Article No. jmbi.1999.2657 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (1999) 288, 141–145

JMB

A Turn Propensity Scale for Transmembrane Helices

Magnus Monné, Marika Hermansson and Gunnar von Heijne*

Department of Biochemistry
Stockholm University
S-106 91 Stockholm, Sweden

Using a model protein with a 40 residue hydrophobic transmembrane segment, we have measured the ability of all the 20 naturally occurring amino acids to form a tight turn when placed in the middle of the hydrophobic segment. Turn propensities in a transmembrane helix are found to be markedly different from those of globular proteins, and in most cases correlate closely with the hydrophobicity of the residue. The turn propensity scale may be used to improve current methods for membrane protein topology prediction.

Keywords: membrane protein; turn propensity; secondary structure; transmembrane helix

**Corresponding author*

Introduction

The formation of tight turns in globular proteins has been studied for decades, both experimentally and by statistical analysis of known structures, and reliable turn propensity scales have been established (Creighton, 1993; von Heijne, 1987). Remarkably, however, essentially nothing is known about the residue characteristics responsible for the formation of tight turns between transmembrane α-helices in integral membrane proteins. This is due in part to the paucity of high-resolution structural information for this class of proteins, but it is nevertheless surprising that no direct means of measuring turn propensities in membrane proteins has been established.

We recently developed a simple experimental system for the analysis of turn formation in transmembrane helices embedded in the membrane of the endoplasmic reticulum (Nilsson & von Heijne, 1998), and could show that a single proline residue placed near the middle of a 40 residue poly(Leu) transmembrane helix efficiently converts the poly(Leu) segment from a single, long transmembrane helix to a tightly spaced pair of transmembrane helices (a ''helical hairpin''). We have now used this system to measure the turn propensities for all the 20 naturally occurring amino acid residues, and present the first experimental propensity scale for the formation of tight turns between transmembrane α-helices. This scale is shown to differ in important respects from turn propensities in globular proteins, and in general correlates better with residue hydrophobicity than with ''classical'' secondary structure propensities.

Abbreviations used: BSA, bovine serum albumin.

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

Results

A system for measuring turn propensities in transmembrane helices under *in vivo*-like conditions

For these studies, we have used the well-characterized *Escherichia coli* protein Lep, which contains two transmembrane helices (H1 and H2) and a large C-terminal domain (P2). When expressed *in vitro* in the presence of dog pancreas microsomes, Lep has been shown to insert into the microsomal membrane with both the N and C termini on the luminal side (Nilsson & von Heijne, 1993), i.e. in the same orientation as it normally inserts into the inner membrane of *E. coli* (Wolfe *et al.*, 1983). Translocation of the P2 domain to the lumenal side is conveniently assayed by the glycosylation of a unique acceptor site for *N*-linked glycosylation (Asn-Ser-Thr) placed 20 residues downstream of H2 (Figure 1(a)). An advantage of this approach is that the microsomal *in vitro* system closely mimics the conditions of *in vivo* membrane protein assembly into the endoplasmic reticulum membrane.

For the studies reported here, H2 was replaced by a 40 residue poly(Leu) segment (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. Given that typical transmembrane helices are ~20-30 residues long (Bowie, 1997), this poly(Leu) stretch should, in principle, be able to form either one long or two closely spaced transmembrane helices. Indeed, the poly(Leu) segment has previously been shown to insert into the micro-

