Will the seven-helix bundle be a common structure for integral membrane proteins?

J.K. Mohana Rao, Paul A. Hargrave+ and Patrick Argos*

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907 and *+Department of Medical Biochemistry (School of Medicine) and Department of Chemistry and Biochemistry (College of Science), Southern Illinois University, Carbondale, IL 62901, USA*

Received 14 April 1983

A prediction algorithm, designed to detect lipid-embedded helical regions in membrane proteins, was applied to the amino acid sequence of a chloroplast thylakoid membrane protein important in photosynthesis. It is suggested that the thylakoid membrane protein consists of 7 transmembrane helices connected by exposed turn segments, similar to current models for bacteriorhodopsin and bovine rhodopsin. This basic structural feature, the 7-helical bundle, may prove to be shared by many integral membrane proteins.

Membrane protein *Thylakoid membrane* *Rhodopsin* *Protein structure* *Bacteriorhodopsin* *Helix*

1. INTRODUCTION

There are few integral membrane proteins for which both the amino acid sequence and details of membrane topography are known. One of the best examples is the purple membrane protein of *Halobacterium halobium*, bacteriorhodopsin (BR). BR consists of a bundle of 7 transmembrane helices and connecting segments as shown by electron diffraction and many additional methods [1]. The vertebrate photoreceptor protein rhodopsin (RHO) is almost certainly another 7-helical protein. Studies by chemical modification, limited proteolysis and secondary structure prediction are all in excellent agreement ([2,3], reviewed in [4]).

The sequence of a protein from spinach chloroplast thylakoid membranes has been reported [5]. The thylakoid membrane protein (TMP) which is a component of photosystem II, regulates electron transport in the chloroplast. It is characterized by rapid synthesis and turnover and the ability to bind the herbicides diuron and atrazine [6,7]. Here, we present evidence which suggests that TMP, like BR and RHO, is a membrane protein containing 7 helices.

2. MATERIALS AND METHODS

We have employed a prediction algorithm designed to detect and delineate hydrophobic helical spans in protein primary sequences [8]. Five physical parameters (hydration potential, membrane-buried transfer free energy, polarity, bulk, and turn conformational preference) are used to produce a smoothed curve as a function of residue sequence number. All primary structural regions with combined values ≥ 0 are predicted as membrane-buried helical regions, while those with values ≤ 0 are considered to be membrane surface-exposed turn regions.

A second technique was used to indicate the number of possible helices in TMP as well as the N-terminus of each helix. Curve segments from the helical spans for the RHO 5-parameter plot were compared to all segments of the comparable TMP

* To whom reprint requests should be addressed

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

plot, and correlation coefficients calculated at each lag value (i.e., the position of the RHO span compared to the N-terminal residue of TMP). The number of peaks and their associated lag values should correspond respectively to the number of helices and their N-terminal amino acid in TMP.

Helical wheels were calculated using a table of membrane-buried helical preference values in which residues preferring a lipid environment have values >1.0 [8]. Helical sidedness was chosen such that $(N_1 - N_2)$ is maximal, where N_1 = the number of residues on one helical side that do not prefer lipid contact, and N_2 = the corresponding number for the other helical side.

3. RESULTS AND DISCUSSION

Application of the prediction algorithm to the primary sequence of the thylakoid membrane protein shows 7 membrane-spanning helical regions (fig.1, table 1). When these 7 general regions of the protein sequence are inspected visually, it is easy to identify spans of 20–28 amino acids which contain predominantly hydrophobic residues and are flanked on either side by hydrophilic residues. To further confirm and refine the helix assignment,

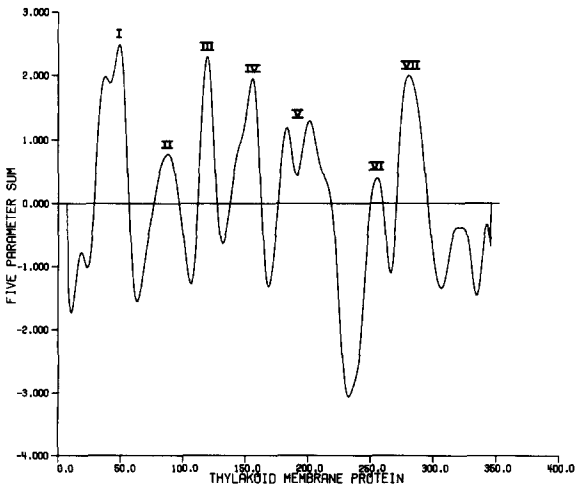

Fig.1. Plot of the amino acid sequence number for thylakoid membrane protein *vs* a weighted 5-parameter characteristic value for a given amino acid [8]. The peaks corresponding to helical regions I–VII are designated. The parametric value is shown as zero at each of the seven terminal residues due to end-effects in the smoothing procedures [8].

correlation *vs* lag curves were determined in order to identify N-terminal residues of the thylakoid protein helices. These procedures led to final selection of each of the helical spans (table 1). These seven helices form the basis for the topological model of the protein presented in fig.2.

