J Membrane Biol (2009) 229:165–168
DOI 10.1007/s00232-009-9178-0

BRIEF COMMUNICATION

Strong Correlation Between Statistical Transmembrane Tendency and Experimental Hydrophobicity Scales for Identification of Transmembrane Helices

Gang Zhao · Erwin London

Received: 20 April 2009 / Accepted: 20 May 2009 / Published online: 12 June 2009

Abstract Direct physical chemistry measurements of the hydrophobicity of amino acids or their derivatives have often been used to estimate the propensity of amino acids to participate in transmembrane helices. In this short note, it is found that there is a very high degree of correlation ($r = 0.944$–0.965) between an average physical chemistry hydrophobicity scale (an average of scales derived, e.g., from the solubility of amino acid derivatives in organic solvents versus water or their binding to hydrophobic particles) and the statistically based transmembrane tendency scale (derived from the relative abundance of residues in known transmembrane and soluble protein sequences (Zhao and London, Protein Sci 15:1987–2001, 2006)). This correlation indicates that, other than hydrophobicity, amino acid properties/interactions that promote or inhibit transmembrane helix formation in a specific membrane protein largely cancel out when averaged over all transmembrane sequences. In other words, other than hydrophobicity, there are no properties of a specific amino acid residue within a hydrophobic segment that have a strong systematic effect upon transmembrane helix formation independent of the remainder of the sequence in that hydrophobic segment. However, proline is an exception to this rule.

Keywords Transmembrane protein · Hydrophobic alpha helices · Hydrophobicity · Biological hydrophobicity · Transmembrane tendency

G. Zhao · E. London (✉)
Department of Biochemistry and Cell Biology,
Stony Brook University, Stony Brook, NY 11794-5215, USA
e-mail: erwin.london@stonybrook.edu

Introduction

Prediction of transmembrane (TM) helices is a classical problem in biochemistry/molecular biology. The utility of scales that predict the TM segments of a protein based on their amino acid sequence is obvious. Considerable effort has been expended in trying to define the most accurate method to derive such scales. Two general approaches have been investigated. In the physical chemistry approach, the partitioning of amino acid-containing molecules between an aqueous and a hydrophobic environment is measured (Kyte and Doolittle 1982; Roseman 1988; Wolfenden et al. 1981). The hydrophobic environment can be a solvent or a lipid bilayer (Kessel et al. 2003; Kyte and Doolittle 1982; Wimley and White 1996). A biological analogue to this approach has been recently developed (Hessa et al. 2005, 2007, 2009; Xie et al. 2007). In this elegant method the ability of simple hydrophobic sequences in a chimeric protein to form a TM structure is used to evaluate the effect of amino acid sequence upon TM insertion (Hessa et al. 2005, 2007, 2009; Xie et al. 2007). The second approach is a statistical one in which the abundance of amino acids within TM helices and soluble sequences is compared. This can be used to develop a simple scale in which each amino acid is assigned a single value representing its propensity to be in a TM sequence (Zhao and London 2006) or as part of a more sophisticated analysis (i.e., hidden Markov methods) in which all the information in TM and surrounding sequences is used to predict TM helices (Kahsay et al. 2005). We previously derived a statistical type of TM tendency scale based on the amino acid composition of sequences in databases of soluble and TM sequences from both eukaryotes and prokaryotes (Zhao and London 2006). It was demonstrated that the TM tendency scale represents a scale at the theoretical limit of accuracy for a statistical

scale in which the abundance of amino acids in soluble and TM proteins is used to define a single average value for the propensity of each type of amino acid to occur in a TM helix. We also found a strong correlation between TM tendency and an experimental scale developed using natural membranes and based on what sequences processed by the translocon will form TM helices (Zhao and London 2006). This indicates that the TM tendency scale effectively captures the average behavior of natural sequences. In this brief note we show that an average physical chemistry-based hydrophobicity scale is highly correlated to TM tendency and consider the implications of this correlation.

Results and Discussion

Table 1 shows an average physical chemistry hydrophobicity scale (AvgH scale) we derived by averaging values from 15 different published scales that we were able to identify in the literature (Abraham and Leo 1987; Black and Mould 1991; Browne et al. 1982; Bull and Breese 1974; Cowan and Whittaker 1990; Deber et al. 2001; Jayasinghe et al. 2001; Kessel et al. 2003; Kyte and Doolittle 1982; Meek 1980; Parker et al. 1986; Roseman 1988; Wilson et al. 1981; Wolfenden et al. 1981). To avoid introducing any bias, we used all the scales we could find, rather than choose a specific subset of scales for detailed analysis. The scales were normalized to have an equal span of values between the most hydrophobic and most hydrophilic amino acid residues and, where necessary, reversed in sign so that larger values correspond to higher hydrophobicity. The resulting AvgH scale shows a close correlation ($r = 0.944$) to the TM tendency scale (Zhao and London 2006). Notice that this degree of correlation is generally higher than that of any individual experimental hydrophobicity scales with the TM tendency scale or their correlation with each other (Table 2). The good correlation between hydrophobicity (AvgH) and TM tendency indicates that they measure a very similar set of overall amino acid properties. The correlation is even better ($r = 0.965$) if Pro, which is an outlier that appears more hydrophobic in the physical chemistry scales than in the TM tendency scale (Fig. 1), is ignored. The fact that Pro has a lower TM tendency than expected based upon its physical chemical hydrophobicity may partly reflect its inability to form a proper backbone hydrogen bond in the context of a helix. This property would not be detected in hydrophobicity measurements using isolated Pro or other non-helix-forming model compounds containing Pro.

