Protein Engineering vol.12 no.7 pp.557–561, 1999

A simple method for predicting transmembrane α helices with better accuracy

M.Michael Gromiha

Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

The prediction of a protein's structure from its amino acid sequence has been a long-standing goal of molecular biology. In this work, a new set of conformational parameters for membrane spanning α helices was developed using the information from the topology of 70 membrane proteins. Based on these conformational parameters, a simple algorithm has been formulated to predict the transmembrane α helices in membrane proteins. A FORTRAN program has been developed which takes the amino acid sequence as input and gives the predicted transmembrane α-helices as output. The present method correctly identifies 295 transmembrane helical segments in 70 membrane proteins with only two overpredictions. Furthermore, this method predicts all 45 transmembrane helices in the photosynthetic reaction center, bacteriorhodopsin and cytochrome c oxidase to an 86% level of accuracy and so is better than all other methods published to date.

Keywords: conformational parameter/prediction/topology/transmembrane α helices

Introduction

Membrane proteins are important for a broad range of processes and functions in biological systems: for example, signal recognition, transport phenomena, energy translocation and conservation in the living cell (von Heijne, 1988; Jennings, 1989; Traxler *et al*., 1993). While the number of published amino acid sequences of membrane proteins is growing exponentially, there are few three-dimensional structures known to date; examples include the photosynthetic reaction center (Deisenhofer *et al*., 1985; Deisenhofer and Michel, 1989; Feher *et al*., 1989), bacteriorhodopsin (Henderson *et al*., 1990), porin (Weiss *et al*., 1991) and cytochrome c oxidase (Iwata *et al*., 1995). Hence, efficient predictive methods can be of help in the modelling of the structures of membrane proteins starting from the amino acid sequence, as obtained from DNA recombinant analysis.

Several methods have been proposed for the prediction of transmembrane α helices in membrane proteins. They are based on several different algorithms—hydrophobicity profiles (Esposti *et al*., 1990; von Heijne, 1992; Ponnuswamy and Gromiha, 1993; Hirokawa *et al*., 1998), neural networks (Rost *et al*., 1995, 1996; Casadio *et al*., 1996; Lohmann *et al*., 1996; Alloy *et al*., 1997), multiple alignment (Cserzo *et al*., 1994; Persson and Argos, 1994, 1997), consensus procedure (Parodi *et al*., 1994) and the dense alignment surface method (Cserzo *et al*., 1997).

In our previous work, we have developed a set of conformation parameters for membrane spanning β-strands and applied them successfully to the prediction of transmembrane β-strands in bacterial porins (Gromiha *et al*., 1997). In this article, a set of conformational parameters for all 20 amino acid residues in the transmembrane α helices of membrane proteins was developed from the topology of 70 membrane proteins. A simple method was proposed for the prediction of transmembrane α helices based on the conformational parameters. This method identifies the membrane spanning regions of 70 membrane proteins to an accuracy of 97% and predicts all the transmembrane segments in three proteins with known three-dimensional structures—PRC from *Rhodopseudomonas viridis*, bacteriorhodopsin and cytochrome c oxidase—to an 86% level of accuracy, better than any other previously published method. This algorithm has been automated with a computer program written in FORTRAN; the predictive results are available from the author.

Materials and methods

Databases

Three sets of data were used for the present study. First, a training set with 70 membrane proteins, whose topology is known experimentally. This data set was used to derive the conformational parameters. Each protein in this set contains from one to 12 membrane spanning segments, and the list of proteins, along with the number of transmembrane segments, is shown in Table I. The amino acid sequence and topology of the membrane spanning segments for all of the 70 membrane proteins were taken from the SWISSPROT database (Bairoch and Boeckmann, 1992). Second, a test set containing the experimental data for three membrane proteins—the photosynthetic reaction center (PRC) from *R.viridis* (Deisenhofer and Michel, 1989), bacteriorhodopsin (Henderson *et al*., 1990) and cytochrome c oxidase (Iwata *et al*., 1995)—whose three-dimensional structures are known at high resolution. The membrane spanning segments in these proteins were taken from the Protein Data Bank (PDB) of the Brookhaven National Laboratory (Bernstein *et al*., 1977). Third, a set of 150 non-homologous globular proteins, whose structures have been determined at high resolution. This set of proteins has representatives from all of the protein structural classes—all-α, all-β, α+β and α/β—and from proteins of varying size. The PDB codes of the proteins used in the present work are given in Table II.

