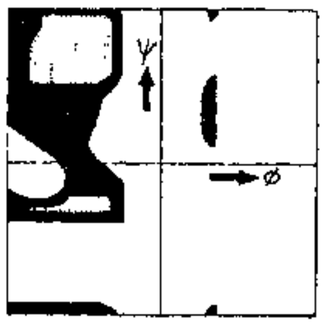

Special issue:
The Ramachandran map and
protein and peptide structures

The prediction of the secondary structure of proteins: Fact or fiction

Gerald D. Fasman
Biochemistry Department, Brandeis University, Waltham, MA 02254-9110, USA

The prediction of the secondary structure of proteins has been widely applied to many areas of protein chemistry. This review gives an overview of its utility, reliability and application to a wide variety of problems in the life sciences.

THE seeds of the protein folding problem were sown by Anfinsen *et al.*[1] in the late 1950s and early 1960s. These classical experiments[1,2] demonstrated that a protein could be denatured and refolded without loss of biological activity. This implied that the amino acid sequence contains all sufficient and necessary information to define the three-dimensional structure of a protein. Since that time, both theoreticians and experimentalists have attempted to further prove and understand the principles underlying this phenomenon[3–6]. Levinthal[7] first stated that the native structure will not be folded by sampling every possible conformation (it would take too long), and the structure attained will probably lie near the minimum of free energy. Anfinsen[2] had proposed that one or more regions of secondary structure, e.g. α-helices, or a two-stranded anti-parallel β-sheet, having marginal stability, would act as nucleation sites and direct the folding. The advent of recombinant DNA techniques has led to an explosion of information concerning sequences and sequence-dependent conformations.

The theoretical efforts could be categorized into three main areas: energetic, heuristic and statistical. Starting with Liquori and coworkers[8] and Ramachandran *et al.*[9], it was first demonstrated that the peptide unit could adopt only certain allowed conformations. Of course all research on protein structure was built on the seminal work of Pauling and Corey[10] who in the early 1950s proposed that several conformations, the α-helix and several β-sheets, were probably the most stable secondary structures available to the polypeptide chain. Since that time the growth of information has been phenomenal and a great deal of it has been derived from the work of the X-ray crystallographer. The approximately 450 X-ray diffraction studies, which have elucidated the three-dimensional conformation of proteins from various sources, have been the main backbone for the development of our understanding of the forces, factors and rationale for the folding of the native structure of proteins.

Appreciating the significance that form follows sequence, the art of the prediction of the secondary structure of proteins (and tertiary structure) had an early start on this problem (for a review see ref. 11). Without delving into the history of the prediction of protein structure, there are some salient facts that have emerged which often bring deep insight to the protein folding problem and are often at odds with the results of the X-ray crystallographer. Several points of controversy will be raised, between fact and fiction (as prediction is often termed) to illustrate that not only does sequence determine secondary structure, but that the environment of crystallization or the composition of the surrounding media can play an important role in determining the final conformation of a protein.

Taking the lead that sequence determines secondary structure, a vast literature was developed using synthetic poly-α-amino acids as models for the conformations of proteins. The work of Blout, Doty, Scheraga and Katchalski laid the groundwork for the facts assembled which showed that each poly-α-amino acid (e.g. poly-L-Lys, poly-L-Glu, etc.) had a preferred structure—at least in a common environment (for a review see ref. 12). Or so it was thought. Work by Doty, Blout and Fasman on poly-L-lysine showed that this monotonous sequence could attain all three conformations, the α-helical, β-sheet, or random conformation. Raising the pH of a poly-L-Lys solution from 7 to 10.5 caused the α-helical conformation to be assumed, as measured by circular dichroism. However,

heating a dilute solution for 15 min at 50°C caused the chain to refold into a β-sheet structure, which was stable on lowering the temperature to room temperature. On cooling overnight at 4°C the chain once again assumed the α-helix[13]. Thus hydrophobic interactions, with their $+\Delta H$, could cause conformational changes, which could also be reversed. This exaggerated example may be the clue for many subtle conformational changes that occur in biological systems. Thus one must add: sequence determines form, but environment also plays an important role. Thus when a ligand binds to a receptor site, the change in environment at the site is very large and it can be sufficient to bring about a conformational change.

