Article No. mb982219 J. Mol. Biol. (1998) 284, 1185–1189

JMB

Breaking the Camel's Back: Proline-induced Turns in a Model Transmembrane Helix

IngMarie Nilsson[1,2] and Gunnar von Heijne[1*]

[1]*Department of Biochemistry Stockholm University, S-106 91 Stockholm, Sweden*

[2]*Department of Biosciences Karolinska Institute, NOVUM S-141 57 Huddinge, Sweden*

**Corresponding author*

We have constructed model membrane proteins with hydrophobic segments of the general composition $Leu_{29}Val\text{-}Leu_n$ where $n = 10$ and 20, and have analyzed their transmembrane topology when inserted into microsomal membranes. These hydrophobic segments span the membrane once, even though they are twice as long as normal transmembrane helices. Strikingly, a single proline residue introduced near the center of the $Leu_{39}Val$ hydrophobic stretch induces the formation of two transmembrane segments separated by a tight turn. These results have implications for our understanding of membrane protein assembly in the endoplasmic reticulum, and for the development of techniques for predicting membrane protein topology.

© 1998 Academic Press

Keywords: membrane protein; protein structure; glycosylation; transmembrane helix; proline

Introduction

In integral membrane proteins of the helix bundle class (von Heijne, 1996), the membrane domain is formed by tightly packed, hydrophobic transmembrane α-helices (TMHs). TMHs can vary substantially in length (typical values range from 20 to 30 residues) and the connecting loops can be as short as three residues (Bowie, 1997). Their hydrophobic character has made it possible to develop hydrophobicity-based prediction methods that have reached rather satisfactory performance levels (Claros & von Heijne, 1994; Jones *et al.*, 1994; Persson & Argos, 1996; Rost *et al.*, 1996). Certain kinds of prediction errors have nevertheless proved difficult to avoid; one of the most common kinds is when a very long, apolar stretch is identified in a polypeptide chain. In such cases, it is often difficult to decide whether to predict a single, long TMH, or two closely spaced TMHs connected by a tight turn. *Ad hoc* rules have been formulated in order to deal with this situation (Rost *et al.*, 1995), though there is little data available on which to base such rules.

The problem we address here is whether very long hydrophobic segments (40-50 residues) can still form a single TMH, and whether proline mutations can induce a tight turn in such very long TMHs. By *in vitro* translation of model proteins in the presence of dog pancreas microsomes, we find that hydrophobic segments composed of up to 50 hydrophobic residues form single transmembrane helices, and that proline mutations in the central ten residues of a 40-residues long hydrophobic stretch induce the formation of a closely spaced pair of transmembrane helices: a "helical hairpin". These results suggest that there is a fine threshold between the formation of one and two TMHs, and open a way to derive an experimental turn propensity scale applicable to membrane proteins.

Results

A model system for studying turn-induction in transmembrane helices

Our model system is based on the well-characterized *Escherichia coli* inner membrane protein leader peptidase (Lep), a protein with two TMHs (H1, H2) that inserts into dog pancreas microsomes with both termini in the lumen (Nilsson & von Heijne, 1993; Figure 1). In the study reported here, the H2 transmembrane segment was replaced by poly-Leu segments (including one Val) flanked by four lysine residues on the N-terminal end and by a Gln-Gln-Gln-Pro stretch on the C-terminal end. All constructs were expressed *in vitro* in the absence or presence of rough microsomes. As illustrated in Figure 1, an acceptor site for *N*-linked glycosylation (Asn-Ser-Thr) placed 20 residues

Abbreviations used: TMHs, transmembrane α-helices; Lep, leader peptidase.