Development of conformational parameters

A conformational parameter set for all 20 amino acid residues has been developed as described below. The frequency of occurrence (%) of amino acid residues was computed in the transmembrane helical part (f_m) of membrane proteins, followed by the occurrence (%) in the whole complex (f_t). The conformational parameter is computed using the equation

$$\alpha_i = (f_m)_i/(f_t)_i.$$

The set of 70 membrane proteins listed in Table I was used to derive the conformational parameters.

Table I. Training set of proteins used in the present study[a]

4f2_human (1)	5ht3_mouse (4)	a2aa_human (7)	a4_human (1)	aa1r_canfa (7)
aa2a_canfa (7)	acha_torca (4)	achd_torca (4)	acm1_human (7)	adt_ricpr (12)
b2ar_human (7)	cb21_pea (3)	cek2_chick (1)	coab_bpfd (1)	cxa1_rat (4)
trbm_human (1)	cyod_ecoli (3)	edg1_human (7)	egfr_human (1)	fce2_human (1)
gaa1_bovin (4)	glp_pig (1)	glpc_human (1)	gmcr_human (1)	gpt_crilo (10)
hema_cdvo (1)	pigr_human (1)	sece_ecoli (4)	secy_bacsu (10)	suis_human (1)
tcb1_rabit (1)	mcp1_ecoli (2)	gtr1_human (12)	hema_ndvu (1)	hb2a_human (1)
ha21_human (1)	myp0_human (1)	atnb_human (7)	atph_pea (2)	aa2b_human (7)
mprd_human (1)	dadr_human (7)	fml1_human (7)	olf2_rat (7)	pafr_cavpo (7)
trfr_mouse (7)	vg74_hsvsa (7)	acm2_human (7)	b3ar_mouse (7)	rdc1_canfa (7)
hema_pi4ha (1)	hg2a_human (1)	il2a_human (1)	lech_human (1)	lha_rhoru (1)
magl_mouse (1)	motb_ecoli (1)	nep_human (1)	oppb_salty (6)	ops2_drome (7)
ops4_drome (7)	opsb_human (7)	nep_human (1)	us27_hcmva (7)	vmt2_iaann (1)
brb2_human (7)	ers1_yeast (7)	mas_mouse (7)	reis_todpa (7)	v2r_human (7)

[a]The number of transmembrane segments in each protein is given in brackets.

Table II. PDB codes for the set of globular proteins used in the present study

1BAB	1BBL	1C5A	1FCS	1FIA	1GCN	1HBB	1HDD
1HIG	1HIL	1IFA	1LE4	1LIG	1MBC	1MBS	1PPT
1RPR	1SAS	1UTG	256B	2CCY	2LH1	2MHB	2MHR
2PDE	2TMV	2ZTA	3CYT	4CPV	4MBN	1CDT	1CID
1DFN	1EGF	1GPS	1HCC	1HIV	1HLE	1IXA	1MDA
1MON	1NXB	1PCY	1PHY	1RDG	1SHF	1TEN	1TIE
1TLK	1TNF	2ACH	2CTX	2ILA	2SNV	3CD4	3EBX
4GCR	7API	8I1B	1AAK	1BBP	1HIP	1LTS	1MS2
1NRC	1OVB	1OVO	1RND	1RNS	1SHA	1SNC	1TFG
1TGS	1THO	1TRX	2CDV	2GAP	2PAD	2SNS	3B5C
3IL8	3RUB	3SGB	4INS	9RSA	1ABA	1CIS	1CSE
1CTF	1DRI	1MLI	1OFV	1PAZ	1SRX	1Q21	2ACQ
2HAD	3ADK	3DFR	4FXN	1ADD	1CTS	1ADS	1AMY
1AOZ	1ARS	1BGU	1BNH	1BYB	1CHM	1GCA	1GPL
1GTR	1LGA	1LPB	1MIN	1MMO	1MPP	1MYP	1NNT
1OXY	1PBE	1PBP	1PDA	1PGS	1PHP	1RPA	1SBP
1SRP	1SRY	1YPT	2CAS	2LGS	2MTA	2POL	2TMD
2TS1	3MDD	7ICD	8ABP	8CAT	8XIM	1FCD	1GLC
	1HGJ	1RBL	2ATC	2BBK	2BPA	2KAU	