The X-ray diffraction-determined structure, although absolutely correct, may not be the relevant biological conformation. Of course the larger the protein's molecular weight, the more probable it is that the X-ray determined structure and the solution structure will be identical. However, for smaller polypeptide chains, such as hormones of approximately 15–50 residues, the environment plays an exceptionally important role. This point will be illustrated via insights gained by prediction algorithms based on statistical studies of X-ray diffraction studies on proteins.

The predictive scheme to be used illustratively will be the Chou–Fasman algorithm, not because of its intrinsic superiority, but because of the familiarity of it to the author. This will illustrate how studies of X-ray determined structures can yield a wealth of information which can be used to question the significance of the individual structural determinations.

The history of the prediction of secondary structures of proteins dates back to the rules of Guzzo[14], which were later complemented by the work of the Scheraga group.

The thesis which should be emphasized is that X-ray crystallography must be taken and used with a discriminating view, as its results, although absolute for each case, must be viewed in the context of the environmental conditions used to obtain the necessary crystals. The desire to know the conformation of biologically active proteins and polypeptides is predicated on the belief that the conformation of these molecules, in part, determines their high degree of specificity and reactivity *in vivo*.

In the Chou–Fasman method, the statistical probability values, called conformational parameters, were obtained for each amino acid residue as found in the X-ray data obtained from 29 proteins[15] as seen in Figure 1. These values have been updated using 64 proteins, and did not vary significantly[16]. These parameters were normalized, based on the fact that in this set of proteins, the average frequency was 38% helices, 20% β-sheets and 33% β-turns. Thus the same residue has the probability of existing in more than one conformation.

What determines its conformational state is the conformational probability of the residues on each side of it. Rough predictive empirical rules were deduced to elucidate the secondary structure. The rules briefly are:
(i) A cluster of four helical residues out of six residues along the protein sequence will nucleate a helix and the helical segment is extended in both directions until an α-tetrapeptide set of breakers with $\langle P_\alpha \rangle < 1$ is reached.
(ii) A cluster of 3-β-formers out of five residues along

P_α			P_β			P_t	f_i		f_{i+1}		f_{i+2}		f_{i+3}	
Glu	1.51	H_α	Val	1.70	H_β	Asn 1.56	Asn	0.161	Pro	0.301	Asn	0.191	Trp	0.167
Met	1.45		Ile	1.60		Gly 1.56	Cys	0.149	Ser	0.139	Gly	0.190	Gly	0.152
Ala	1.42		Tyr	1.47		Pro 1.52	Asp	0.147	Lys	0.115	Asp	0.179	Cys	0.128
Leu	1.21		Phe	1.38		Asp 1.46	His	0.140	Asp	0.110	Ser	0.125	Tyr	0.125
Lys	1.16	h_α	Trp	1.37	h_β	Ser 1.43	Ser	0.120	Thr	0.108	Cys	0.117	Ser	0.106
Phe	1.13		Leu	1.30		Cys 1.19	Pro	0.102	Arg	0.106	Tyr	0.114	Gln	0.098
Gln	1.11		Cys	1.19		Tyr 1.14	Gly	0.102	Gln	0.098	Arg	0.099	Lys	0.095
Trp	1.08		Thr	1.19		Lys 1.01	Thr	0.086	Gly	0.085	His	0.093	Asn	0.091
Ile	1.08		Gln	1.10		Gln 0.98	Tyr	0.082	Asn	0.083	Glu	0.077	Arg	0.085
Val	1.06		Met	1.05		Thr 0.96	Trp	0.077	Met	0.082	Lys	0.072	Asp	0.081
Asp	1.01	I_α	Arg	0.93	i_β	Trp 0.96	Gln	0.074	Ala	0.076	Thr	0.065	Thr	0.079
His	1.00		Asn	0.89		Arg 0.95	Arg	0.070	Tyr	0.065	Phe	0.065	Leu	0.070
Arg	0.98	i_α	His	0.87		His 0.95	Met	0.068	Glu	0.060	Trp	0.064	Pro	0.068
Thr	0.83		Ala	0.83		Glu 0.74	Val	0.062	Cys	0.053	Gln	0.037	Phe	0.065
Ser	0.77		Ser	0.75	b_β	Ala 0.66	Leu	0.061	Val	0.048	Leu	0.036	Glu	0.064
Cys	0.70		Gly	0.75		Met 0.60	Ala	0.060	His	0.047	Ala	0.035	Ala	0.058
Tyr	0.69	b_α	Lys	0.74		Phe 0.60	Phe	0.059	Phe	0.041	Pro	0.034	Ile	0.056
Asn	0.67		Pro	0.55	B_β	Leu 0.59	Glu	0.056	Ile	0.034	Val	0.028	Met	0.055
Pro	0.57	B_α	Asp	0.54		Val 0.50	Lys	0.055	Leu	0.025	Met	0.014	His	0.054
Gly	0.57		Glu	0.37		Ile 0.47	Ile	0.043	Trp	0.013	Ile	0.013	Val	0.053

