic forms generated by each of the TMD mutants shown in Fig. 3A were quantitated, and the ratios between any two of the forms ( $C_{tm}PrP^{sec}PrP$  (A),  $C_{tm}PrP^{Ntm}PrP$  (B), and  $secPrP^{Ntm}PrP$  (C)) were plotted as described in the legend to Fig. 2. Ratios that were  $>20$  or  $<0.05$  are represented on the graph as either 20 or 0.05, respectively, and are indicated with *asterisks*. In addition, PrP mutants that have been shown in previous studies to result in neurodegenerative disease in transgenic mice (7, 8) are indicated in *boldface type*. *wt*, wild type.

few (if any) specific sequence motifs associated with them, instead being defined by very general features such as hydrophobicity (22, 23, 25, 26). Thus, polymorphic changes in either the signal sequence or TMD could potentially have effects on PrP that influence susceptibility to neurodegenerative disease. This raised the possibility that in the case of PrP, the signal sequence and TMD are under selective pressure to avoid polymorphic changes that could significantly affect topogenesis of this protein.

We explored this idea by examining the evolutionary conservation of these domains in PrP. If achieving the precise ratio of topologic forms of PrP is important to its normal function or critical to avoiding neurodegenerative disease, one might ex-

pect sequence conservation of these domains. As a first step, we compiled and compared the TMDs and signal sequences of all available mammalian PrP sequences (27, 28). Seventy-six mammalian PrPs have been cloned and sequenced through the TMD region, of which 42 have complete sequence data for the signal sequence. Alignment of the TMDs revealed a remarkable degree of conservation (25 of 29 residues are conserved across all 76 species), with no polymorphic changes in the hydrophobic domain from residues 113 through 128 (Fig. 6A).

By contrast, the signal sequence, while overall quite well conserved, contained only 4 of 22 residues that were invariant among all species (Fig. 6B). The most variable part of the signal sequences appeared to be the n-region. The signal sequences appeared to fall into one of two classes: those with an uncharged and shorter n-region (composed of MANLSYW-) and those with a longer, positively charged n-region (composed of MVKSHIGSW-). Such differences across species were surprising given that changes in the charge of the n-region can clearly impact PrP topology, which by extrapolation from prior studies (4, 5) may affect susceptibility to neurodegenerative disease. Thus, we wished to determine the topologic consequences of the polymorphic changes in the signal sequence, particularly the n-region. To explore this, we analyzed chimeras between a representative signal sequence with the charged n-region (bovine) and one with an uncharged n-region (Syrian golden hamster) (Fig. 6C).