Table 1 AvgH scale with values ordered from most to least hydrophobic

Residue	AvgH
Phe	8.88
Leu	8.85
Ile	8.38
Trp	8.26
Val	7.19
Tyr	6.63
Met	6.58
Pro	5.84
Ala	5.46
Cys	5.15
Gly	4.55
Thr	4.28
Ser	3.82
His	3.47
Gln	3.18
Asn	2.80
Lys	2.24
Asp	1.74
Glu	1.71
Arg	1.61

The strong correlation between the AvgH and TM scales leads to the conclusion that (except for Pro) factors other than hydrophobicity have little systematic effect on whether a sequence forms a TM helix. This is not to say that specific polar interactions between residues do not have a crucial importance in TM helix formation in individual membrane proteins. Instead, it implies that there are no properties of amino acids other than hydrophobicity that strongly promote or interfere with TM helix formation without regard to what other residues are present in the TM helix.

This conclusion is based upon the use of AvgH, an average experimental hydrophobicity scale; and it is reasonable to ask: Why should this AvgH scale give a more accurate result than the individual experimental scales from which it is derived? There are various physical chemical methods to assess hydrophobicity (solvent partition, HPLC retention times, surface tension, vapor pressure, and computational methods) and various amino acid derivatives that can be used for such measurements. We made the assumption that most such scales derived from these approaches would have peculiarities specific to the system used to measure hydrophobicity. For example, if the solubility of some amino acid derivatives in different hydrophobic solvents/environments is measured, then the specific properties of that solvent/hydrophobic environment (polarity, hydrogen bonding capability) and the way it interacts with the specific amino acid derivative chosen could influence the scale in some systematic fashion. This is supported by the data in Table 2, which show that the correlation between different experimental hydrophobicity scales is often poor. However, if these peculiarities are

Table 2 Correlation between different experimental physical chemistry hydrophobicity scales and TM tendency

	KD	CW	AL	BB	Deber et al.	Roseman	Wolfenden et al.	Wilson et al.	Browne et al., TFA	Browne et al., HFBA	BM	Parker et al.	Meek	Jayasinghe et al.	Kessel et al.
TM tendency	0.88	0.94	0.85	0.79	0.85	0.92	0.75	0.78	0.75	0.63	0.91	0.87	0.75	0.90	0.81
KD	–	0.87	0.76	0.67	0.80	0.79	0.86	0.67	0.56	0.37	0.84	0.73	0.52	0.68	0.83
CW		–	0.88	0.80	0.79	0.96	0.80	0.82	0.75	0.63	0.96	0.90	0.77	0.91	0.81
AL			–	0.85	0.81	0.91	0.71	0.86	0.78	0.64	0.88	0.85	0.62	0.78	0.72
BB				–	0.84	0.79	0.50	0.83	0.81	0.73	0.86	0.91	0.76	0.74	0.55
Deber et al.					–	0.73	0.52	0.84	0.69	0.55	0.80	0.87	0.57	0.71	0.58
Roseman						–	0.80	0.82	0.76	0.62	0.96	0.86	0.73	0.91	0.84
Wolfenden et al.							–	0.53	0.50	0.27	0.80	0.53	0.44	0.57	0.93
Wilson et al.								–	0.65	0.49	0.84	0.83	0.57	0.76	0.56
Browne et al., TFA									–	0.93	0.77	0.85	0.79	0.73	0.54
Browne, et al., HFBA										–	0.60	0.79	0.84	0.68	0.34
BM											–	0.87	0.78	0.86	0.82
Parker et al.												–	0.81	0.86	0.59
Meek													–	0.82	0.51
Jayasinghe et al.														–	0.70

The scales shown are from the following references: TM tendency (Zhao and London 2006); KD (Kyte and Doolittle 1982); CW (Cowan and Whittaker 1990); AL (Abraham and Leo 1987); BB (Bull and Breese 1974); Deber et al. (2001); Roseman (1988); Wolfenden et al. (1981); Wilson et al. (1981); Browne et al., TFA or HFBA (Browne et al. 1982); BM (Black and Mould 1991); Parker et al. (1986); Meek (1980); Jayasinghe et al. (2001); and Kessel et al. (2003)