Figure 1. Conformational parameters for α-helical, β-sheet and β-turn residues in 29 proteins. P_α, P_β, P_t are conformational parameters of α-helical, β-sheet and β-turns. f_i, f_{i+1}, f_{i+2}, f_{i+3} are bend frequencies in the four positions of the β-turn (refs. 15, 19). H_α, H_β, etc., as defined previously (refs. 15, 19).

the protein sequence will nucleate a β-sheet, which is extended in both directions until a β-tetrapeptide set of breakers with $\langle P_\beta \rangle < 1.0$ is reached. (iii) When regions contain both α- and β-forming residues, the overlapping region is helical if $\langle P_\alpha \rangle > \langle P_\beta \rangle$ and vice versa.

A third structure found is the β-turn. The polypeptide folds back on itself (a 180° reversal of chain direction) with an H-bond usually found between residues 1 and 4. Venkatachalam[17] was the first to characterize the various β-turns (11 in all). Four hundred fiftynine turns were found and the frequencies evaluated for each amino acid in each position[18]. β-turns were not previously noted by X-ray crystallographers. To predict β-turns the following is used: the probability of bend occurrence at residue i is calculated from $p_t \geq 0.75 \times 10^{-4}$ ($\approx 1.5 \times \langle p_t \rangle$) as well as $\langle P_t \rangle > 1.00$ and $\langle P_\alpha \rangle < \langle P_t \rangle > \langle P_\beta \rangle$ (ref. 19).

To give an example of the use of the predictive scheme, in Figure 2 is seen the computer output of the secondary structure of bovine pancreatic trypsin inhibitor (BPTI), a 58-residue polypeptide. 87% of the helical, and 95% of β-residues were correctly predicted. With knowledge of the disulfides it is possible to complete the folding to give the complete 3-D structure.