The thylakoid protein shares many structural characteristics in common with BR and RHO. The average helical length is about 24 for each protein. Correlations of the amino acid compositions in the predicted helical regions range between 0.8 and 0.9 for the intercomparisons. The percentage of strongly polar and charged residues in the helical regions of the structural models of BR [9], RHO [3] and TMP are, respectively, 21%, 12% and 13% (we are considering Asp, Glu, Lys, Arg, Thr, Ser, Gln and Asn as strongly polar and charged residues). Percentages of such residues in the turn regions are, respectively, 45%, 50% and 50%. The percentages of hydrophobic residues in the helical regions of BR, RHO, and TMP are, respectively, 68%, 76% and 68% (we are considering Ala, Phe, Ile, Leu, Met, Val, Tyr and Trp as hydrophobic). Only two charged residues (one Glu and one Arg) are predicted to be buried (fig.2) and they might conveniently form an ion pair. For each protein the N- and C-terminal non-helical regions have the longest residue length.

Helical wheels were constructed for the 7 predicted helices of TMP (fig.3) in an attempt to discern which surface would face the protein interior and which surface would face the lipid bilayer. Helices of TMP compare quite favorably with those of BR and RHO (table 2). This observation (that hydrophobic membrane-spanning helices have opposing surfaces which vary considerably in their polarity) may prove to be a general feature associated with helix-packing requirements for multihelix integral membrane proteins.

Helices in known soluble proteins rarely contain glycine or proline in their middle and C-terminal portions [10]. The predicted TMP helices contain 16 glycines and 9 prolines, whereas those of BR contain 13 glycines and 5 prolines, and helices of RHO 8 glycines and 7 prolines. The electron diffraction structure of BR shows that some of the helices are bent or kinked [1], presumably as a result of their proline and glycine content. Occurrence of these residues in membrane-buried helices would appear to be more common than in helices

Table 1

Assignment of predicted helical spans in thylakoid membrane protein

Helix designation	Segment from 5-parameter curve	N-Terminal residue from correlation curve	Assigned helical span	Helical length (residues)
I	29– 55	30	30– 55	26
II	77– 97	76	77– 97	21
III	113–127	110	105–127	23
IV	139–163	141	141–163	23
V	177–219	192	192–219	28
VI	251–260	243	246–265	20
VII	273–295	271	273–295	23

The segments are given as amino acid number spans in the protein primary structure. The 5-parameter curve is shown in fig.1