E-mail address of the corresponding author: gunnar@biokemi.su.se

0022–2836/98/491185–05 $30.00/0 © 1998 Academic Press

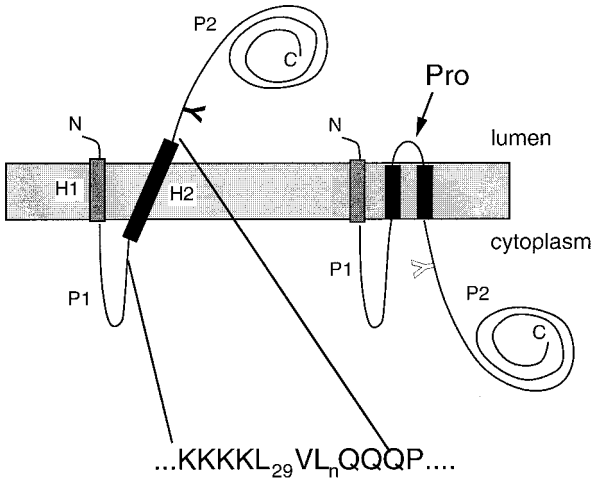

Figure 1. Model protein used in this study. The H2 transmembrane segment in Lep was replaced with the indicated stretch of residues, where $n = 10$ and 20. A glycosylation acceptor site was placed 20 residues downstream of H2 (counting from the first Gln after the hydrophobic stretch). Depending on the lumenal or cytoplasmic localization of the large C-terminal domain, the glycosylation acceptor site will either be modified (**Y**) or not (Ⓨ).

downstream of H2 served as a reporter to distinguish between a lumenal (glycosylated) and a cytoplasmic (non-glycosylated) localization of the C-terminal P2 domain, i.e. between the formation of a single TMH or a helical hairpin.

A 50 residues long hydrophobic stretch spans the membrane once

We have previously reported that model poly-Leu segments with between eight and 29 leucine residues and one valine form a single TMH when placed in the H2 position (Nilsson *et al.*, 1998). To determine whether very long poly-Leu stretches would still form only a single TMH, we made two additional constructs ($L_{39}V$, $L_{49}V$) composed of a stretch of 29 leucine residues, one valine, and an additional 10 or 20 leucine residues. As seen in Figure 2, both constructs were efficiently glycosylated, demonstrating that the P2 domain was translocated to the lumen and hence that the poly-Leu segments both formed a single TMH.

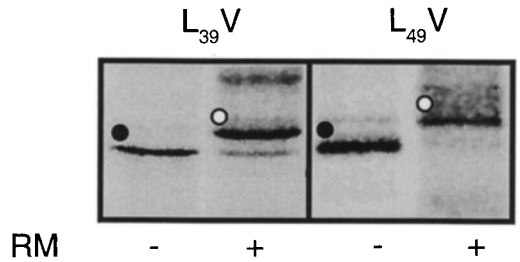

Figure 2. Long, uniformly hydrophobic segments form single transmembrane helices. The $L_{39}V$ and $L_{49}V$ constructs were translated *in vitro* in the absence (−) and presence (+) of rough microsomes and analyzed by SDS-PAGE. Filled and open circles indicate the non-glycosylated and glycosylated forms of the proteins, respectively.

Proline breaks a 40 residues long transmembrane helix when placed in the central ten positions

Given that the hydrophobic segments in the $L_{39}V$ and $L_{49}V$ constructs should, in principle, be long enough to form a pair of TMHs, we thought it possible that a helix-breaking residue such as proline introduced near the middle of the poly-Leu stretch might be able to induce the formation of a tight turn and thus give rise to a helical hairpin rather than a single, continuous TMH. In this case, the P2 domain would be located in the cytoplasm, and hence not become glycosylated, cf. Figure 1.

A series of Leu → Pro replacements were made in the $L_{39}V$ construct, and the degree of glycosylation was determined for each mutant. As seen in Figure 3(a) and (b), proline mutations in the ten residues at either end of the hydrophobic stretch allowed efficient glycosylation, whereas the constructs with a proline in any one of the ten central positions were not glycosylated, strongly suggesting that a helical hairpin is formed.

The topologies suggested by the glycosylation assay were further corroborated by protease protection experiments on the $L_{39}V$ and $L_{39}V$-P22 constructs (Figure 3(c). As expected, the P2 domain in the $L_{39}V$ construct was protected by the microsomal membrane (lane 3), whereas it was fully accessible to proteinase K in the $L_{39}V$-P22 construct (lane 6).