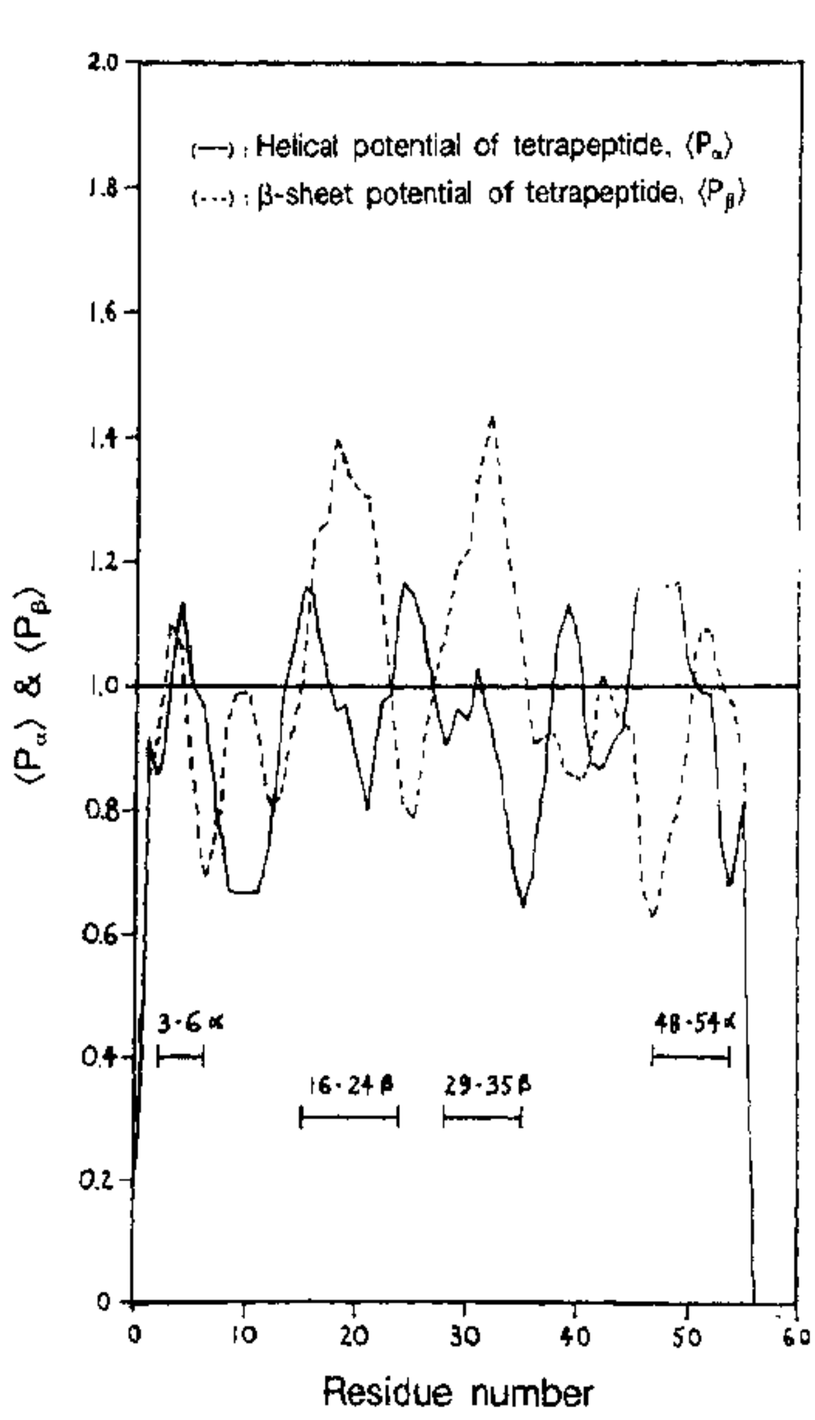

Figure 2. Predicted conformational profile of pancreatic trypsin inhibitor (———) average helical potential $\langle P_\alpha \rangle$, of tetrapeptides i to $i+3$; (– – –) average β-sheet potential $\langle P_\beta \rangle$, of tetrapeptide i to $i+3$. The α- and β-regions found by X-ray are shown as horizontal bars near the bottom of the figure (ref. 42).

To illustrate the use of the predictive scheme in biological systems, some examples will be given. Examples of how these predictions differ from the conclusions from X-ray diffraction studies will also be given.

Conformational changes. These were predicted for concanavalin A. X-ray diffraction structural determination showed 2% α-helical structure and high β-content in concanavalin A (refs. 20 and 21). However in 70% 2-chloroethanol, 55% helicity can be induced as measured by circular dichroism (CD)[22]. If one adds up the regions which have high helical potential $P_\alpha > 1.0$ (Figure 3), although they have still higher β-potential, the total number is 47%—close to that found in chloroethanol[15]. Thus the predictive scheme has the potential to seek out regions with potential for conformational change.

Glucagon—a 29-residue hormone. To test if data obtained from high-molecular-weight proteins could be used with low-MW polypeptides, the structure of glucagon was predicted. The prediction suggested two conformations. Thus between residues 19 and 27 two conformations are possible, i.e. a conformational change is possible. Is this factual? The first conformation was found in dilute solutions, the second in gels, as determined by infrared spectroscopy[23,24]. By choosing the correct concentration it was possible to follow this conformational change by CD[25]. The X-ray diffraction-determined structure (crystals obtained at pH 9.5) indicated 55% helix with three kinks[26]. Thus for small polypeptides the crystal structure may be significantly different from the more dynamic structure in solution. Thus small environmental changes can induce conformational changes—e.g. ligand binding to receptors. Thus the X-ray-determined structure, although beautiful and correct, may be irrelevant to biological function.

