Biophysical Chemistry, 34 (1989) 69–77
Elsevier

BIOCHE 01374

A theoretical method for distinguishing between soluble and membrane proteins

Norikuni Yanagihara, Makiko Suwa and Shigeki Mitaku

Department of Material Systems Engineering, Faculty of Technology, Tokyo University of Agriculture and Technology, Nakamachi 2-24-16, Koganei-shi, Tokyo 184, Japan

Received 17 January 1989
Revised manuscript received 11 May 1989
Accepted 15 May 1989

Amino acid sequence; Hydrophobicity; Protein morphology; Membrane protein; Protein, soluble; Fourier transformation

A method for distinguishing between membrane and soluble proteins in an amino acid sequence was developed, using only two parameters associated with the hydrophobicity: the average hydrophobicity and the power spectral density of period longer than 30 residues. The power spectral density was calculated by a maximum entropy method of Fourier transformation. Membrane proteins could be distinguished from soluble proteins with a distinction rate as high as 97%. This fact strongly suggests that the morphology of proteins, i.e., membrane or soluble forms, is determined thermodynamically through the hydrophobicity of polypeptides.

1. Introduction

Higher order structures of proteins are formed due to the effects of various noncovalent forces: electrostatic interactions, van der Waals forces, hydrophobic interactions and hydrogen bonds [1]. In order to establish the mechanism of formation of higher order structures of proteins, the characteristics of and the role played by each type of noncovalent interaction require clarification. Unlike polar interactions, which vary with distance according to a dependence of r^{-n}, the free energy of hydrophobic interaction is proportional to the surface area of a molecule [2]. This kind of interaction occurs ubiquitously in aqueous solutions of organic substances and is considered to be the driving force for the formation of various aggregated structures [3]. For example, a number of amphiphiles show micellar, hexagonal and lamellar phases. The hydrocarbon chains of amphiphiles undergo aggregation together within the structure due to hydrophobic interaction, whereas hydrophilic groups stabilize the microstructure. Since proteins are also amphiphilic, one of the most significant effects contributed by hydrophobic interactions to the formation of higher order structures of proteins may be the determination of the gross morphology, i.e., whether they become globular or membranous.

In this respect, amino acid sequence analysis has revealed significant differences in hydrophobicity between membrane and soluble proteins, which are the major categories of natural proteins:

(1) The hydrophobicity of membrane proteins with several membrane-spanning helices is very high as a result of contact with the lipid bilayer [4–9], whereas soluble proteins possess a hydrophobic core surrounded by a hydrophilic surface [10–12].

Correspondence address: S. Mitaku, Department of Material Systems Engineering, Faculty of Technology, Tokyo University of Agriculture and Technology, Nakamachi 2-24-16, Koganei-shi, Tokyo 184, Japan.

(2) Highly hydrophobic segments of about 25 residues have been observed in the amino acid sequence of membrane proteins corresponding to the thickness of transmembrane helices [4–9,13], while the amphiphilic nature of the helices becomes significant at the surface of soluble proteins [14–16].

(3) The hydration of soluble proteins, to which hydrophobic interaction is intimately related, appears to be responsible for their denaturation [17].

The above facts suggest that the hydrophobicity of the amino acid sequence is an important factor in the formation of the gross structure of soluble and membrane proteins.

If the hydrophobic interaction is indeed the major factor determining protein morphology, then discrimination between membrane and soluble proteins should be possible simply through analysis of the hydrophobicity of the amino acid sequences. In this paper, we describe a method that facilitates discrimination between soluble and membrane proteins through the combination of two parameters calculated from the hydrophobicity: the average hydropathy index of the entire amino acid sequence and the power spectral density of periods longer than 30 residues. Analysis of more than 3000 amino acid sequences from the National Biomedical Research Foundation (NBRF) data base revealed that intrinsic membrane proteins with several membrane-spanning helices may be distinguished from soluble proteins with a reliability of 97%. The conditions under which a polypeptide becomes a natural protein are also discussed.