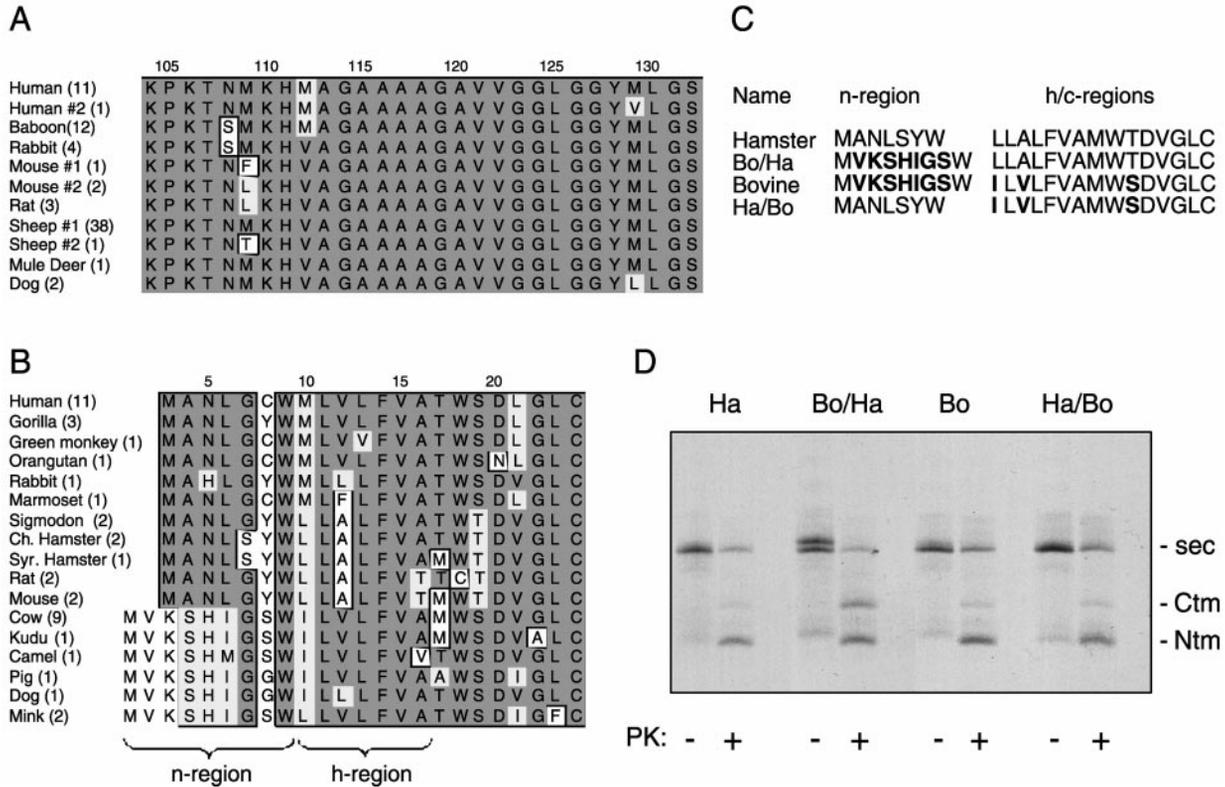
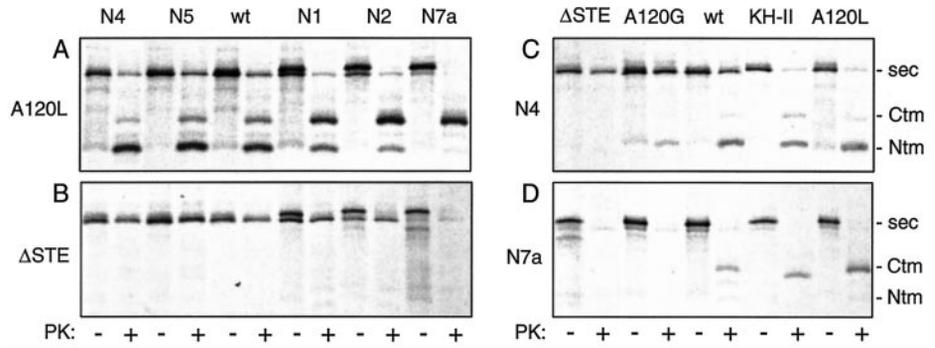
We found that a bovine/hamster hybrid signal containing the bovine n-region (*Bo/Ha*) resulted in a dramatic increase in  $C_{tm}PrP$  (Fig. 6D). This level of  $C_{tm}PrP$  is significantly more than is generated with the A117V mutation, which leads to human neurodegenerative disease. It is comparable to the KH-II and AV3 mutations, which lead to early and rapid development of disease in mice (*e.g.* compare with Fig. 3B). Quite remarkably, introducing the three very conservative polymorphisms into the h-region of bovine PrP (Leu to Ile, Ala to Val, and Thr to Ser) in a complete reversion to the normal ratio of topologic forms of PrP (Fig. 6D). Surprisingly, these three polymorphic changes in the absence of the charged bovine n-region (*Ha/Bo*) have no effect on topology. Thus, it appears that the topologically significant polymorphic changes in the n-region are fully and completely reversed by compensatory changes in the h-region, which by themselves are functionally silent. These results argue that the signal sequence, although displaying much more sequence variation than the highly conserved TMD, is functionally conserved with respect to its effect on PrP topology.

The data in Fig. 6 suggest that by conserving the topologic function of the signal and the exact sequence of the TMD, the unusual features of PrP topogenesis have also been conserved. It should be noted that, although the hydrophobic core of the TMD is highly conserved, there are a few variations in the residues immediately preceding it, changes that could conceivably affect topology. However, a comparative analysis of the topology of hamster, two mouse variant, and human PrPs has not revealed differences in the ratios of topologic forms that are generated (Ref. 5 and data not shown). Thus, it appears that the precise effects of the signal and TMD on the generation of each topologic form have been conserved.