In 1975 it was suggested that by changing a few residues in the 19–27 sequence of glucagon, one could lock either structure in, and then the conformation could be determined by CD. The biological activity could then be assessed to see which of these two structures is the important biological one[27]. Eleven years later Hruby *et al.*[28] performed this experiment replacing [Lys-17, Lys-18, Glu-21], which changed the helical probability from $\langle P_\alpha \rangle$ 1.04 to 1.13. In a standard biological assay this analogue was 500% more potent than the native material, and the CD showed it to be locked in the α-helical conformation.

There has been considerable interest in the manner in which precursor proteins are synthesized, the prepro sequences, and the manner in which they are transported across the membrane. Rosenblatt *et al.*[29] had synthesized the prepro parathyroid hormone, a 30-

	Helical regions				β-Sheet regions			
	X-ray	Predicted	$\langle P_\alpha \rangle$	$\langle P_\beta \rangle$	X-ray	Predicted	$\langle P_\beta \rangle$	$\langle P_\alpha \rangle$
Concanavalin A[15]	–	38–43	1.13	1.08	4–9	3–12	1.18	1.08
	81–85	81–86	1.13	1.08	25–29	25–29	1.28	0.90
	–	155–160	1.16	1.08	48–55	47–55	1.16	0.95
	–	180–189	1.17	1.00	59–66	60–67	1.14	1.01
					73–78	73–80	1.13	0.97
					92–97	88–96	1.15	1.05
					106–116	106–113	1.14	0.98
					125–132	124–134	1.11	1.09
					140–144	140–144	1.21	1.17
					173–177	173–177	1.13	1.06
					190–199	190–200	1.18	1.12
					209–215	209–215	1.19	1.03
					–	229–234	1.11	1.08

Figure 3. Comparison of experimental and predicted helical and β-sheet regions in a protein not included in computing the conformational parameters P_α and P_β.

amino acid polypeptide, and showed that it had full biological activity. The pre sequence (the signal peptide) of many proteins, seems to have a hydrophobic core, which may aid in their penetrating the membrane. The predicted sequence is seen in Figure 4. Again two structures were suggested: (a) 20% α, 57% β and (b) 83% α, 0% β. The conformation of the polypeptide was investigated in an aqueous solvent and a solvent of similar dielectric constant to a lipid membrane. The CD spectra in these two environments showed that in an aqueous buffer a CD curve was obtained, yielding 27% α-helix, 43% β-sheet. In hexafluoroisopropanol (a low constant dielectric solvent, similar to a lipid bilayer), a CD curve was found, yielding 46% α, 0% β. Thus this polypeptide, the signal peptide, could adopt either conformation and perhaps the α-helical structure forms the hydrophobic helix predicted to penetrate the membrane. Recently Briggs and Gierasch[30] have examined the *E. coli* λ receptor protein whose signal peptide was predicted to be α-helical by the Chou–Fasman method. They synthesized the signal sequence from the wild type and several mutant proteins; these