2. Methods

2.1. Average hydrophobicity

We employed plotting of the hydropathy index according to the method of Kyte and Doolittle [7] for the analysis. We also examined other indices [13,18,19], however, only slight differences between the indices were found. The values of the hydropathy index were as follows: Ile, 4.5; Val, 4.2; Leu, 3.8; Phe, 2.8; Cys, 2.5; Met, 1.9; Ala, 1.8; Gly, −0.4; Thr, −0.7; Ser, −0.8; Trp, −0.9; Tyr, −1.3; Pro, −1.6; His, −3.2; Glu, −3.5; Gln, −3.5; Asp, −3.5; Asn, −3.5; Lys, −3.9; Arg, −4.5. In order to obtain the average hydropathy index of a protein, the amino acid sequence was converted to a hydropathy index sequence, i.e., $X(1)$, $X(2), \ldots, X(N)$. The average hydropathy index $\langle X \rangle$ was then calculated from the expression:

$$\langle X \rangle = \left\{ \sum_{i=1}^{N} X(i) \right\} / N \tag{1}$$

2.2. Spectral analysis

The spectral characteristics of a stationary series of numbers may be estimated using autoregressive models. In this case, the hydropathy index sequence of a protein was analyzed by using a procedure based on maximum entropy, namely, the autoregression method first described by Akaike [20]. In an autoregression model, a series of hydropathy indices is represented by the following equation.

$$X(n) = \sum_{m=1}^{M} a_m X(n-m) + a_0 + \epsilon(n) \tag{2}$$

where $X(n)$ denotes the hydropathy index of the n-th amino acid, $\epsilon(n)$ white noise and M the order of the autoregression. In calculation of the power spectral density, $X(n)$ is first replaced by

$$X^*(n) = X(n) - \left\{ \sum_{n=1}^{N} X(n) \right\} / N, \tag{3}$$

N designating the total number of the series. Thus, it is not difficult to show that, provided the best coefficients $\{ a_m^{(M)}: m = 1, 2, \ldots, M \}$ and the least mean square of the residuals R_M can be established,

$$R_M = \left[\sum_{n=1}^{N} \left\{ X^*(N) - \sum_{m=1}^{M} a_m^{(M)} X^*(n-m) \right\}^2 \right] / N, \tag{4}$$

then the power spectral density $p(f)$ at frequency f for the autoregression model of order M can be

obtained as

$$p(f) = \frac{R_M}{\left|1 - \sum_{m=1}^{M} a_m^{(M)} \exp(-i2\pi f_m)\right|^2} \quad (5)$$

However, in using this procedure, the necessity for determining the order M of the autoregressive model represents a difficult task. In order to do so, the final prediction error (FPE) is used in the maximum entropy method of Akaike [20],

$$(\mathrm{FPE}) = \frac{N+M+1}{N}\mathrm{MIN}(R_M), \quad (6)$$

in which $\mathrm{MIN}(R_M)$ denotes the minimum of the mean square of residuals for the order M. The autoregression model in which the FPE is at a minimum is assumed to be the best form. In many cases of spectral analysis of the amino acid sequence, however, we were unable to determine the minimum of FPE. In such cases, the order of M was assumed to be 30, since the results obtained via calculation were sufficiently stable and the resolution of the periods was good enough on this assumption, as shown in section 2.3.

2.3. *Test calculation of spectral analysis*

It is known that, in spite of the high resolution of the periodicity, the maximum entropy method entails the disadvantage of a low degree of reliability for the magnitude of the power spectral density, particularly when the stationary sequence of numbers is short [21]. The hydropathy index sequences in our calculation appear to be of sufficient length, since the degree of polymerization of proteins lies mostly between 100 and 500. However, we have evaluated the width for the range of variation in values of the calculated spectral density by using model data. Fig. 1 shows an example of such a calculation. The model data comprised the sum of six sine functions with periods of 2.8, 3.6, 5, 10, 20 and 40 residues of the same intensity as shown in fig. 1a. Fig. 1b depicts an example of a calculated spectrum for a length of 200. The periods of the calculated spectral peaks were in