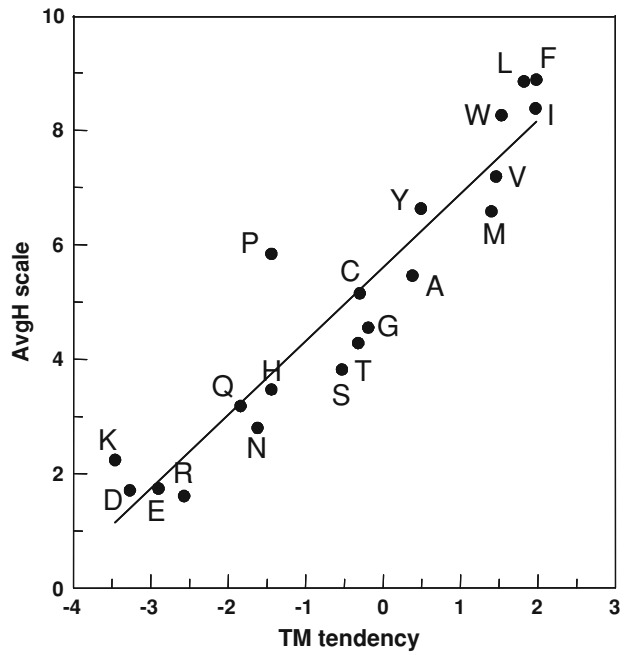

Fig. 1 Correlation between TM tendency and average hydrophobicity values. Zero values are arbitrary for these scales. Notice that the TM tendency of Pro is significantly lower than that expected based on its physical chemical hydrophobicity. See text for details

specific to each individual scale, they should tend to cancel out in the average of different scales. As the comparison to TM tendency shows, this assumption is likely to be valid.

Acknowledgement This work was supported by NIH grant GM 48596.

References

Abraham DJ, Leo AJ (1987) Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. Proteins 2:130–152

Black SD, Mould DR (1991) Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. Anal Biochem 193:72–82

Browne CA, Bennett HP, Solomon S (1982) The isolation of peptides by high-performance liquid chromatography using predicted elution positions. Anal Biochem 124:201–208

Bull HB, Breese K (1974) Surface tension of amino acid solutions: a hydrophobicity scale of the amino acid residues. Arch Biochem Biophys 161:665–670

Cowan R, Whittaker RG (1990) Hydrophobicity indices for amino acid residues as determined by high-performance liquid chromatography. Pept Res 3:75–80

Deber CM, Wang C, Liu LP, Prior AS, Agrawal S, Muskat BL, Cuticchia AJ (2001) TM Finder: a prediction program for transmembrane protein segments using a combination of

hydrophobicity and nonpolar phase helicity scales. Protein Sci 10:212–219

Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, Nilsson I, White SH, von Heijne G (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature 433:377–381

Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, Nilsson I, White SH, von Heijne G (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature 450:1026–1030

Hessa T, Reithinger JH, von Heijne G, Kim H (2009) Analysis of the transmembrane helix integration in the endoplasmic reticulum in *S. cerevisae*. J Mol Biol 386:1222–1228

Jayasinghe S, Hristova K, White SH (2001) Energetics, stability, and prediction of transmembrane helices. J Mol Biol 312:927–934

Kahsay RY, Gao G, Liao L (2005) An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics 21:1853–1858

Kessel A, Shental-Bechor D, Haliloglu T, Ben-Tal N (2003) Interactions of hydrophobic peptides with lipid bilayers: Monte Carlo simulations with M2delta. Biophys J 85:3431–3444

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Meek JL (1980) Prediction of peptide retention times in high-pressure liquid chromatography on the basis of amino acid composition. Proc Natl Acad Sci USA 77:1632–1636

Parker JM, Guo D, Hodges RS (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 25:5425–5432

Roseman MA (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. J Mol Biol 200:513–522

Wilson KJ, Honegger A, Stotzel RP, Hughes GJ (1981) The behaviour of peptides on reverse-phase supports during high-pressure liquid chromatography. Biochem J 199:31–41

Wimley WC, White SH (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat Struct Biol 3:842–848

Wolfenden R, Andersson L, Cullis PM, Southgate CC (1981) Affinities of amino acid side chains for solvent water. Biochemistry 20:849–855

Xie K, Hessa T, Seppala S, Rapp M, von Heijne G, Dalbey RE (2007) Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. Biochemistry 46:15153–15161

Zhao G, London E (2006) An amino acid "transmembrane tendency" scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: relationship to biological hydrophobicity. Protein Sci 15:1987–2001