Table III. Conformational parameters for all 20 amino acid residues in the transmembrane α helices of inner membrane proteins

Residue	f_t	f_m	α
Ala	7.572	10.101	1.334
Asp	4.112	0.705	0.171
Cys	2.194	2.330	1.062
Glu	5.075	0.855	0.168
Phe	5.278	9.011	1.707
Gly	6.636	6.624	0.998
His	2.144	1.137	0.530
Ile	6.629	11.952	1.803
Lys	4.171	0.479	0.115
Leu	10.321	16.753	1.623
Met	2.594	3.664	1.413
Asn	4.108	2.030	0.494
Pro	4.991	2.809	0.563
Gln	3.258	1.118	0.343
Arg	4.623	0.780	0.169
Ser	7.140	5.835	0.817
Thr	6.110	5.121	0.838
Val	7.755	12.403	1.599
Trp	1.844	2.349	1.274
Tyr	3.443	3.946	1.146

Prediction of transmembrane α helices

Primary rule

Consider the amino acid sequence of a membrane protein. If the conformational parameter of a particular amino acid is ⩾0.80 (average value obtained from the set of 70 membrane proteins), then the index of priority assigned to that residue is 1, and if the value is <0.8, the index is taken to be zero.

$$\alpha_i \geqslant 0.8 \qquad \text{priority index} = 1$$
$$\alpha_i < 0.8 \qquad \text{priority index} = 0$$

Here, i varies from 1 to N, where N is the total number of residues.

Secondary rules

A set of secondary rules has been formulated to predict the transmembrane helical segments.

S1

Search for a continuous sequence of 18 points with higher priority (priority index = 1), with a maximum of three non-adjacent exceptions; pick up the segments and append the overlapping segments.

Table IV.

(a) The priority index of residues in bacteriorhodopsin

1011100011	1111111111	1111111101	1111000100	1111111101	1111111111
1111111101	1100001111	0110111110	1111011111	0100111111	1110111111
1111111011	1101111111	1111111111	1111111010	1100011111	0110011111
1111011111	1101111101	0101111111	0111011111	1110101111	0101000110
1111111					

(b) Prediction of transmembrane helices in bacteriorhodopsin

Step 1	Step 2	Step 3	Final predicted segments
9–34	9–34	–	9–34
41–72	41–72	–	41–72
77–102	82–102	–	82–102
105–162	105–157	105–127	105–127
–	–	134–157	134–157
176–200	176–200	–	176–200
202–230	202–223	–	202–223

S2

Collect all four consecutive overlapping residues with each of the appended segments obtained from S1 (e.g., if the appended segment is 1–25, the overlapping four residues are 1–4, 2–5, . . ., 22–25); check whether two zeros are present within any of these four-residue segments. If so, cut the segment with the high priority residue (priority index = 1) as the terminal one, and select the longer segment as the transmembrane helix.

S3

Longer segments (more than 40 residues) are divided into two segments; the terminal residues are fixed so that each segment contains a minimum number of zeros and sufficient number of residues (minimum of 18 residues) to be a transmembrane helix.

Accuracy of prediction

The accuracy of predicted segments was computed using the equation

$$\text{Accuracy } (\%) = [N - (N_u + N_o)]/N$$

where N, N_u and N_o are the total number of residues, the number of residues underpredicted and the number of residues overpredicted in a particular protein, respectively.

Results and discussions

Conformational parameter set for membrane spanning α helices

The set of conformational parameters for all 20 amino acid residues are given in Table III. It can be seen that the residues Ala, Cys, Phe, Gly, Ile, Leu, Met, Ser, Thr, Val, Trp and Tyr are more prevalent in the transmembrane α helices. It is interesting to note that all of the aromatic residues prefer the transmembrane regions, consistent with the study of Sciffer *et al*. (1992) on the importance of tryptophan residues in membrane proteins. Also, proline is not a favored residue (α = 0.563) in the transmembrane regions, as indicated by studies of Deber *et al*. (1990) and of the α helices of globular proteins (Barlow and Thornton, 1988; Gromiha and Ponnuswamy, 1995). It is noteworthy that the polar residues, Ser and Thr, prefer the membrane region of transmembrane helical proteins despite only marginally higher α values (0.82 and 0.84, respectively) than average (0.80).