To rule out the trivial possibility that protease sensitivity and lack of glycosylation was a consequence of lack of membrane insertion of the poly-Leu segment, we deleted H1 from a number of constructs with glycosylation sites in the P1 and P2 domains (Figure 4). As expected, the ΔH1-$L_{39}V$ construct was efficiently glycosylated in the P2 domain (Figure 4(a), lane 1) but not in the P1 domain on the N-terminal side of H2 (Figure 4(a), lane 2). The ΔH1-$L_{39}V$-P22 construct, in contrast, was not modified on either acceptor site (Figure 4(a), lanes 3 and 4). The ΔH1-$L_{39}V$-P22 construct was quantitatively retained in the membrane pellet after a sodium carbonate wash when the microsomes were present during translation but not when added after translation (Figure 4(b)), demonstrating efficient co-translational membrane insertion of the helical hairpin.

Discussion

How long can a hydrophobic transmembrane segment be? Here, we have made poly-Leu segments with up to 50 contiguous hydrophobic residues, and find that they still insert into microsomal membranes as a single TMH. The fact that a single

rather than a pair of TMHs is formed suggests that all hydrophobic residues are buried in the membrane and that the TMH is thus either strongly tilted or, if perpendicular to the membrane, that the regions near its ends bury themselves in the bilayer surface.

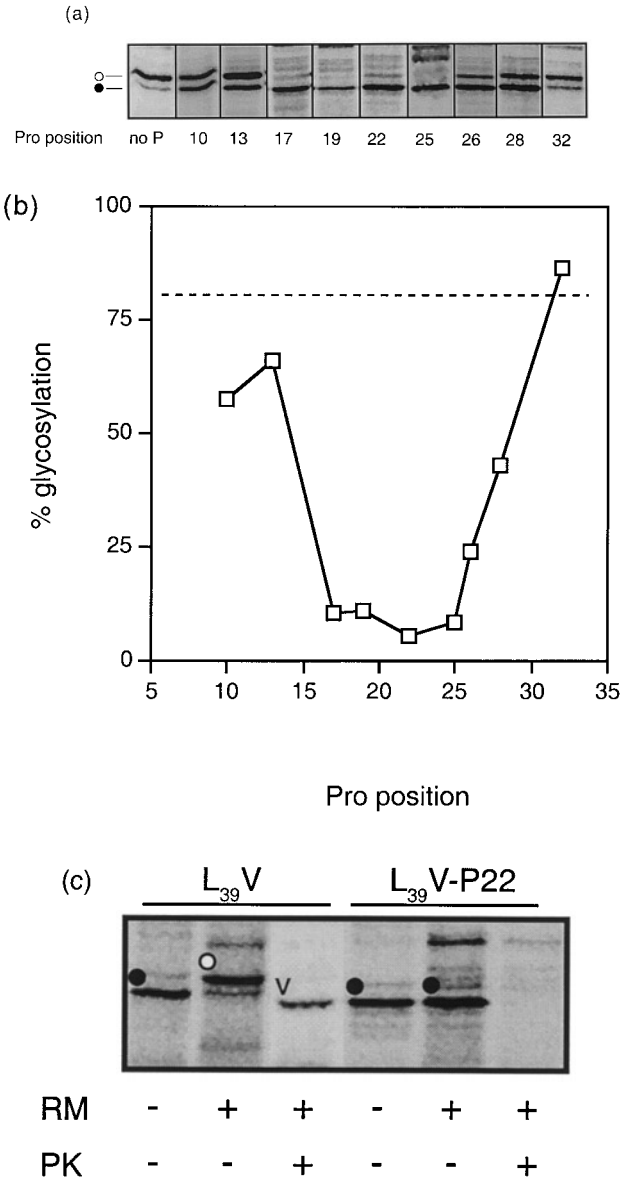

Figure 3. Single proline mutations in the center of the $L_{39}V$ TMH cause the formation of a helical hairpin. (a) The indicated constructs were translated *in vitro* in the absence or presence of rough microsomes (RM) and analyzed by SDS-PAGE. Filled and open circles indicate the non-glycosylated and glycosylated forms of the proteins, respectively. (b) Quantitation of the gels shown in (a). The percentage glycosylation was calculated as $100 \times I_+/(I_+ + I_-)$, where I_+ (I_-) is the intensity of the glycosylated (non-glycosylated) band. The broken line indicates the glycosylation efficiency of the $L_{39}V$ construct without proline residues. Proline positions are counted from the N-terminal end of the $L_{39}V$ stretch. (c) The $L_{39}V$ and $L_{39}V$-P22 constructs were translated in the absence or presence of rough microsomes, and then subjected to proteinase K (PK) treatment. Filled and open circles indicate the non-glycosylated and glycosylated forms of the proteins, respectively, and the protease-protected P2-fragment in the $L_{39}V$ construct is indicated by the arrow.