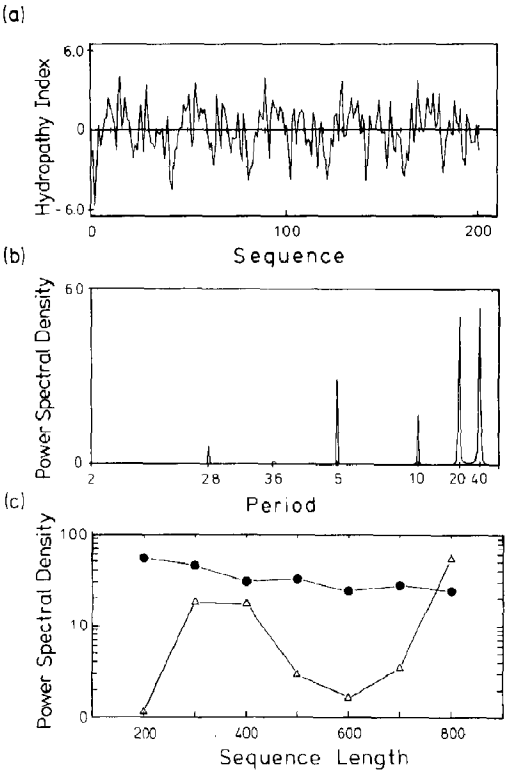

Fig. 1. Model calculation of spectral analysis via the maximum entropy method. Model data comprising superposition of six sine functions of the same amplitude (a). Periods in the original data reproduced by spectral analysis: 2.8, 3.6, 5, 10, 20 and 40 residues (b). Power spectral densities for periods of 40 (●) and 3.6 (△) residues as a function length of sequences (c).

good agreement with the six periods in the model sequence. However, the peak values were scattered over a considerable range, particularly at lower frequencies. For example, the power spectral density at 3.6 residues is much smaller than that at 40 residues. The factor of scatter in the values of the intensity is one of the disadvantages entailed by the use of the maximum entropy method. The intensity of a calculated peak is less reliable compared to that obtained using the fast Fourier transform (FFT) method, although the resolution of the period is remarkably high, particularly for a noisy sequence such as that of the amino acids [22]. We then carried out a systematic calculation

of the model data, varying the length of the sequence between 200 and 800 residues. The power spectral densities at 3.6 and 40 residues are compared in fig. 1c. In all calculations the period of the spectral peak agreed well with the sine function for the model data and the magnitude of the power spectral density remained stable over the range of periods beyond 20 residues, although the power spectral density exhibited considerable variation in values for shorter periods. As we have only used a power spectral density of more than 30 residues in length for studies on discriminating between the proteins, the results of calculation are assumed be sufficiently reliable.

2.4. Categories of proteins

Data on proteins in the NBRF data base were used in the present work for calculations. Since fragments with short sequences of amino acids are unsuitable for periodicity analysis, 3609 proteins other than the fragments were studied. The proteins were classified into three categories: multi-helical membrane proteins; single-helical membrane proteins; soluble proteins including peripheral membrane proteins. The first category includes typical membrane proteins with several transmembrane helices. Of the 3609 proteins in the database (version 4.0), 120 were assigned as being multi-helical membrane proteins. The second group comprises those proteins with only one transmembrane helix which anchors a large soluble functional segment; 95 proteins fell within this classification. The third category contained 2563 soluble proteins which are stable in aqueous solution. The total number of proteins in these three categories was 2778. The remaining 831 proteins were of miscellaneous classification: membrane proteins with a high β-sheet content, hypothetical forms inferred from the DNA sequence and those which we could not assign to the above three categories.