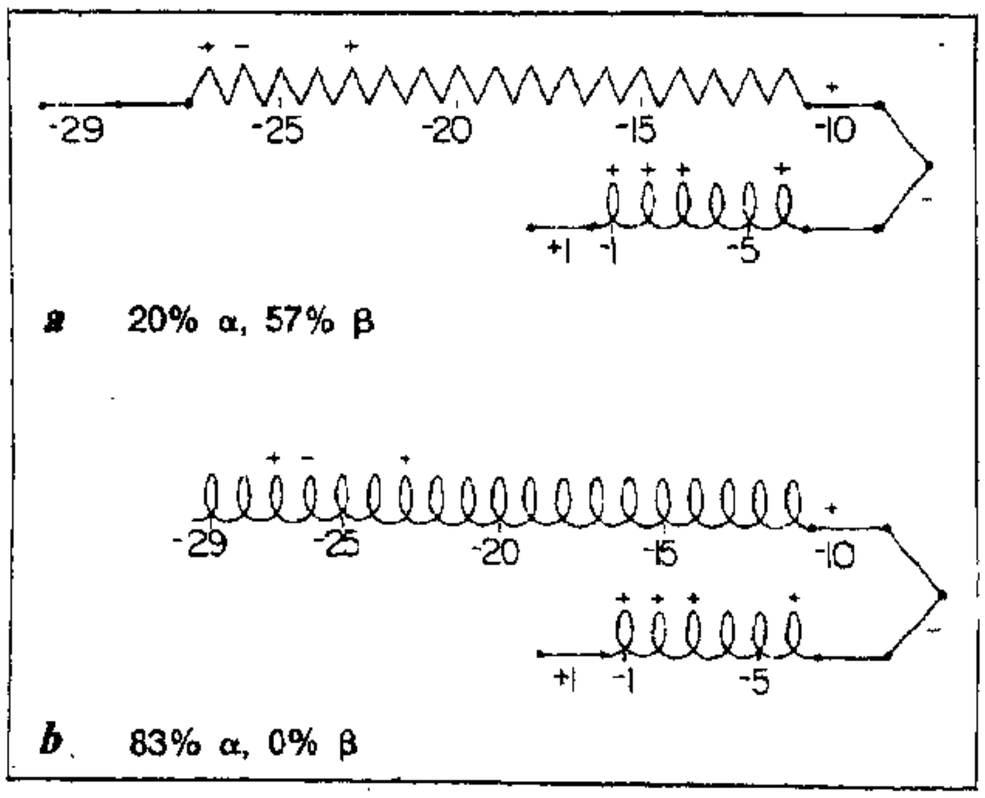

Figure 4. Predicted secondary structure of [D-Tyr^{+1}]prepro PTH-(-29 through +1)-amide. Conformation (*a*) has a high helical content of β-sheet (∿∿∿). Conformation (*b*) is predominantly α-helical (ℓℓℓℓℓ) and devoid of β-sheet. Random coil is indicated by (•——•). The β-turn is indicated between residues −10 and −7 (ref. 29).