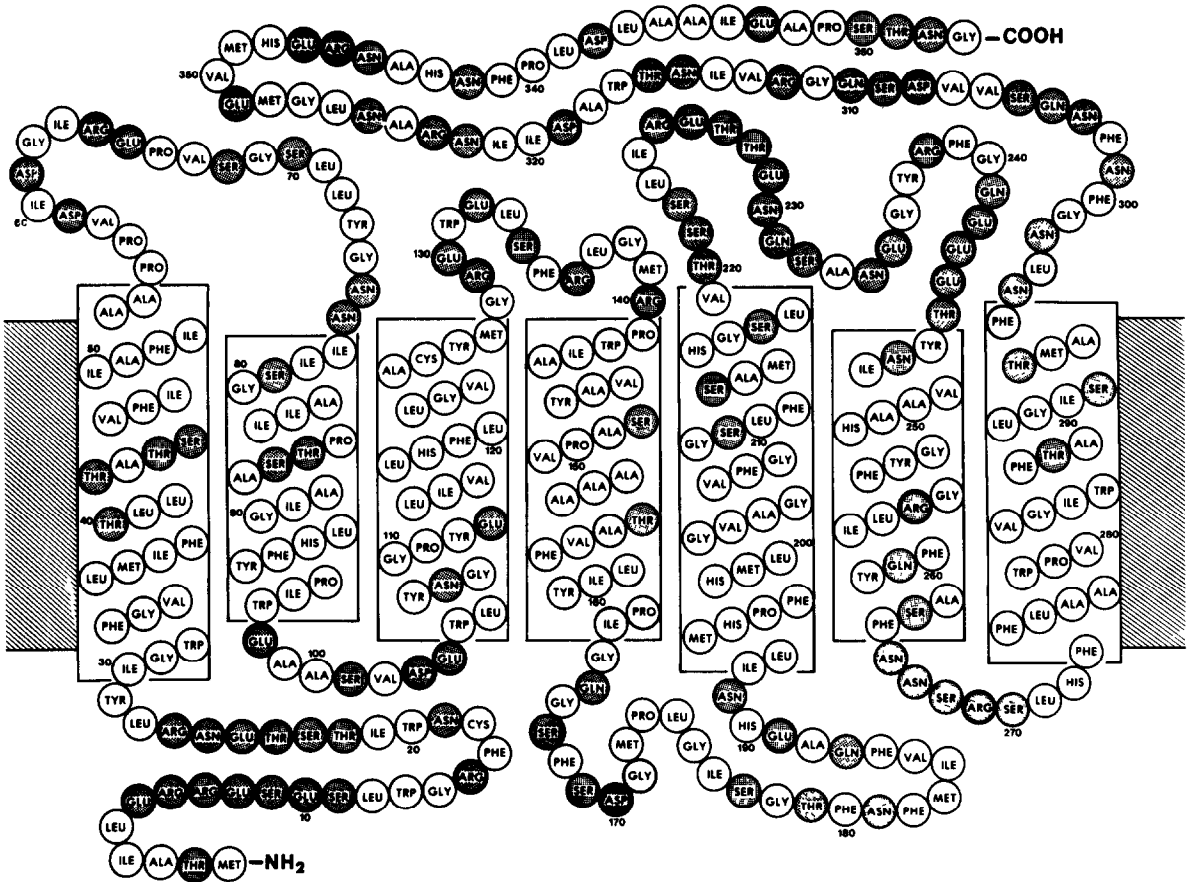

Fig.2. A depiction of the suggested helical and turn regions in thylakoid protein. The residues within the large rectangles are predicted as helices buried within the lipid bilayer while the remaining amino acids constitute the exposed turn segments. The shaded residues correspond to the charged (in bold circles) or strongly polar amino acids.

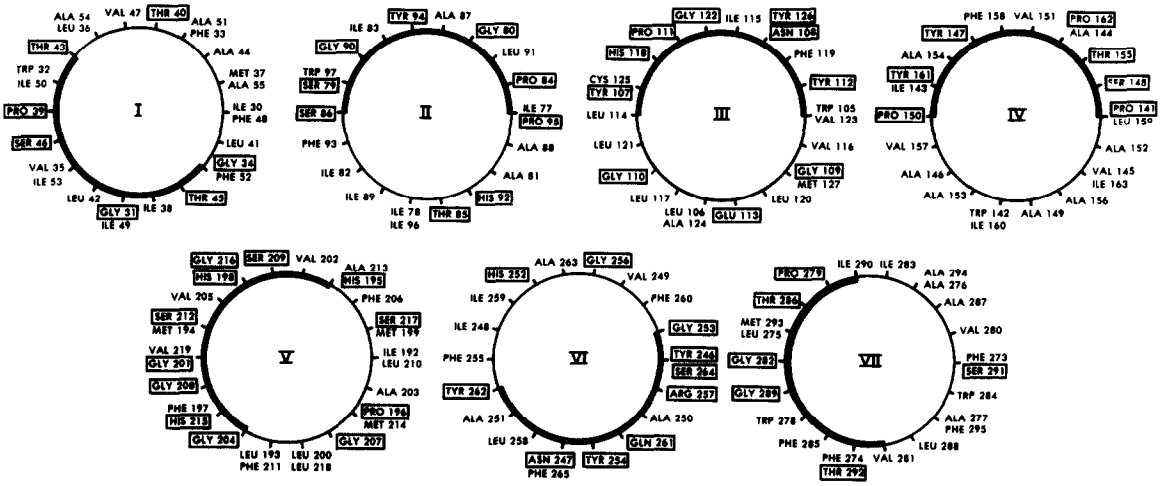

Fig.3. Helical wheels for the 7 predicted helical segments in thylakoid membrane protein. Amino acids with membrane-buried propensities <1.0 (i.e., not preferring lipid contact) are boxed. The side of the helices suggested to face the protein interior is indicated by thick lines.

in soluble proteins. It is also noteworthy that threonine and serine are the most frequently occurring polar residues in the predicted helical spans. There are 20 Ser + Thr in BR, 12 in RHO and 16 in TMP. Hydrogen bonding between their side chain oxygens and peptide bond atoms should provide helical stability [10] and eliminate the need for interaction with other amino acids in the protein interior.