3. Results

The hydrocarbon region of lipid bilayers is rather nonpolar, having a dielectric constant of

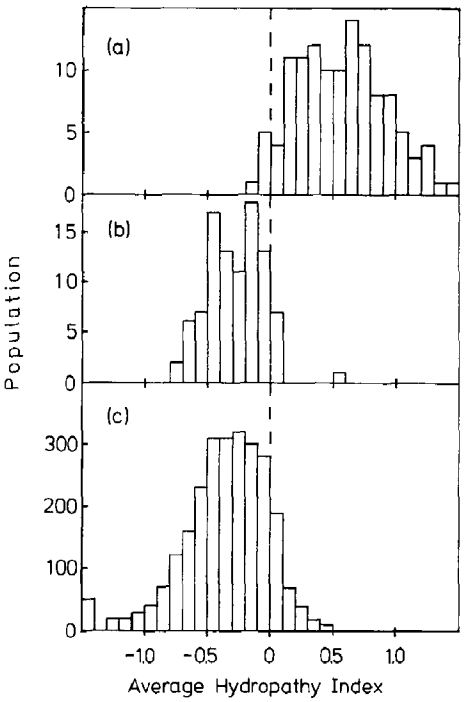

Fig. 2. Population of multi-helical membrane (a), single-helical membrane (b) and soluble (c) proteins plotted as a function of average hydropathy index for the whole amino acid sequence.

2–4 [22], which results in the highly hydrophobic nature of membrane proteins. On the other hand, soluble proteins should be less hydrophobic because of their polar aqueous environment. Fig. 2 shows histograms of three types of proteins, i.e., multi-helical and single-helical membrane proteins, and soluble proteins, as a function of the average hydropathy index. The following points should be noted from the histograms: (1) multi-helical membrane proteins are more hydrophobic than soluble proteins as well as single-helical membrane proteins, however, a definitive distinction between membrane and soluble proteins is difficult to make, due to the histograms overlapping with each other. (2) The histogram for single-helical membrane proteins is almost the same as that of the soluble proteins.

The greater hydrophobicity of multi-helical membrane proteins as compared to soluble proteins indicates that the average hydropathy index may be used as one of the parameters for dis-

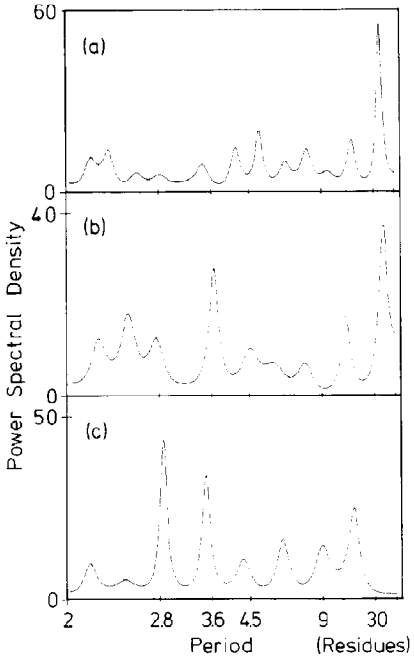

Fig. 3. Power spectral density calculated via the maximum entropy method plotted as a function of period for bovine rhodopsin (a), human glycophorin (b) and chicken cytochrome *c* (c).

criminating between proteins. The other relates to the periodicity in membrane proteins. The sharp contrast between neighboring segments, the hydrophobic portion being situated in the membrane and the hydrophilic part exposed to water, results in a strong periodicity of hydrophobicity. Since the length of transmembrane helices is determined by the thickness of the hydrocarbon region of the lipid bilayer, i.e., about 3 nm, a periodicity of greater than 30 residues should be observed in the amino acid sequence.