Remarkably, a single Leu → Pro replacement in any one of the ten central positions in the $L_{39}V$ TMH is enough to induce the formation of a pair of TMHs, a helical hairpin (Engelman & Steitz, 1981). The loss of hydrogen bonds and the steric problems caused by the proline residue thus make it energetically more favorable to place it in a tight turn near the membrane/water interface than to force it into the center of a long, membrane-embedded TMH. To be effective as a turn-promoter, the proline needs to be some 15 residues away from either end of the hydrophobic stretch, suggesting that a hydrophobic segment needs to be at least ~30 residues long before it can form two rather than one TMH. This has obvious implications for topology prediction methods, but the turn-inducing propensities of other residues besides proline need to be measured before a set of consistent rules can be formulated.

The transmembrane topology of a membrane protein is determined during the insertion of the nascent polypeptide chain into the Sec61 translocon in the microsomal membrane (Do *et al.*, 1996; Liao *et al.*, 1997; Mothes *et al.*, 1997). An internal signal-anchor sequence like H2 presumably enters the translocon in a loop or hairpin conformation, (Figure 5, top), which may facilitate the formation of a permanent helical hairpin in the presence of a turn-promoting residue (Figure 5, bottom). A stop-transfer sequence, on the other hand, presumably enters the translocon in a stretched conformation, and the effect of turn-promoting residues may be different; this possibility will be addressed in future studies.

Materials and Methods

Enzymes and chemicals

Unless otherwise stated, all enzymes were from Promega. T7 DNA polymerase, [^{35}S]Met, ribonucleotides, deoxyribonucleotides, dideoxyribonucleotides, and the cap analog m7G(5′)ppp(5′)G were from Amersham-Pharmacia (Uppsala, Sweden). Plasmid pGEM1, DTT, BSA, Sp6 RNA polymerase, RNasin and rabbit reticulocyte lysate were from Promega. Proteinase K, spermidine and PMSF were from Sigma. Oligonucleotides were from Kebo Lab (Stockholm, Sweden).

DNA manipulations

For cloning into and expression from the pGEM1 plasmid, the 5′ end of the *lep* gene was modified, first, by the introduction of an *Xba*I site and, second, by changing the context 5′ to the initiator ATG codon to a ''Kozak consensus'' sequence (Kozak, 1989). Thus, the 5′ region of the gene was modified to: ...ATAACCCTCTAGAGCCACCATGGCGAAT... (*Xba*I site and initiator codon underlined). Replacement of the H2 region in Lep was performed by first introducing *Bcl*I and *Nde*I restriction sites in codons 59 and 80 flanking the H2 region and