#### DISCUSSION

In this study, we have analyzed a series of topology-altering mutants of the signal sequence and TMD of PrP to gain insight into how an initially homogeneous population of nascent chains can be made into three topologic forms. Several important conclusions can be drawn from our analyses. The first is that, in addition to its well established role in targeting proteins to the ER (14), the PrP signal sequence appears to encode deter-

**FIG. 5. Analysis of signal and TMD double mutants.** The translocation and topology of PrP constructs containing various signal sequences (indicated above the gels) combined with either the A120L (A) or ΔSTE (B) TMD mutation were analyzed as described in the legend to Fig. 1C. Also shown are the results of the analysis of PrP constructs containing various TMD mutations (indicated above the gels) containing either the N4 (C) or N7a (D) signal sequence. *wt*, wild type; *PK*, proteinase K.



**FIG. 6. Evolutionary conservation of PrP topogenic determinants.** A and B, the regions of the TMD and signal sequence of all available mammalian PrP species were assembled and aligned using ClustalW (48). The TMDs of 76 PrP sequences, representing 11 distinct variants, are shown. The signals of 42 PrP sequences (fewer PrP clones have been sequenced in this region than in the TMD region), representing 17 variants, are shown. A representative species for each sequence variant is indicated, along with the total number of species found to contain that sequence. *Ch.*, Chinese; *Syr.*, Syrian. C, shown is a diagram of Syrian hamster (*Ha*) and bovine (*Bo*) chimeric signal sequences containing different combinations of the n- and h-regions. D, the translocation and topology of PrPs containing the chimeric signal sequences shown in C were analyzed. *PK*, proteinase K.

minants for an additional role in directing topology. Interestingly, this second role could be perturbed with mutations to impact topology without significantly affecting overall translocation efficiency. Thus, it appears that the PrP signal sequence has two separate functions, in targeting and topogenesis, which is generally not seen with N-terminal signal sequences. This is consistent with the observation that the precise action of the topogenesis function of the PrP signal is evolutionarily conserved (Fig. 6) and that signal sequences from other proteins do not carry out this function in the same manner of the PrP signal.<sup>2</sup>

The second surprising finding in this study is the manner in which the signal and TMD act together to direct topology. Although it was entirely anticipated that the TMD would be involved in membrane integration, it was unexpected to find that its action was highly dependent on the signal. Thus, the combination mutants have provided the tools to generate nearly homogeneous populations of each of the topologic forms.

These constructs will be useful in identifying and distinguishing topology-specific interactions with the protein translocation or protein folding machinery, as well as for studying the differential metabolism of the individual topologic forms. Finally, the results obtained with the double mutants, together with the analyses of the individual mutants, have suggested a useful model for understanding PrP topogenesis (see below).

*An Unexpected Role for the Signal Sequence in PrP Topogenesis*—How does the PrP signal sequence play a role in determining PrP topology? Our interpretation of the quantitative analysis of the signal mutants (Fig. 2) is that the signal directs the segregation of nascent chains into two populations: <sup>Ctm</sup>PrP versus <sup>Ntm</sup>PrP<sup>sec</sup>PrP. One feature that distinguishes <sup>Ctm</sup>PrP from both the <sup>sec</sup>PrP and <sup>Ntm</sup>PrP forms is localization of the N-terminal domain: it resides in the cytosol for <sup>Ctm</sup>PrP, but is in the lumen for the <sup>Ntm</sup>PrP and <sup>sec</sup>PrP forms (see Fig. 1B). It is therefore plausible to view the action of the signal sequence as directing the localization of the N terminus of PrP. Signals such

as N2 and N6 favor a cytosolic localization for the N terminus, whereas the N4 and N10 signals favor a luminal disposition.

The mechanism by which the signal sequence is able to direct localization of the N terminus is currently not clear. Its actions are likely to be mediated, at least in part, by interactions with components of the protein translocation apparatus. After targeting to the ER membrane, signal sequences are recognized by components of the translocation channel (29, 30). It is thought that this recognition is important for achieving tight binding of the ribosome to the translocon and gating of the translocation channel (29, 31). Gating of the translocation channel involves closure of the ribosome-membrane junction (effectively shielding the nascent chain from the cytosol), followed by opening of a luminal gate (31) possibly composed of BiP (32). Only after the luminal gate opens is the N terminus of the nascent chain allowed access to the ER lumen. Thus, one attractive model of signal sequence action is that the signal is able to control localization of the N terminus by modulating the gating properties of the translocation channel. Future studies will focus on determining which of these post-targeting interactions between a signal sequence and components of the translocon play a key role in PrP topogenesis.