In order to study the periodicity, we have calculated the power spectral density of the hydropathy index sequence obtained from the amino acid sequence. Fig. 3 shows examples of the power spectral density of three kinds of proteins: bovine rhodopsin, human glycophorin and chicken cytochrome *c*. Rhodopsin, a multi-helical membrane protein, showed a significant peak at a period of 42 residues as expected. However, the power spectrum of cytochrome *c*, a soluble protein, showed no strong peak for periodicities in the range above 30 residues. Although single-helical membrane proteins are similar to soluble proteins in the histogram of fig. 2, the power spectral density of glycophorin displayed a significant peak at a period of about 50 residues which is characteristic of membrane proteins. Therefore, figs. 2 and 3 suggest that multi-helical membrane proteins are characterized by high values of the average hydropathy index, $\langle X \rangle$, and a power spectral density greater than 30 residues, P. Low $\langle X \rangle$ and high P values are typically found for single-helical

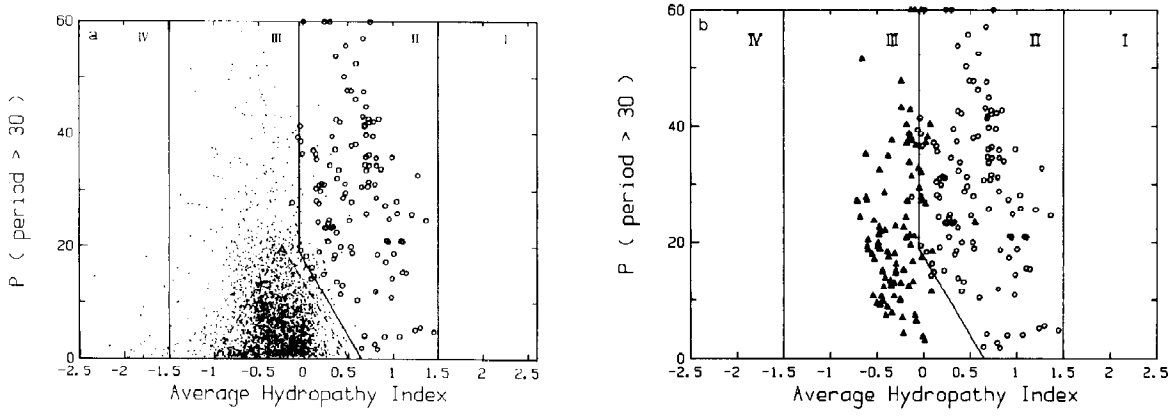

Fig. 4. All multi-helical membrane proteins (○) and soluble proteins (dots) plotted in $\langle X \rangle$ vs. P space (a). Corresponding plot for multi-helical (○) and single-helical (▲) membrane proteins (b). Boundaries distinguishing the various kinds of protein. The triangular area delineated by the broken lines contains 75% of the soluble proteins. See text for details of the boundaries between areas.

membrane proteins, whereas soluble proteins exhibit low values for both parameters.

The above calculation has been performed for the sequences of 3609 proteins from the NBRF database. Fig. 4a and b shows plots for proteins of the multi-helical membrane, soluble and single-helical membrane types in $\langle X \rangle$ vs. P space. Because the number of soluble proteins is large, namely 2563, the symbol corresponding to this group of proteins is a dot, multi-helical and single-helical membrane proteins being designated by empty circles and filled triangles, respectively. According to the distribution of the three types of protein, we have divided the $\langle X \rangle$ vs. P space into four areas, I–IV. The boundaries between the four areas are as follows:

$$\text{(I/II)}\ \langle X \rangle = 1.5,$$

$$\text{(II/III)}\ \langle X \rangle = -0.05 \quad \text{and}$$

$$P = -27.1\langle X \rangle + 17.6, \qquad (7)$$

$$\text{(III/IV)}\ \langle X \rangle = -1.5.$$

It is apparent from fig. 4a that the distribution of multi-helical membrane proteins differs from that of soluble proteins to the extent of being almost exclusive of the latter. The boundary between areas II and III consists of two straight lines. For the case of periodicities beyond 30 residues being sufficiently large, a protein with even a negative average hydropathy index, $\langle X \rangle \geqq -0.05$, may become one of the multi-helical membrane type. However, a higher average hydrophobicity is required for multi-helical membrane proteins with lower periodicity, viz., $P \geqq -27.1\langle X \rangle + 17.6$. In any event, the transmembrane helices of such proteins must be sufficiently hydrophobic with respect to length as well as magnitude.