Prediction of membrane spanning helices in membrane proteins

The present predictive method was applied to a set of 70 membrane proteins, which were used to derive the conformational parameters. The topology of all these proteins were known experimentally and the proteins within this set contain single and multiple membrane spanning segments traversing from two to 12 times. This method predicts the correct topology of 68 inner membrane proteins with experimentally determined topologies and correctly identifies 295 transmembrane segments with only two overpredictions.

Prediction of transmembrane helices in PRC, bacteriorhodopsin and cytochrome c oxidase

The predictive method was applied to three proteins—PRC, bacteriorhodopsin and cytochrome c oxidase—whose three-dimensional structures have been determined at high resolution. These proteins were not used to derive the conformational parameters and the jack-knife test determines the accuracy of the prediction.

As a working example, here the primary and secondary rules (explained in the methods section) are applied for the prediction of the transmembrane helices in the protein bacteriorhodopsin.

Primary rule

Consider the first five residues in the amino acid sequence of bacteriorhodopsin, AQITG. The conformational parameters for these residues are 1.334, 0.343, 1.803, 0.838 and 0.998, respectively. Hence, as per the primary rule, the priority index values of the five residues are 1, 0, 1, 1 and 1, respectively. The computed priority index values for each residues in bacteriorhodopsin are given in Table IV(a).

Secondary rules

Step 1 (search for segments with high priority values). First, rule S1 is applied to select the segments. The sequence was searched to find a stretch of 18 high priority residues with a maximum of three non-adjacent exceptions. The overlapping segments thus found—9–26, 10–27, ..., 17–34—were appended to give the first segment, 9–34. A continuous search identified the next segment, 41–72. In a similar manner, the whole sequence was searched. The segments obtained from step 1 are given in Table IV(b) (column 1).

Table V. Prediction of transmembrane α helices in PRC, bacteriorhodopsin and cytochrome c oxidase

Protein	Subunit and helix	Experiment[a]	Predicted
Photosynthetic reaction center (1PRC)	L1	33–53	21–53
	L2	84–111	80–102
	L3	116–139	114–134
	L4	171–198	172–198
	L5	226–249	232–259
	M1	52–76	47–74
	M2	111–137	110–133
	M3	143–166	144–161
	M4	198–223	199–229
	M5	260–284	266–289
Bacteriorhodopsin (2BRD)	1	10–32	9–34
	2	38–62	41–72
	3	80–100	82–102
	4	108–127	105–127
	5	137–157	134–157
	6	167–191	176–200
	7	203–225	202–223
Cytochrome c oxidase (1OCC)	A1	12–40	15–37
	A2	51–86	53–89
	A3	95–117	108–129
	A4	141–170	136–171
	A5	183–212	181–211
	A6	228–261	243–263
	A7	270–286	268–289
	A8	299–327	301–326
	A9	336–357	334–359
	A10	371–400	370–394
	A11	407–433	408–426
	A12	447–478	446–478
	B1	15–45	27–48
	B2	60–87	61–87
	C1	16–34	16–34
	C2	41–66	37–58
	C3	73–105	78–109
	C4	129–152	127–147
	C5	156–183	159–182
	C6	191–224	184–220
	C7	233–255	233–261
	D1	77–103	78–97
	G1	13–37	9–33
	I1	12–52	19–40
	J1	26–54	30–55
	K1	9–35	13–40
	L1	18–44	19–40
	M1	12–35	16–37
Total number of residues			2357
Number of residues correctly predicted			2026
Accuracy of prediction (%)			86.0

[a]Data from Nikiforovich (1998). PDB codes are given in brackets.

Step 2 (selection of the transmembrane segments). Consider the first segment 9–34; there are no two low priority values among any of the four consecutive residues in this segment. Hence, residues 9–34 were selected as the first segment. A similar pattern was observed for the segments 41–72 and 176–200. For segment 77–102, two low priority values for residues 81 and 84 were present near the N-terminus and hence the segment was cut at residue 81 and the segment 82–102 was selected; for segment 105–162, two low priority residues were observed at positions 158 and 160, near the C-terminus, and it was therefore cut at residue 157 and the segment 105–157 was selected as a transmembrane helix. Similarly, the segment 202–223 was selected as there were two low priority residues at 224 and 226. The selected segments from step 2 are given in column 2 of Table IV(b).

Step 3 (division of longer segments). A longer segment 105–157 of 53 residues was observed. The segment was divided into two segments in such a way that each of the segments contain a minimum number of low priority values and a sufficient number of residues (minimum of 18 residues) to be a transmembrane helix. Hence, the segments 105–127 and 134–157 were obtained.