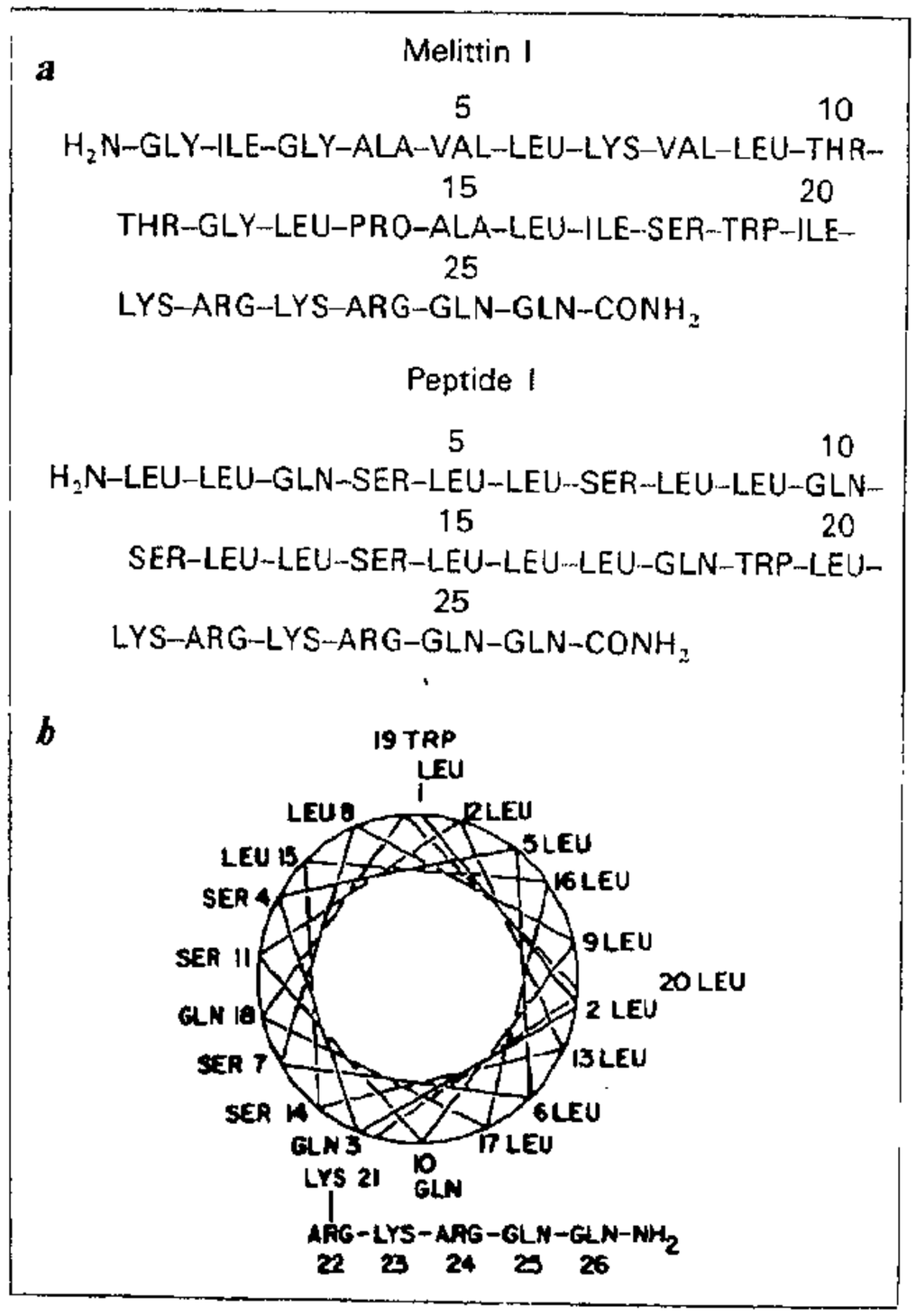

Figure 5. Amino acid sequences of (*a*) melittin I and (*b*) peptide I. Axial projection of α-helical region of peptide I showing the relative location of the side-chains with the segregation of the hydrophobic and hydrophilic residues (ref. 31).