**The Role of the TMD in PrP Topogenesis**—The generation of either the  $^{Ctm}PrP$  or  $^{Ntm}PrP$  form requires integration of PrP into the lipid bilayer. This step is likely to be mediated by the TMD (33) and is supported by the observation that mutations in or near the TMD can significantly impact the generation of the membrane-integrated forms of PrP (Refs. 4–6 and 34 and this study). Previous studies of model transmembrane proteins suggest that membrane integration is a multistep process (35) that involves lateral gating of the translocation channel to allow the TMD access to the lipid bilayer (33). At present, the exact mechanism of gating and the role of the TMD in directing translocon gating are not well understood. However, it is reasonable to hypothesize that mutations in the TMD region of PrP act by either altering the partitioning of the TMD into the lipid bilayer or influencing the lateral gating properties of the translocon. Mutations that alter the hydrophobicity of the TMD (e.g. A117V, AV3, A120L, and A120G) may affect its lipid partitioning properties, as has been shown for other TMDs (33).

By contrast, mutations in the hydrophilic domain preceding the TMD (a domain that has been termed STE for stop transfer effector (34)) may act by disrupting or enhancing its interaction with components of the translocon to influence its lateral gating properties. Consistent with this idea, previous studies have shown that the STE domain, when placed adjacent to a heterologous potential membrane-spanning domain, can influence its integration into the lipid bilayer (34). Similar effects on membrane integration were observed with the STE-like sequence preceding the IgM TMD (36). In these studies, cross-linking experiments additionally suggested that the STE of IgM may act within the translocon to effect membrane integration of the adjacent TMD (36). Thus, it seems likely that the topologic effects seen with the mutations preceding the TMD of PrP could similarly act by interactions with components of the translocon to direct integration of the adjacent TMD.

**A Conceptual Model for Understanding PrP Topogenesis**—It is apparent from this and previous studies that a key step in generating multiple topologic forms during PrP topogenesis is the integration of some of the nascent PrP chains into the lipid bilayer. However, the topologic consequence of integration is dependent on the orientation of the nascent chain. If the nascent chain is in an orientation with the N terminus in the cytosol, then TMD-mediated integration would result in  $^{Ctm}PrP$ , whereas lack of integration would generate cytosolic PrP. Alternatively, if the nascent chain is in an orientation with the

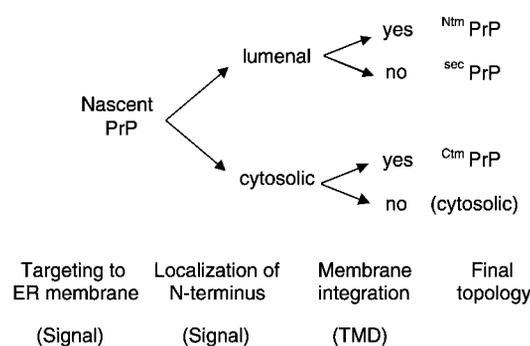


FIG. 7. **Schematic model of PrP topogenesis.** Following targeting to the ER, a homogeneous population of nascent PrP is segregated, by action of the signal sequence, into chains whose N termini are luminal versus cytosolic. Membrane integration, mediated by the TMD, determines the final topology achieved by each population of PrP chains. See "Discussion" for details.

N terminus in the ER lumen, then TMD-mediated integration results in  $^{Ntm}PrP$ , whereas lack of integration results in  $^{sec}PrP$ .

As discussed above, one principal determinant of PrP orientation with respect to localization of the N terminus is the signal sequence. This action of the signal, combined with the action of the TMD in mediating membrane integration, suggests a unifying model for how the three topologic forms of PrP can be generated (Fig. 7). We propose that, in addition to mediating targeting of PrP to the ER membrane, the signal sequence determines localization of the N terminus of PrP. Subsequently, the TMD determines whether PrP chains will be integrated into the lipid bilayer or not. Thus, the specific combination of decisions at the signal- and TMD-mediated steps determines the topology achieved by any given nascent chain. Heterogeneity of topologic forms therefore arises from the relative "inefficiencies" of the signal- and TMD-mediated steps, *i.e.* the signal does not quantitatively direct translocation of the N terminus, and the TMD does not quantitatively integrate into the membrane.