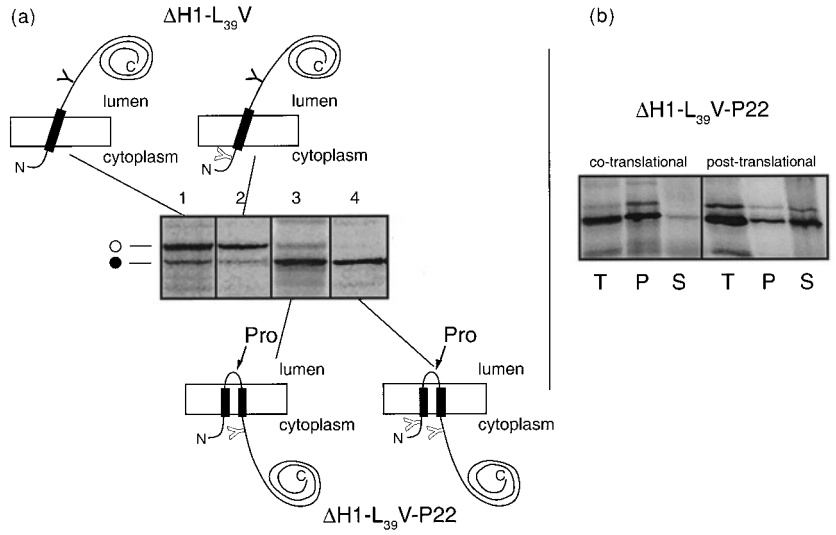

Figure 4. Topology mapping and alkaline extraction of the ΔH1-L_{39}V and ΔH1-L_{39}V-P22 constructs. (a) The indicated constructs were translated *in vitro* in the presence of rough microsomes and analyzed by SDS-PAGE. Filled and open circles indicate the non-glycosylated and glycosylated forms of the proteins, respectively. Construct ΔH1-L_{39}V (lane 1) lacks residues 5-46, i.e. H1 and about two-thirds of the P1 domain. The ΔH1-L_{39}V-derived construct in lane 2 has an additional glycosylation site added at the N terminus (see Materials and Methods). The two constructs shown in lanes 3 and 4 are the same as those in lanes 1 and 2, respectively, except that a Leu in position 22 in the middle of the hydrophobic segment has been replaced by Pro. The glycosylation acceptor site in the P1 domain was placed 21 residues upstream of H2, a distance previously shown to allow efficient glycosylation of a lumenally exposed N-terminal tail (Nilsson *et al.*, 1998; Nilsson & von Heijne, 1993). (b) Construct ΔH1-L_{39}V-P22 was translated *in vitro* either in the presence of rough microsomes (lanes marked co-translational) or with the microsomes added post-translationally (lanes marked post-translational). The microsomes were subjected to a sodium carbonate wash before loading onto the gel. T, total sample; P, pellet; S, supernatant.

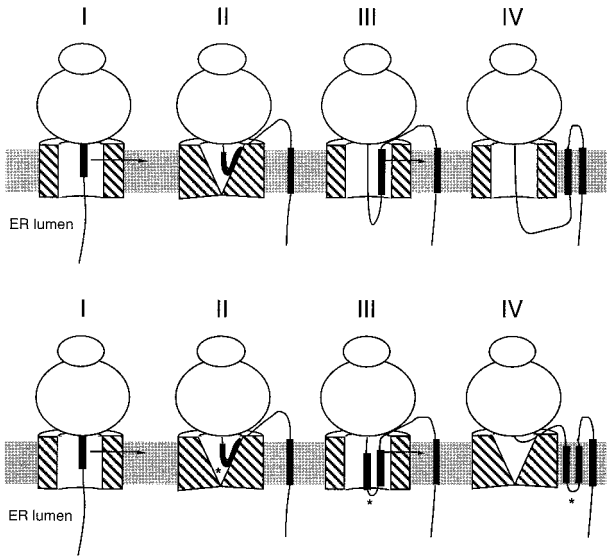

Figure 5. Models for the insertion of the H2 segment as a single transmembrane helix (top) and as a helical hairpin (bottom). A turn-inducing Pro residue is indicated by an asterisk. The N-terminal H1 segment targets the ribosome-nascent chain complex to the translocon and the N-terminal tail is translocated through the open translocation channel (I). H1 then moves laterally out of the translocon, and the channel closes. H2 enters the closed channel in a loop conformation (II), and triggers its re-opening. In the absence of a turn-inducing residue, H2 forms a single transmembrane helix and translocation continues (III, IV; top); when a turn-inducing residue is present, a permanent helical hairpin is formed (III; bottom), and the channel closes again (IV; bottom). The model is adapted from Mothes *et al.* (1997).

then replacing the *Bcl*I-*Nde*I fragment by the appropriate double-stranded oligonucleotides. Residues 57-81 in H2 were replaced by poly-Leu sequences of the general design PGLIKKKKL$_{29}$VL$_n$QQQP, where $n = 10$, 20. The ΔH1 constructs were made by deleting residues 5-46 in Lep. When indicated (Figure 4), an N-terminal glycosylation acceptor site (underlined) was added such that the N-terminal sequence became MA<u>NST</u>K$_{47}$... Site-specific mutagenesis used to add *Bcl*I and *Nde*I restriction sites at the 3′ and 5′ ends of H2 in Lep and to introduce Asn-Thr-Ser acceptor sites for *N*-linked glycosylation was performed according to the method of Kunkel (Geisselsoder *et al.*, 1987; Kunkel, 1985). Glycosylation acceptor sites were designed as described previously (Nilsson *et al.*, 1994), i.e. by replacing three appropriately positioned codons with codons for the acceptor tri-peptide Asn-Ser-Thr. All mutants were confirmed by DNA sequencing of plasmid or single-stranded M13 DNA using T7 DNA polymerase.