0022-2836/99/160141–05 $30.00/0

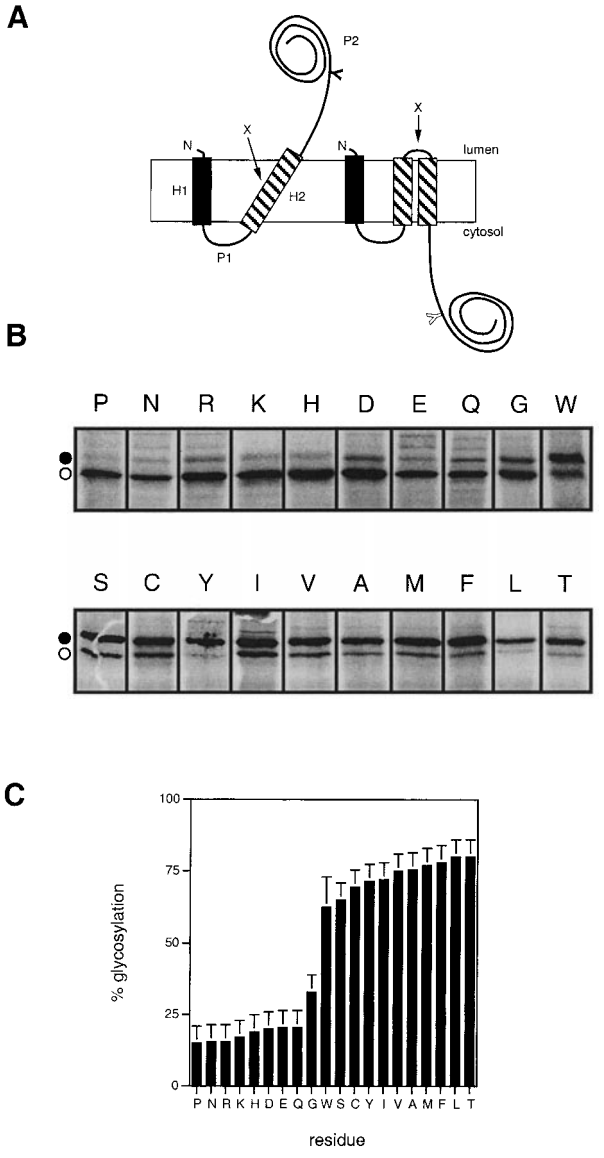

Figure 1. (a) Model protein used in this study. The H2 transmembrane segment in Lep was replaced with a stretch of residues of the general design $LIK_4L_{21}XL_7VL_{10}Q_3P$, where X is one of the 20 naturally occurring amino acid residues. A glycosylation acceptor site was placed 20 residues downstream of H2 (counting from the first Gln residue after the hydrophobic stretch). Depending on the luminal or cytoplasmic localization of the P2 domain, the glycosylation acceptor site will either be modified (**Y**) or not (Y). Note that the tilted conformation of the model single-spanning transmembrane H2 helix has not been experimentally proven. (b) The indicated L22 → X mutants were translated *in vitro* in the presence of rough microsomes and analyzed by SDS-PAGE. Black and white dots indicate the glycosylated and non-glycosylated forms of the proteins, respectively. (c) Quantification of the gels shown in (b). The percentage glycosylation was calculated as $100\ I_+/(I_+ + I_-)$, where I_+ (I_-) is the intensity of the glycosylated (non-glycosylated) band. From duplicate experiments on all the 20 constructs, the typical error in the determination of glycosylation efficiency was $\leqslant 5\,\%$ (bars), except for W where the error was $\pm 10\,\%$.

somal membrane as a single transmembrane helix with the P1 loop in the cytoplasm and the P2 domain in the microsomal lumen, and it was observed that the introduction of a Pro residue near the middle of the poly(Leu) stretch results in the formation of a helical hairpin in the membrane and localization of the P2 domain to the cytoplasmic side (Nilsson & von Heijne, 1998).

Using the same poly(Leu) construct, we have now measured turn propensities for all the 20 naturally occurring amino acids by substituting each residue (X) for Leu22 near the middle of the poly(Leu) segment, and expressing the L22 → X constructs *in vitro* in the presence of dog pancreas microsomes (Figure 1(b)). Based on the quantification of the glycosylation efficiencies given in Figure 1(c), a scale of turn propensities can be derived from this set of data (Table 1).

To rule out that lack of glycosylation is the result of inefficient insertion of the poly(Leu) stretch into the microsomal membrane rather than formation of a helical hairpin structure, a segment encompassing H1 and part of the P1 domain (residues 5-46) was deleted from two poorly glycosylated constructs (L22 → R and L22 → E), and membrane insertion of the L22 → R(Δ5-46) and L22 → E(Δ5-46) constructs was monitored by alkaline extraction of the microsomes. This treatment is known to remove peripherally bound membrane proteins, but leaves properly inserted transmembrane proteins in the membrane pellet (Fujiki *et al.*, 1982). In addition, the P2 domain (residues 79-323) was expressed alone to make sure that any membrane association observed for the Δ5-46 constructs was due only to the poly(Leu) segment. As seen in Figure 2, the two Δ5-46 constructs remained with the membrane pellet, whereas the P2 domain,

Table 1. Turn propensities for amino acid residues in a transmembrane helix

Residue	Turn propensity
A	0.5
C	0.6
D	1.6
E	1.6
F	0.4
G	1.3
H	1.6
I	0.6
K	1.6
L	0.4
M	0.5
N	1.7
P	1.7
Q	1.6
R	1.7
S	0.7
T	0.4
V	0.5
W	0.7
Y	0.6

The turn propensity is defined as $(1 - f_X)/\mu(1 - f_X)$, where f_X is the fraction of glycosylated molecules in the L22 → X mutant and $\mu(1 - f_X)$ is the mean value of $1 - f_X$ over all 20 residues. The typical error in the propensity values is ±0.05.