Fig. 4b depicts the distribution of single-helical and multi-helical membrane proteins in $\langle X \rangle$ vs. P space. It is remarkable that most single-helical membrane proteins are located within area III and are clearly distinguishable from their multi-helical counterparts, single-helical membrane proteins namely occurring in the same area as soluble proteins. The low hydrophobicity of single-helical membrane proteins seems reasonable, the greater

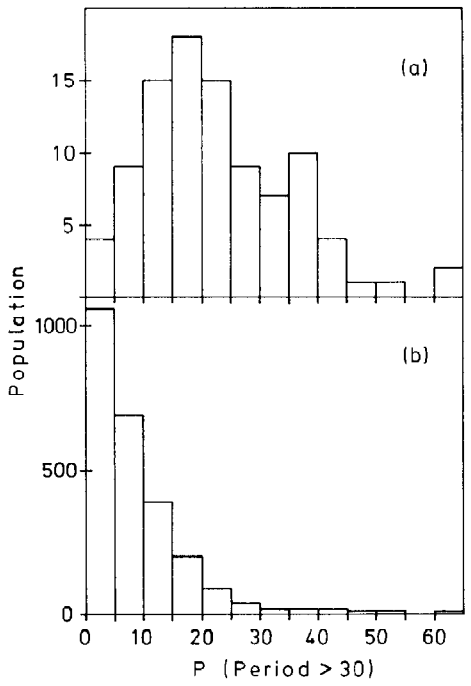

Fig. 5. Population of single-helical membrane (a) and soluble (b) proteins plotted as a function of power spectral density for periodicities longer than 30 residues.

proportion of these proteins protrudes from the membrane into the aqueous phase. The difference between single-helical membrane proteins and the soluble type lies in the existence of a transmembrane helix anchoring large soluble segments of single-helical membrane proteins. This structure should result in periodicity occurring in the hydropathy index for single-helical membrane proteins. Fig. 5 shows histograms for both of these types of protein as a function of the power spectral density beyond 30 residues. The population of soluble proteins decreases monotonically from a power spectral density of zero. However, the population of single-helical membrane proteins shows a peak around a power spectral density of 20, which is apparently different from that of soluble proteins. Although completely distinguishing between the two types is a difficult task due to overlap in the region of short periodicity, it would appear reasonable to assume that the gross struc-

Table 1

Proportion of proteins in four areas of average hydropathy index vs. power spectral density space

Type of protein	Proportion (%)				Total number of proteins
	I	II	III	IV	
Soluble proteins	0	1.8	96.6	1.6	2563
Multi-helical membrane proteins	0	97.5	2.5	0	120
Single-helical membrane proteins	0	11.6	88.4	0	95
Miscellaneous	0.4	15.6	83.8	0.2	831
Total	0.1	8.4	90.3	1.2	3609

ture of the proteins is also determined by their hydrophobicity.

About 75% of soluble proteins are clustered into a small triangular region around the origin in fig. 4a which is delineated by the broken lines: $P = 27.1(\langle X \rangle + 1.0)$ and $P = -27.1(\langle X \rangle - 0.5)$. The densely populated area of soluble proteins is not necessarily triangular, since there is no theoretical basis for such a distribution. The lines were drawn for convenience's sake. However, it seems certain that most soluble proteins are formed by a subtle balance of the hydrophobic and hydrophilic residues within their amino acid sequences.

The boundaries of eq. 7 were also derived from inspection of the entire set of data points in fig. 4a and b, rather than from theoretical considerations. The boundaries appear to be justified based on the good separation among the three kinds of protein. The populations of all kinds of proteins, i.e., multi- and single-helical membrane, soluble, and miscellaneous proteins, in areas I–IV are listed in table 1. Out of 120 multi-helical membrane proteins 117 were allocated to area II, constituting 97.5% of this type of protein. 2477 soluble proteins belong to area III which accounts for 96.6% of the total for this group. Thus, typical intrinsic membrane proteins may be discriminated almost completely from soluble proteins by hydrophobicity analysis alone.