Such a model reconciles the qualitatively different effects of the signal and TMD mutants (Fig. 2 versus Fig. 4). Mutations in the signal sequence that affect topology would favor localization of the N terminus to either the lumen or cytosol. However, they would not significantly affect the process of membrane integration (mediated by the TMD), thereby explaining why the  $^{sec}PrP/^{Ntm}PrP$  ratio was largely unaffected by any of the signal mutants. On the other hand, TMD mutants would impact the membrane integration step, thereby increasing or decreasing  $^{Ntm}PrP$  and  $^{Ctm}PrP$ . Since these mutants would not affect the signal-mediated step, they would not have a significant impact on the  $^{Ctm}PrP/^{Ntm}PrP$  ratio.

The effects of the combination mutants (Fig. 5) are also readily interpreted in light of this paradigm. For example, the N4 signal sequence favors a luminal disposition for the N terminus. Fusing this signal to a TMD mutation that favors membrane integration (e.g. A120L) should generate predominantly  $^{Ntm}PrP$ , whereas a TMD that does not favor membrane integration (e.g.  $\Delta$ STE or A120G) should result in  $^{sec}PrP$ . By contrast, the N7a signal favors a cytosolic localization for the N terminus. With this signal sequence, a TMD that favors integration results in  $^{Ctm}PrP$ , whereas a TMD that does not favor integration results in cytosolic PrP. Thus, Fig. 5 demonstrates that by systematically and independently manipulating the localization of the N terminus (with mutations in the signal) and membrane integration (with mutations in the TMD), PrP can be predictably directed toward achieving any of the topologic forms.

Previous studies of single-spanning membrane proteins have

demonstrated that the key factors in determining topology include the length of the TMD (37–39), charged residues flanking the TMD (36, 40–42), and the folding properties of domains flanking the TMD (43). Although these features are clearly important for some membrane proteins, the orientation taken by the PrP TMD appears largely to be mediated by sequences not included in or adjacent to the TMD. This is supported by the demonstration that the same TMD (for example, A120L or KH-II) can be made to span the membrane efficiently in either orientation, dependent on the action of the signal sequence (Fig. 5). Thus, the generation of each topologic form is not decided by a single domain within PrP, but rather is dependent on the combinatorial action of at least two domains.

Cooperation between more than one topogenic element is not unprecedented in the biogenesis of multi-membrane-spanning proteins like the multidrug-resistant protein MDR1 (44, 45). In these studies, more than one TMD was proposed to act in concert to mediate proper membrane integration during the assembly of this complex membrane protein. In addition, topologic heterogeneity resulting from the inefficient action of a topogenic element has also been observed, again in multi-membrane-spanning proteins such as MDR1 and CFTR (46, 47). Thus, it may be that the biogenesis of PrP involves simplified variations of events that occur more commonly in multi-membrane-spanning proteins. Thus, mechanistic studies of combinatorial control of PrP topology may yield insights into similar events occurring more generally in more complex membrane proteins.

**Implications for Prion Disease**—The combinatorial nature of PrP topogenesis has several implications for the role of PrP in neurodegenerative disease. First, it appears that the manipulation of either step can potentially affect PrP topology in a manner that would lead to neurodegenerative disease. This has already been demonstrated in the case of the TMD, where mutations favoring the transmembrane forms of PrP lead to the generation of increased  $C^{tm}$ PrP and development of neurodegeneration in humans and transgenic mice (4, 5). One prediction based on the *in vitro* studies presented here would be that signal sequence mutations that lead to increased  $C^{tm}$ PrP (such as N2 and N7) may also lead to neurodegeneration in transgenic mice. The present work has generated a wide array of tools to test this and other hypotheses in future studies.