Table 2

Comparison of the level of helical sidedness possible in thylakoid membrane protein, bovine rhodopsin and bacteriorhodopsin

Helix	Thylakoid protein		Bovine rhodopsin		Bacterio-rhodopsin	
	(N_1)	(N_2)	(N_1)	(N_2)	(N_1)	(N_2)
I	6	1	5	2	9	1
II	7	2	8	1	7	3
III	7	3	6	2	7	1
IV	8	0	4	0	7	3
V	9	2	6	1	5	2
VI	8	2	4	0	7	2
VII	5	1	7	2	6	3
Mean	7.1	1.6	5.7	1.1	6.9	2.1
Mean (N_1-N_2)	5.5		4.6		4.8	

N_1 refers to the number of residues on one side of a helical wheel that do not prefer lipid contact while N_2 refers to the number of such residues on the opposite side of the wheel [8]. Sides N_1 and N_2 were chosen such that their difference is maximal

Little experimental information is available which may be used to aid construction of a membrane topographic model for TMP. It has been reported that this M_r ~32000 protein is digested by trypsin to yield a M_r 18000 fragment which is further digested to M_r 16000 [6,7]. In [5] the TMP sequence was analyzed and it was concluded that only cleavage at Arg^{64} and Arg^{238} followed by cleavage at Arg^{225} could yield fragments of appropriate size. Such cleavages would be consistent with our proposed model since all of the postulated digestion sites are located on the same (carboxyl-terminal) side of the membrane. Proteolysis may yield fragments from the remainder of the molecule which are too small or too heterogeneous to be easily detected by the gel techniques employed [6,7].

Although the information concerning the primary and secondary structures of integral membrane proteins is currently rather scanty, the transmembrane helix will undoubtedly emerge as the basic structural element. The dominant structural feature of both BR and RHO is a 7-helical bundle, and we now suggest that TMP possesses a similar structure. We have compared the primary structures of all 3 proteins and find no statistically significant relationship. The 3 proteins are diverse

in origin, coming from plant, bacterial and mammalian sources. Although they are all involved in light-sensitive processes, and both BR and RHO are retinyl-proteins, there is no reason to believe that they are closely evolutionarily related. Rather, it seems more reasonable to us that the packing of 7 helices together in integral membrane proteins may represent a uniquely stable arrangement which has been achieved by processes of convergent evolution and that we may expect to see more examples of such proteins in the future.

ACKNOWLEDGEMENTS

P.A. acknowledges support from the National Institutes of Health (grant no. GM27682) and the American Cancer Society (Faculty Research Award no. FRA173). P.A.H. acknowledges support from the National Institutes of Health (grant nos. EY1275 and EY2875).

REFERENCES

[1] Henderson, R. (1977) Annu. Rev. Biophys. Bioeng. 6, 87–109.
[2] Ovchinnikov, Yu.A. (1982) FEBS Lett. 148, 179–191.
[3] Hargrave, P.A., McDowell, J.H., Curtis, D.R., Wang, J., Juszczak, E., Fong, S.-L., Mohana Rao, J.K. and Argos, P. (1983) Biophys. Struct. Mech., in press.
[4] Hargrave, P.A. (1982) Prog. Retinal Res. 1, 1–51.
[5] Zurawski, G., Bohnert, H.J., Whitfeld, P.R. and Bottomley, W. (1982) Proc. Natl. Acad. Sci. USA 79, 7699–7703.
[6] Mattoo, A.K., Pick, U., Hoffman-Falk, H. and Edelman, M. (1981) Proc. Natl. Acad. Sci. USA 78, 1572–1576.
[7] Steinback, K.E., McIntosh, L., Bogorad, L. and Arntzen, C.J. (1981) Proc. Natl. Acad. Sci. USA 78, 7463–7467.
[8] Argos, P., Mohana Rao, J.K. and Hargrave, P.A. (1982) Eur. J. Biochem. 128, 565–575.
[9] Engelman, D.M., Henderson, R., McLachlan, A.D. and Wallace, B.A. (1980) Proc. Natl. Acad. Sci. USA 77, 2023–2027.
[10] Argos, P. and Palau, J. (1982) Int. J. Prot. Pept. Res. 19, 380–393.