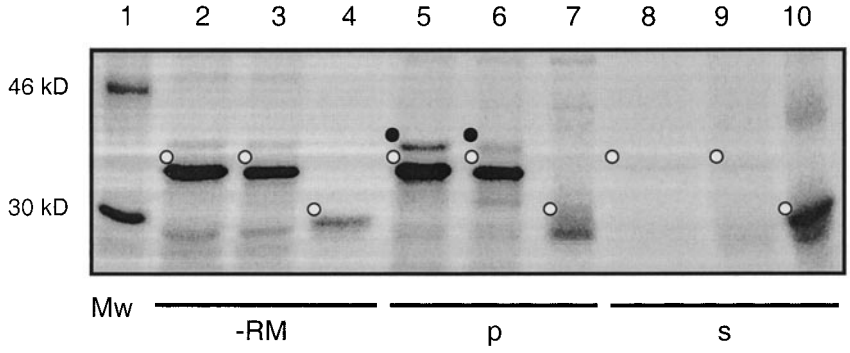

Figure 2. Alkaline extraction of the L22 → R(Δ5-46) (lanes 2, 5, 8), L22 → E(Δ5-46) (lanes 3, 6, 9), and P2 (lanes 4, 7, 10) constructs. Constructs were translated *in vitro* in the absence (lanes 2-4) or presence (lanes 5-10) of rough microsomes. In lanes 5-10, microsomes were subjected to a sodium carbonate wash before loading onto the gel. p, pellet; s, supernatant. Black and white dots indicate glycosylated and non-glycosylated forms of the proteins, respectively. The Δ5-46 constructs lack H1 and about two-thirds of the P1 domain, and the P2 construct lacks residues 2-78, i.e. the entire H1-P1-H2 domain. Lane 1 contains molecular mass markers as indicated.

when expressed alone, was found exclusively in the supernatant. We conclude that the degree of glycosylation seen for the different L22 → X mutants accurately reflects the fraction of molecules that insert with a single transmembrane segment *versus* a helical hairpin, and that it can thus be used as a basis for the turn propensity scale presented in Table 1.

Discussion

We have used a simple *in vivo*-like system where the membrane topology adopted by a 40 residue long model transmembrane segment can be used to directly infer a turn propensity scale relevant for transmembrane α-helices. As seen in Figure 1(c), hydrophobic residues (L, F, A, Y, V, and I) do not induce a turn in the poly(Leu) helix, whereas charged or highly polar residues do. In addition, the two classical helix breakers Pro and Gly both induce a turn (Pro somewhat more efficiently than Gly). Interestingly, despite their polar nature, Ser and Thr do not have high turn propensities. It is known from helices in globular proteins that Ser, Thr, and Cys side-chains can form hydrogen bonds to the polypeptide backbone (Gray & Matthews, 1984), which might increase their apparent hydrophobicity when present in a transmembrane helix, making turn formation less favorable. Consistent with this, Ser and Thr are rather frequently found in transmembrane helices, in contrast to Asn and Gln (von Heijne, 1992). Finally, Trp is known to have the strongest preference for the lipid-water interface of all the amino acid residues (Wimley & White, 1996), which may explain its somewhat higher turn propensity compared to, e.g. Phe and Tyr. Although the data reported in Figure 1(c) and Table 1 are largely consistent with a two-tier system where residues have either a high or a low turn propensity, it may be possible to provide a better discrimination between different residues in the transition region between high and low turn propensity by inserting pairs of residues in the middle of the model transmembrane segment; such studies are in progress.

The turn propensity scale derived here deviates significantly from the turn propensities observed in globular proteins (Figure 3, top panel). Thus, while the charged and highly polar amino acids all have high turn propensities in the transmembrane helix context, this is not the case in globular proteins. Ser, in contrast, has a rather high turn propensity in globular proteins but not in transmembrane helices. Pro and Gly are turn promoters in both contexts. The correlation between the turn propensity scale and the so-called interface hydrophobicity scale (Wimley & White, 1996) is not very strong (Figure 3, middle panel), whereas the correlation with helical propensities in *n*-butanol (Liu & Deber, 1998) is somewhat better (Figure 3, bottom panel). Good correlations are also obtained with some statistically defined hydrophobicity scales (von Heijne, 1992; results not shown).