The second implication is for the mechanism of pathogenesis of transmissible prion diseases. Previous studies have suggested that one of the consequences of accumulated PrP<sup>Sc</sup> is to lead, perhaps indirectly, to the generation of increased  $C^{tm}$ PrP (4). Although the mechanism relating PrP<sup>Sc</sup> accumulation to  $C^{tm}$ PrP generation remains unknown, the present study identifies two key steps in the generation of  $C^{tm}$ PrP that could potentially be modulated *in trans*. Further studies will be required to determine whether one consequence of PrP<sup>Sc</sup> accumulation is to impact either the signal- or TMD-mediated step in PrP biogenesis in a manner that results in increased  $C^{tm}$ PrP generation.

Finally, the evolutionary conservation of the determinants of PrP topology was surprising for two reasons. First, the domains that are conserved, the signal sequence and TMD, are traditionally thought to be highly variable sequences that contain minimal (if any) exact sequence requirements (22, 23, 25, 26). That they are functionally conserved to make a precise ratio of topologic forms argues further that topologic dysregulation of PrP is detrimental to the organism. The second surprising aspect of this conservation is that the ability to make the transmembrane forms of PrP has remained constant during evolution. Given that generation of the  $C^{tm}$ PrP transmembrane form may be involved in the pathogenesis of both genetic and

transmissible forms of prion disease, it is not clear why this feature of PrP topogenesis should be conserved. This is especially puzzling given that single amino acid changes outside of the mature region of PrP (in the signal sequence) can significantly reduce the level of  $C^{tm}$ PrP generated, which may reduce the susceptibility to neurodegeneration. It is possible that any susceptibility to neurodegeneration is manifested primarily after reproductive age, thereby minimizing the evolutionary pressure against the generation of  $C^{tm}$ PrP. Alternatively, a more provocative possibility is that the different topologic forms have normal functions. The importance of such normal functions may outweigh the potentially increased susceptibility to neurodegeneration, thereby explaining the conservation of topologic regulation of PrP. Thus, an understanding of the mechanisms of PrP topogenesis may not only have implications for its role in neurodegenerative disease, but may provide some insight into the presently unknown normal functions of PrP.

**Acknowledgments**—We thank Devarati Mitra and Jeff Salerno for help with some of the constructs; Brigitte Bogert, Jeff Salerno, and Erik Snapp for valuable comments on this manuscript; and Paul Randazzo and Doug Lowy for encouragement and stimulating discussions. We are also indebted to Vishu Lingappa and members of his laboratory, especially Tom Rutkowski, for many interesting and productive discussions and for providing several of the TMD mutation plasmids.