Expression *in vitro*

Synthesis of mRNA from pGEM1 by SP6 RNA polymerase and translation in reticulocyte lysate in the presence of dog pancreas microsomes were performed as described (Liljeström & Garoff, 1991). Proteinase K treatment was performed by adding $CaCl_2$ (13 mM final concentration) and proteinase K (67 μg/ml final concentration) to the translation mixture. After incubating 20 minutes on ice, PMSF was added (670 μg/ml final concentration) and the sample was further incubated for five minutes on ice before SDS-PAGE analysis. Expression of polypeptides was either in the presence of rough microsomes or with rough microsomes added post-translationally to the reaction mixture. Sodium carbonate extraction of microsomes was carried out as described (Sakaguchi *et al.*, 1987). Proteins were analyzed by SDS-PAGE and gels were quantified on a Fuji

BAS1000 phosphoimager using the MacBAS 2.1 software. The extent of glycosylation of a given mutant was calculated as the quotient between the intensity of the glycosylated band divided by the summed intensities of the glycosylated and non-glycosylated bands.

Acknowledgments

This work was supported by grants from the Swedish Cancer Foundation, the Swedish Natural and Technical Sciences Research Councils, and the Göran Gustafsson Foundation to G.v.H.

References

Bowie, J. U. (1997). Helix packing in membrane proteins. *J. Mol. Biol.* **272**, 780–789.

Claros, M. G. & von Heijne, G. (1994). TopPred II: an improved software for membrane protein structure prediction. *CABIOS,* **10**, 685–686.

Do, H., Falcone, D., Lin, J., Andrews, D. W. & Johnson, A. E. (1996). The cotransational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell,* **85**, 369–378.

Engelman, D. M. & Steitz, T. A. (1981). The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell,* **23**, 411–422.

Geisselsoder, J., Witney, F. & Yuckenberg, P. (1987). Efficient site-directed *in vitro* mutagenesis. *BioTechniques,* **5**, 786–791.

Jones, D. T., Taylor, W. R. & Thornton, J. M. (1994). A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry,* **33**, 3038–3049.

Kozak, M. (1989). Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell Biol.* **9**, 5073–5080.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA,* **82**, 488–492.

Liao, S., Lin, J., Do, H. & Johnson, A. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell,* **90**, 31–41.

Liljeström, P. & Garoff, H. (1991). Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor. *J. Virol.* **65**, 147–154.

Mothes, W., Heinrich, S., Graf, R., Nilsson, I., von Heijne, G., Brunner, J. & Rapoport, T. (1997). Molecular mechanisms of membrane protein integration into the endoplasmic reticulum. *Cell,* **89**, 523–533.

Nilsson, I. & von Heijne, G. (1993). Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* **268**, 5798–5801.

Nilsson, I., Whitley, P. & von Heijne, G. (1994). The C-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J. Cell Biol.* **126**, 1127–1132.

Nilsson, I., Sääf, A., Whitley, P., Gafvelin, G., Waller, C. & von Heijne, G. (1998). Proline-induced disruption of a transmembrane α-helix in its natural environment. *J. Mol. Biol.* PEW 251/98.

Persson, B. & Argos, P. (1996). Topology prediction of membrane proteins. *Protein Sci.* **5**, 363–371.

Rost, B., Casadio, R., Fariselli, P. & Sander, C. (1995). Transmembrane helices predicted at 95% accuracy. *Protein Sci.* **4**, 521–533.

Rost, B., Fariselli, P. & Casadio, R. (1996). Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci.* **5**, 1704–1718.

Sakaguchi, M., Mihara, K. & Sato, R. (1987). A short amino-terminal segment of microsomal cytochrome P-450 functions both as an insertion signal and as a stop-transfer sequence. *EMBO J.* **6**, 2425–2431.

von Heijne, G. (1996). Principles of membrane protein assembly and structure. *Prog. Biophys. Mol. Biol.* **66**, 113–139.

Edited by F. Cohen

(Received 15 May 1998; received in revised form 1 September 1998; accepted 1 September 1998)