In summary, we have measured the turn propensities for all the 20 naturally occurring amino acids placed in the middle of a poly(Leu) segment that is long enough to form either a single or two closely spaced transmembrane helices. Most of our results can be explained by hydrophobicity: all hydrophobic residues prefer the membrane environment over the membrane-water interface region (Wimley & White, 1996), and have low turn propensities. Conversely, the charged and highly polar residues induce turn formation in order to avoid the membrane interior. Thus, intrinsic conformational preferences become largely irrelevant in the context of a transmembrane helix, as observed previously in peptide studies of α *versus* β-structure formation in water, detergent, and lipid vesicle environments (Deber & Li, 1995; Li & Deber, 1994). In the context of the microsomal membrane, Pro behaves as a strongly polar residue, presumably because its inclusion in a transmembrane helix necessitates the disruption of at

least one hydrogen bond. Perhaps the most surprising result is that Gly has such a high turn propensity, since it neither has a polar side-chain, nor disrupts backbone hydrogen bonds when in a helix. Apparently, its exceptional conformation flexibility suffices to make the turn conformation preferred over the intact transmembrane helix. A possible mechanism for turn formation in a transmembrane helix during its insertion into the membrane of the endoplasmic reticulum has been suggested previously (Nilsson & von Heijne, 1998).

Finally, we anticipate that the turn propensity scale presented here will improve our ability to distinguish between cases of a single long and two closely spaced transmembrane helices when predicting membrane protein topology from amino acid sequence information.

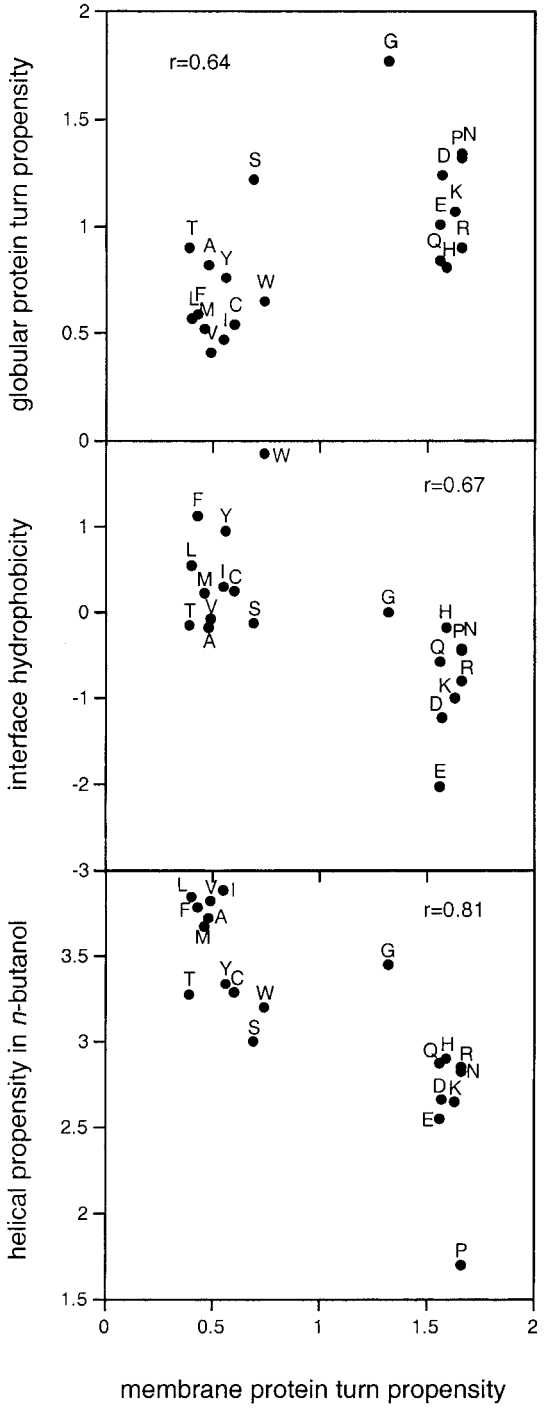

Figure 3. Turn propensities in transmembrane helices are different from turn propensities in globular proteins. Top, the turn propensity scale from Table 1 is plotted against a typical turn propensity scale for globular proteins (Williams *et al.*, 1987); middle, the interface hydrophobicity scale described by Wimley & White (1996); bottom, a scale of helical propensities in *n*-butanol (Liu & Deber, 1998). Correlation coefficients are indicated in the respective panels.