The final predicted segments are given in Table IVb (last column).

In a similar way, the membrane spanning segments were predicted in the other two membrane proteins. The predicted transmembrane helical segments for all three proteins are shown in Table V, along with the experimentally-derived transmembrane segments, presented for comparison.

From Table V, it can be seen that the present method identifies all of the 45 transmembrane segments, and predicts 2026 residues correctly—that is, a predictive accuracy of 86%.

Testing the prediction of transmembrane-like helices in globular proteins

The present method was applied to a set of 150 soluble globular proteins to check whether this method predicts any transmembrane segments in these proteins. It correctly excluded 99% of the considered proteins to be of globular type. This result confirms that the present method excludes transmembrane-like helices present in globular proteins. This method was compared with three other methods, DAS (Cserzo *et al*., 1994), TMPRED (Rost *et al*., 1995) and SOSUI (Yanagihara *et al*., 1989), where the number of proteins correctly excluded were 125, 138 and 150, respectively.

Comparison with other methods

Recently, Nikiforovich (1998) proposed a non-statistical procedure for the prediction of transmembrane helical segments based on energy calculations and predicted the membrane spanning helices in PRC, bacteriorhodopsin and cytochrome c oxidase. The predicted results have been compared with four other recent methods, SOSUI (Yanagihara *et al*., 1989), DAS (Cserzo *et al*., 1994), TMPRED (Rost *et al*., 1995) and core (Nikiforovich, 1998). The transmembrane α helices for all

Table VI. Comparison of predictive ability (%) of six other methods with the present method for PRC, bacteriorhodopsin and cytochrome c oxidase

Method	Protein			
	PRC	Bacterio-rhodopsin	Cytochrome c oxidase	Average
Core (Nikiforovich, 1998)	83.9	79.4	76.8	78.8 (0)
DAS (Cserzo *et al*., 1994)	85.6	81.0	84.8	81.0 (0)
SURHYD (Ponnuswamy and Gromiha, 1993)	90.4	75.3	83.4	84.3 (1)
TMPRED (Rost *et al*., 1995)	88.8	86.2	82.6	84.5 (1)
SOSUI (Yanagihara *et al*., 1989)	89.1	80.2	84.9	85.4 (0)
SIGNAL (Nikiforovich, 1998)	91.4	86.6	84.6	86.5 (1)
Present work	86.1	81.0	86.7	86.0 (0)

Number of missed segments are given in brackets.

three proteins were predicted with the present method and the results compared with those of the five methods used by Nikiforovich (1998), together with the results of our earlier method, based on the hydrophobicity profile (Ponnuswamy and Gromiha, 1993). The results, presented in Table VI, show that the present method predicts all the transmembrane α helices in cytochrome c oxidase with a higher level of accuracy than all other methods, whereas the methods TMPRED (Rost *et al*., 1995) and SIGNAL (Nikiforovich, 1998) failed to identify one of the transmembrane segments. The method SIGNAL (Nikiforovich, 1998) and SURHYD (Ponnuswamy and Gromiha, 1993) have higher accuracies for the protein PRC. The higher accuracy of PRC by our previous method, SURHYD (Ponnuswamy and Gromiha, 1993) may be due to the fact that the information was taken from this protein to compute the surrounding hydrophobicity scale and the same scale was used for predictive purposes. Also, by considering all three proteins on both counts, number of transmembrane segments correctly predicted and percentage accuracy of prediction, the prediction performance of the present method is satisfactorily high among other methods. Further, this method predicts the transmembrane α helices for all proteins in the set with an accuracy of >80% and an average accuracy of 86%.

Conclusions

In this paper, a new set of conformational parameters for membrane spanning α helices was developed from the information about the topology of 70 membrane proteins. Then a primary rule was proposed to predict the transmembrane α helices of inner membrane proteins based on the application of conformational parameters. Based on the results, a set of secondary rules was proposed to extract the segments. The present method identified the membrane spanning helices in 70 membrane proteins with an accuracy of 97%. Furthermore, this method predicts all of the transmembrane α helices with an accuracy >80% for the proteins PRC, bacteriorhodopsin and cytochrome c oxidase individually and with an overall accuracy of 86%. These accuracy levels are superior to other methods published to date.