Peripheral membrane proteins were allocated to the category of soluble proteins in this work, since they are solubilized by mild treatments such as increasing the ionic strength. This indicates that the mechanism of binding to membranes is quite different between intrinsic and peripheral membrane proteins, the latter being essentially of the soluble type. In fact, peripheral membrane proteins selected from the NBRF database [8] had the same distribution in $\langle X \rangle$ vs. P space as soluble proteins. Most were located within the triangular region in fig. 4 (data not shown).

Analysis of the miscellaneous proteins listed in table 1 also provides useful information. Membrane proteins with a high β-sheet content had low hydrophobicity and were located around the $\langle X \rangle$-axis in region III, despite being embedded in the membrane. The mechanism of stabilization of these proteins in membranes has not yet been clarified. However, it is presumably related to the fact that most of the membrane proteins with a high β-sheet content are the outer membrane proteins of bacteria. Since they must traverse the aqueous medium to reach the outer membrane after synthesis in the cytoplasm, their hydrophilic nature seems reasonable. Data for many hypothetical proteins have been plotted in area II, suggesting that they are multi-helical membrane proteins.

Another important point is that substantially no natural proteins were found in areas I and IV. About 20 soluble proteins were ascertained to occur in area IV, representing only 1.2% of the total for this type. Most were nucleic-acid-binding proteins having a large number of basic residues. They are considered to form intimate complexes with acidic nucleic acids, neutralizing the acidic groups in nucleic acids through the basic side chains. With this specific exception, no natural proteins were found in area IV. This fact, together with the absence of natural proteins in area I, suggests that a polypeptide that is excessively hydrophobic or hydrophilic cannot form a stable three-dimensional structure.

4. Discussion

The present calculation has revealed the possibility of hydrophobic interactions participating in

the formation of protein structure, as summarized in the following.

(1) Two parameters calculated from the hydrophobicity sequence permitted almost complete distinction to be made between membrane and soluble proteins.

(2) The $\langle X \rangle$-P space could be divided into four areas. Multi-helical membrane proteins were allocated to area II and those of the soluble type to area III.

(3) Areas I and IV, in which $\langle X \rangle > 1.5$ and $\langle X \rangle < -1.5$, respectively, contained few natural proteins.

The greater than 2500 soluble proteins analyzed in this work include all families of proteins with various values of α-helix and β-sheet contents. Despite this diversity in protein structure, all soluble proteins were found to occupy area III. Furthermore, 75% thereof was clustered into a very small region within area III. However, multi-helical membrane proteins were assigned to area II, despite having a helical content similar to that of the globin-type soluble proteins. The fact that multi-helical proteins could be distinguished almost completely from soluble proteins clearly demonstrates the role played by hydrophobic interactions in determining the morphology of proteins, independently of their fine structure.

Klein et al. [8] have shown that intrinsic membrane proteins may be distinguished from the peripheral type via analysis of the hydrophobicity. They employed long hydrophobic helical segments in intrinsic membrane proteins and succeeded in distinguishing between both types of membrane protein, i.e., intrinsic and peripheral, with a reliability of 99%. Although they investigated only membrane proteins, their work appears to be in accord with the present data in that hydrophobicity is essential for differentiating between transmembrane proteins and other types.

The determination of protein morphology by data on hydrophobic interactions seems to be analogous to the degree of polymorphism of various amphiphilic molecules. As discussed in detail by Israelachvili [3], the morphology of aggregate structures of amphiphiles, i.e., micellar, hexagonal and lamellar phases, is governed by several factors. One of the most crucial of these is the ratio of the volume of the hydrophobic moiety to that of the hydrophilic group. When the hydrophobic group of an amphiphilic molecule is larger than the hydrophilic group, the molecule prefers the lamellar phase, while small micellar aggregates are observed for amphiphiles with a smaller hydrophobic group. It appears that the variation between proteins, which also belong to a class of amphiphilic molecules [2], corresponds to the mesophases of amphiphiles. With respect to volume, multi-helical membrane proteins have large hydrophobic segments compared to the hydrophilic portion. This is similar to the case of formation of lamellae by detergents possessing large hydrophobic groups. In contrast, soluble proteins have smaller hydrophobic segments, corresponding to micelle formation for detergents with single hydrocarbon chains. This analogy strongly suggests that the divergence of natural polypeptides into membrane and soluble proteins is the result of hydrophobic interactions.