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, bovine serum albumin (BSA), SP6 RNA polymerase, RNasin and rabbit reticulocyte lysate were from Promega. Spermidine was from Sigma. Oligonucleotides were from Cybergene (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 (Johansson *et al.*, 1993; Kozak, 1989). 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 then replacing the *Bcl*I-*Nde*I fragment by the appropriate double-stranded oligonucleotides. Residues 59-80 in H2 were replaced by poly(Leu) sequences of the general design $LIK_4L_{21}XL_7VL_{10}Q_3P$, where X is one of the 20 naturally occurring amino acids; the Val residue results from the inclusion of a *Spe*I restriction site. The Δ(5-46) and P2 constructs were made by deleting, respectively, residues 5-46 and 2-78 in Lep. 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-Ser-Thr acceptor sites for N-linked glycosylation was performed according to the Kunkel method (Geisselsoder *et al.*, 1987; Kunkel, 1987) . Glycosylation acceptor sites were designed as described (Nilsson *et al.*, 1994), i.e. by replacing three appropriately positioned codons with codons for the acceptor tripeptide Asn-Ser-Thr. For the L22 → X substitutions, the QuickChange™ site-directed mutagenesis kit from Stratagene was used. Some of the primers were designed with a degenerate base in the second position of the codon for X in order to get more than one mutant from the primer pair. All mutants were confirmed by DNA sequencing of plasmid using T7 DNA polymerase.

Expression *in vitro*

The constructs in pGEM1 were transcribed by SP6 RNA polymerase for one hour at 37 °C. The transcription mixture was as follows: 1-5 μg DNA template, 5 μl of 10 × SP6 H-buffer (400 mM Hepes-KOH (pH 7.4),

60 mM Mg acetate, 20 mM spermidine-HCl), 5 μl of 1 mg/ml BSA, 5 μl of 10 mM m7G(5′)ppp(5′)G, 5 μl of 50 mM DTT, 5 μl of rNTP mix (10 mM ATP, 10 mM CTP, 10 mM UTP, 5 mM GTP), 18.5 μl of H_2O, 1.5 μl of RNase inhibitor (50 units), 0.5 μl of SP6 RNA polymerase (20 units). Translation was performed in reticulocyte lysate in the presence and absence of dog pancreas microsomes (Liljeström & Garoff, 1991) . Sodium carbonate extraction of microsomes was carried out as described (Sakaguchi *et al*., 1987) . Translation products were analyzed by SDS-PAGE and gels were quantified on a Fuji BAS1000 phosphoimager using the MacBAS 2.31 software. The glycosylation efficiency 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.

Creighton, C. (1993). *Proteins: Structures and Molecular Properties*, 2nd edit., W. H. Freeman & Co., New York.

Deber, C. M. & Li, S. C. (1995). Peptides in membranes: helicity and hydrophobicity. *Biopolymers,* **37**, 295-318.

Fujiki, Y., Hubbard, A. L., Fowler, S. & Lazarow, P. B. (1982). Isolation of intracellular membranes by means of sodium carbonate treatment. *J. Cell Biol.* **93**, 97-102.

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

Gray, T. M. & Matthews, B. W. (1984). Intrahelical hydrogen bonding of serine, threonine and cysteine residues within alpha-helices and its relevance to membrane-bound proteins. *J. Mol. Biol.* **175**, 75-81.

Johansson, M., Nilsson, I. & von Heijne, G. (1993). Positively charged amino acids placed next to a signal sequence block protein translocation more efficiently in *Escherichia coli* than in mammalian microsomes. *Mol. Gen. Genet.* **239**, 251-256.

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. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367-382.

Li, S.-C. & Deber, C. M. (1994). A measure of helical propensity for amino acids in membrane environments. *Nature Struct. Biol.* **1**, 368-373.

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.

Liu, L. P. & Deber, C. M. (1998). Uncoupling hydrophobicity and helicity in transmembrane segments. α-Helical propensities of the amino acids in non-polar environments. *J. Biol. Chem.* **273**, 23645-23648.

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. & von Heijne, G. (1998). Breaking the camel's back: proline-induced turns in a model transmembrane helix. *J. Mol. Biol.* **284**, 1185-1189.

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.

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. (1987). *Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit?*, Academic Press, San Diego.

von Heijne, G. (1992). Membrane protein structure prediction - hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225**, 487-494.

Williams, R., Chang, A., Juretic, D. & Loughran, S. (1987). Secondary structure predictions and medium range interactions. *Biochim. Biophys. Acta,* **916**, 200-204.

Wimley, W. C. & White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Struct. Biol.* **3**, 842-848.

Wolfe, P. B., Wickner, W. & Goodman, J. M. (1983). Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope. *J. Biol. Chem.* **258**, 12073-12080.

Edited by F. E. Cohen

(Received 16 December 1998; received in revised form 22 February 1999; accepted 22 February 1999)