Availability of the program

The executable file, MICHELP, is available from the author and will be distributed upon request.

Acknowledgments

The author wishes to thank Dr V.T.Jacob for critical reading of the manuscript.

References

Aloy,P., Cedano,J., Oliva,B., Aviles,F.X. and Querol,E. (1997) *CABIOS*, **13**, 231–234.
Bairoch,A. and Boeckmann,B. (1992) *Nucleic Acids Res.*, **20**, 2019–2022.
Barlow,D.J. and Thornton,J.M. (1988) *J. Mol. Biol.*, **201**, 601–609.
Bernstein,F.C., Koetzle,T.F., Williams,G.J.B., Meyer,E.F., Brice,M.D., Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **112**, 535–542.
Casadio,R., Fariselli,P., Taroni,C. and Compiani,M. (1996) *Eur. Biophys J.*, **24**, 165–178.
Cserzo,M., Bernassau,J.-M., Simon,I. and Maigret,B. (1994) *J. Mol. Biol.*, **243**, 388–396.
Cserzo,M., Wallin,E., Simon,I., von Heijne,G. and Elofsson,A. (1997) *Protein Engng*, **10**, 673–676.
Deber,C.M., Glibowicka,M. and Woolley,G.A. (1990) *Biopolymers*, **29**, 149–157.
Deisenhofer,J. and Michel,H. (1989) *Science*, **245**, 1463–1473.
Deisenhofer,J., Epp,O., Sinning,I. and Michel,H. (1985) *J. Mol. Biol.*, **246**, 429–457.
Esposti,D.M., Crimi,M. and Venturoli,G. (1990) *Eur. J. Biochem.*, **190**, 207–219.
Feher,G., Allen,J.P., Okamura,M.Y. and Rees,D.C. (1989) *Nature*, **339**, 111–116.
Gromiha,M.M. and Ponnuswamy,P.K. (1995) *Int. J. Pept. Protein Res.*, **45**, 225–240.
Gromiha,M.M., Majumdar,R. and Ponnuswamy,P.K. (1997) *Protein Engng*, **10**, 497–500.
Henderson,R., Baldwin,J.M., Ceska,T.A, Zemlin,F., Beckmann,E. and Downing,K.H. (1990) *J. Mol. Biol.*, **213**, 899–929.
Hirokawa,T., Boon-chieng,S. and Mitaku,S. (1998) *Bioinformatics*, **14**, 378–379.
Iwata,S., Ostermeier,C., Ludwig,B. and Michel,H. (1995) *Nature*, **376**, 660–669.
Jennings,M.L. (1989) *Annu. Rev. Biochem.*, **58**, 999–1027.
Lohmann,R., Schneider,G. and Wrede,P. (1996) *Biopolymers*, **38**, 13–29.
Nikiforovich,G.V. (1998) *Protein Engng*, **11**, 279–283.
Parodi,L.A., Granatir,C.A. and Maggiora,G.M. (1994) *CABIOS*, **10**, 527–535.
Persson,B. and Argos,P. (1994) *J. Mol. Biol.*, **237**, 182–192.
Persson,B. and Argos,P. (1997) *J. Protein Chem.*, **16**, 453–457.
Ponnuswamy,P.K. and Gromiha,M.M. (1993) *Int. J. Pept. Protein Res.*, **42**, 326–341.
Rost,B., Casadio,R., Fariselli,P. and Sander,C. (1995) *Protein Sci.*, **4**, 521–533.
Rost,B., Fariselli,P. and Casadio,R. (1996) *Protein Sci.*, **5**, 1704–1718.
Sciffer,M., Chang,C.-H. and Stewans,F.J. (1992) *Protein Engng*, **5**, 213–214.
Traxler,B., Boyd,D. and Beckwith,J. (1993) *J. Memb. Biol.*, **132**, 1–11.
von Heijne,G. (1988) *Biochim. Biophys Acta*, **947**, 307–333.
von Heijne,G. (1992) *J. Mol. Biol.*, **225**, 487–494.
Weiss,M.S., Kreush,A., Shiltz,E., Nestel,U., Welte,W., Weckesser,J. and Schulz,G.E. (1991) *FEBS Lett.*, **280**, 379–382.
Yanagihara,N., Suwa,M. and Mitaku,S. (1989) *Biophys Chem.*, **34**, 69–77.

Received August 14, 1998; revised February 4, 1999; accepted March 18, 1999