The index range of Kyte and Doolittle is between -4.5 and 4.5. As shown in fig. 5, however, the region of natural proteins is rather small. The average hydropathy index of natural proteins ranges between only -1.5 and 1.5 (most soluble proteins have negative values of the hydropathy index while those of membrane proteins are positive). Polypeptides in area I, in which the average hydropathy index is above 1.5, are too hydrophobic to form stable membrane proteins. That signifies that large polar segments are required for forming a membrane protein, otherwise the structure of the lipid bilayer itself would be altered or stacking of membranes would be caused. Area IV comprises polypeptides with an average hydropathy index lower than -1.5. Polypeptides of this type presumably assume random coil structures, at least in part, thus impeding functional processes. It is remarkable that more than 75% of the soluble proteins are located within the small triangular region in fig. 4. In the future designing of a synthesis for pure enzyme, it is recommended that the information on hydrophobicity shown in fig. 4 be taken into account.

Acknowledgement

This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan (no. 61460248).

References

1 G.E. Schulz and H. Schirmer, Principle of protein structure (Springer, New York, 1979).
2 C. Tanford, The hydrophobic effect (Wiley, New York, 1980).
3 J.N. Israelachvili, Intermolecular and surface forces (Academic Press, London, 1985).
4 R.A. Capaldi and G. Vanderkooi, Proc. Natl. Acad. Sci. U.S.A. 69 (1972) 930.
5 D.M. Engelman, A. Goldman and T.A. Steitz, Methods Enzymol. 88 (1982) 81.
6 Yu.A. Ovchinnikov, FEBS Lett. 148 (1982) 179.
7 J. Kyte and R.F. Doolittle, J. Mol. Biol. 157 (1982) 105.
8 P. Klein, M. Kanehisa and C. DeLisi, Biochim. Biophys. Acta 815 (1985) 468.
9 S. Mitaku, S. Hoshi, T. Abe and R. Kataoka, J. Phys. Soc. Jap. 53 (1984) 4083.
10 J.P. Segrest and R.J. Feldmann, J. Mol. Biol. 87 (1974) 853.
11 W. Kauzmann, Adv. Protein Chem. 14 (1959) 1.
12 M.F. Perutz, J.C. Kendrew and H.C. Watson, J. Mol. Biol. 13 (1965) 669.
13 D. Eisenberg, R.M. Weiss, T.C. Terwilliger and W. Wilcox, Faraday Symp. Chem. Soc. 17 (1982) 109.
14 M. Schiffer and A.B. Edmundson, Biophys. J. 7 (1967) 121.
15 D. Eisenberg, R.M. Weiss and T.C. Terwilliger, Nature 299 (1982) 371.
16 S. Mitaku, S. Hoshi and R. Kataoka, J. Phys. Soc. Jap. 54 (1985) 2047.
17 T. Ooi and M. Oobatake, J. Biochem. 103 (1988) 114.
18 P. Argos, J.K.M. Rao and P.A. Hargrave, Eur. J. Biochem. 128 (1982) 565.
19 R.M. Sweet and D. Eisenberg, J. Mol Biol. (1983).
20 M. Akaike, Ann. Inst. Stat. Math. 21 (1969) 243.
21 M. Hino, Spectral analysis (Asakura, Tokyo, 1977) (in Japanese).
22 K. Kano, H. Gota and T. Ogawa, Chem. Lett. 1981 (1981) 655.