## REFERENCES

- Prusiner, S. B. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13363–13383
- Johnson, R. T., and Gibbs, C. J., Jr. (1998) *N. Engl. J. Med.* **339**, 1994–2004
- Weissmann, C. (1999) *J. Biol. Chem.* **274**, 3–6
- Hegde, R. S., Tremblay, P., Groth, D., DeArmond, S. J., Prusiner, S. B., and Lingappa, V. R. (1999) *Nature* **402**, 822–826
- Hegde, R. S., Mastroianni, J. A., Scott, M. R., DeFea, K. A., Tremblay, P., Torchia, M., DeArmond, S. J., Prusiner, S. B., and Lingappa, V. R. (1998) *Science* **279**, 827–834
- Stewart, R. S., and Harris, D. A. (2001) *J. Biol. Chem.* **276**, 2212–2220
- Hsiao, K. K., Cass, C., Schellenbert, G. D., Bird, T., Devine-Gage, E., Wisniewski, H., and Prusiner, S. B. (1991) *Neurology* **41**, 681–684
- Tranchant, C., Doh-ura, K., Warter, J. M., Steinmetz, G., Chevalier, Y., Hanauer, A., Kitamoto, T., and Tateishi, J. (1992) *J. Neurol. Neurosurg. Psychiatry* **55**, 185–187
- Mastroianni, J. A., Curtis, M. T., Oberholtzer, J. C., DaCosta, M. M., DeArmond, S., Prusiner, S. B., and Garbern, J. Y. (1995) *Neurology* **45**, 2042–2050
- Mallucci, G. R., Campbell, T. A., Dickinson, A., Beck, J., Holt, M., Plant, G., de Pauw, K. W., Hakin, R. N., Clarke, C. E., Howell, S., Davies-Jones, G. A., Lawden, M., Smith, C. M., Ince, P., Ironside, J. W., Bridges, L. R., Dean, A., Weeks, I., and Collinge, J. (1999) *Brain* **122**, 1823–1837
- Blobel, G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1496–1500
- Wessels, H. P., and Spiess, M. (1988) *Cell* **55**, 61–70
- Lipp, J., Flint, N., Haeuptle, M. T., and Dobberstein, B. (1989) *J. Cell Biol.* **109**, 2013–2022
- Walter, P., and Johnson, A. E. (1994) *Annu. Rev. Cell Biol.* **10**, 87–119
- Stahl, N., Borchelt, D. R., Hsiao, K., and Prusiner, S. B. (1987) *Cell* **51**, 229–240
- Ma, J., and Lindquist, S. (1999) *Nat. Cell Biol.* **1**, 358–361
- DeFea, K. A., Nakahara, D. H., Calayag, M. C., Yost, C. S., Mirels, L. F., Prusiner, S. B., and Lingappa, V. R. (1994) *J. Biol. Chem.* **269**, 16810–16820
- Hegde, R. S., Voigt, S., and Lingappa, V. R. (1998) *Mol. Cell* **1**, 85–91
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Schagger, H., and von Jagow, G. (1987) *Anal. Biochem.* **166**, 368–379
- Lopez, C. D., Yost, C. S., Prusiner, S. B., Myers, R. M., and Lingappa, V. R. (1990) *Science* **248**, 226–229
- von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105
- von Heijne, G. (1990) *J. Membr. Biol.* **115**, 195–201
- Silhavy, T. J., Benson, S. A., and Emr, S. D. (1983) *Microbiol. Rev.* **47**, 313–344
- Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
- von Heijne, G. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 167–192
- Schätzl, H. M., DaCosta, M., Taylor, L., Cohen, F. E., and Prusiner, S. B. (1995) *J. Mol. Biol.* **245**, 362–374
- Wopfner, F., Weidenhofer, G., Schneider, R., von Brunn, A., Gilch, S., Schwarz, T. F., Werner, T., and Schätzl, H. M. (1999) *J. Mol. Biol.* **289**, 1163–1178
- Jungnickel, B., and Rapoport, T. A. (1995) *Cell* **82**, 261–270
- Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T. A. (1996) *J. Cell Biol.* **134**, 25–35
- Crowley, K. S., Liao, S., Worrell, V. E., Reinhart, G. D., and Johnson, A. E. (1994) *Cell* **78**, 461–471
- Hamman, B. D., Hendershot, L. M., and Johnson, A. E. (1998) *Cell* **92**, 747–758
- Heinrich, S. U., Mothes, W., Brunner, J., and Rapoport, T. A. (2000) *Cell* **102**, 233–244
- Yost, C. S., Lopez, C. D., Prusiner, S. B., Myers, R. M., and Lingappa, V. R. (1990) *Nature* **343**, 669–672

35. Do, H., Falcone, D., Lin, J., Andrews, D. W., and Johnson, A. E. (1996) *Cell* **85**, 369–378
36. Falcone, D., Do, H., Johnson, A. E., and Andrews, D. W. (1999) *J. Biol. Chem.* **274**, 33661–33670
37. Wahlberg, J. M., and Spiess, M. (1997) *J. Cell Biol.* **137**, 555–562
38. Eusebio, A., Friedberg, T., and Spiess, M. (1998) *Exp. Cell Res.* **241**, 181–185
39. Harley, C. A., Holt, J. A., Turner, R., and Tipper, D. J. (1998) *J. Biol. Chem.* **273**, 24963–24971
40. Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschien, C., Wessels, H. P., and Spiess, M. (1991) *J. Biol. Chem.* **266**, 973–978
41. Parks, G. D., and Lamb, R. A. (1991) *Cell* **64**, 777–787
42. Spiess, M. (1995) *FEBS Lett.* **369**, 76–79
43. Denzer, A. J., Nabholz, C. E., and Spiess, M. (1995) *EMBO J.* **14**, 6311–6317
44. Skach, W. R., and Lingappa, V. R. (1993) *J. Biol. Chem.* **268**, 23552–23561
45. Borel, A. C., and Simon, S. M. (1996) *Cell* **85**, 379–389
46. Skach, W. R., Calayag, M. C., and Lingappa, V. R. (1993) *J. Biol. Chem.* **268**, 6903–6908
47. Tector, M., and Hartl, F. U. (1999) *EMBO J.* **18**, 6290–6298
48. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680