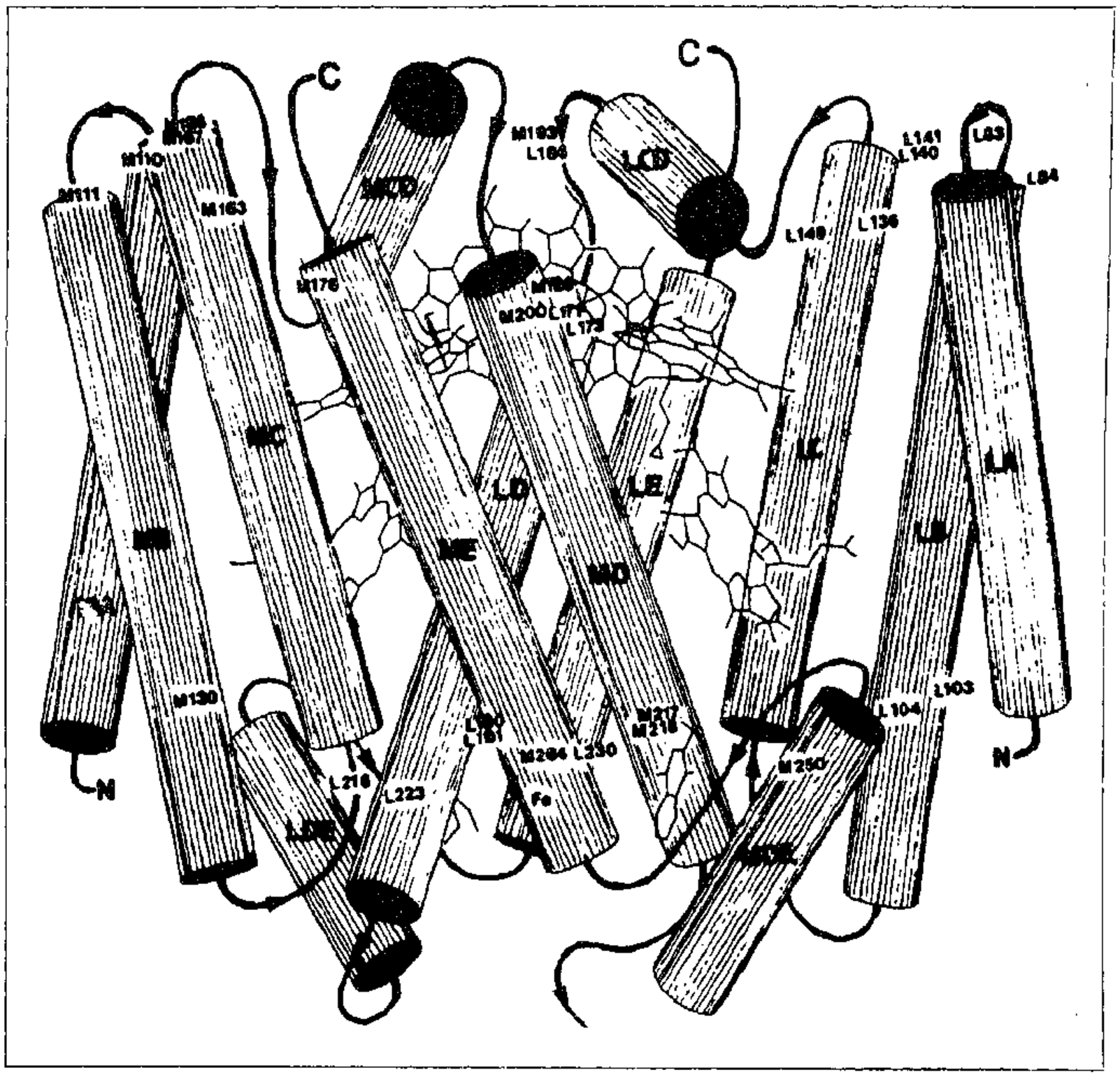

Figure 6. Column model for the core of the reaction center from *Rsp viridis* (ref. 32).

latter are biosynthesized in the *E. coli*, but are not exported as is the wild type. The prediction of the mutant signals suggested that they were not helical. CD measurements on these synthetic-signals confirmed the prediction, namely, the mutants did not assume a helical conformation.

The late Professor Tom Kaiser designed a cytotoxic peptide, similar to melittin (bee venom activity)[31]. The desired α-helix was an amphiphilic α-helix, hydrophobic on one side, hydrophilic on the other (Figure 5). Choosing residues from P_α table (Figure 1) Leu was used wherever there was a hydrophobic residue. The CD showed the new synthetic peptide was α-helical (69%), and it caused hemolysis of erythrocytes and disrupted bilayers as did melittin.

Using the conformational parameters, it is possible to predict whether a single amino acid mutation might cause a conformational change with a consequential change of biological activity. Such a study was performed with the *lac* repressor–*lac* operator interaction. The *lac* repressor (347 amino acids) structure was predicted. The N-terminal region is necessary for binding to the 24 base pair DNA operator. This predicted structure contains 37% α-helix and 35% β-sheet, while CD estimated 40% α and 42% β, in fair agreement. To identify which amino acids of the *lac* repressor are involved in contact with the bases of the DNA, various mutants have been obtained and binding studies performed. The structures of five *lac* repressor mutants were predicted. The first two mutants indicate a loss of repression due to the single amino acid replacements. The next three mutants show no conformational change and retention of binding. The last two examples show that β-turn alterations can cause a loss of activity. Thus it is possible to predict a conformational change based on a single amino acid substitution which may cause a change of biological activity.

Membrane proteins have recently become of major interest and especially how their conformations play a role in transport. Utilizing the *only* membrane protein whose structure has been determined, the *Rhodopseudomonas viridis* reaction center, its secondary structure was predicted by several methods and a comparison of these predictions with the determined structure was made. The X-ray-determined structure[32], is shown in Figure 6. There are three separate chains, named L, M and H (plus other chromophores) which form subunits, which, as α-helices and a small amount of β-sheet, criss-

cross the membrane. It is interesting to note they do not *all* cross the membrane in a perpendicular fashion as most frequently hypothesized.

The various prediction schemes were compared to the X-ray-determined structure.

The well-known *Kyte–Doolittle hydrophobicity profile*[33] (Figure 7) is fairly accurate, but it does not give the correct lengths of the helices. The Kyte–Doolittle algorithm does not predict α-helices or β-sheets, but predicts hydrophobic sequences which may be transmembrane regions. Also using the default value of a window of 11 gives 2 less helices than found. Using a window of 9 gives the correct answer, but how is one to know which size window to use?

The Klein–Kanehisa–DeLisi[34] *method* of determining integral vs peripheral sequences uses the method of discriminant analysis to predict the integral and

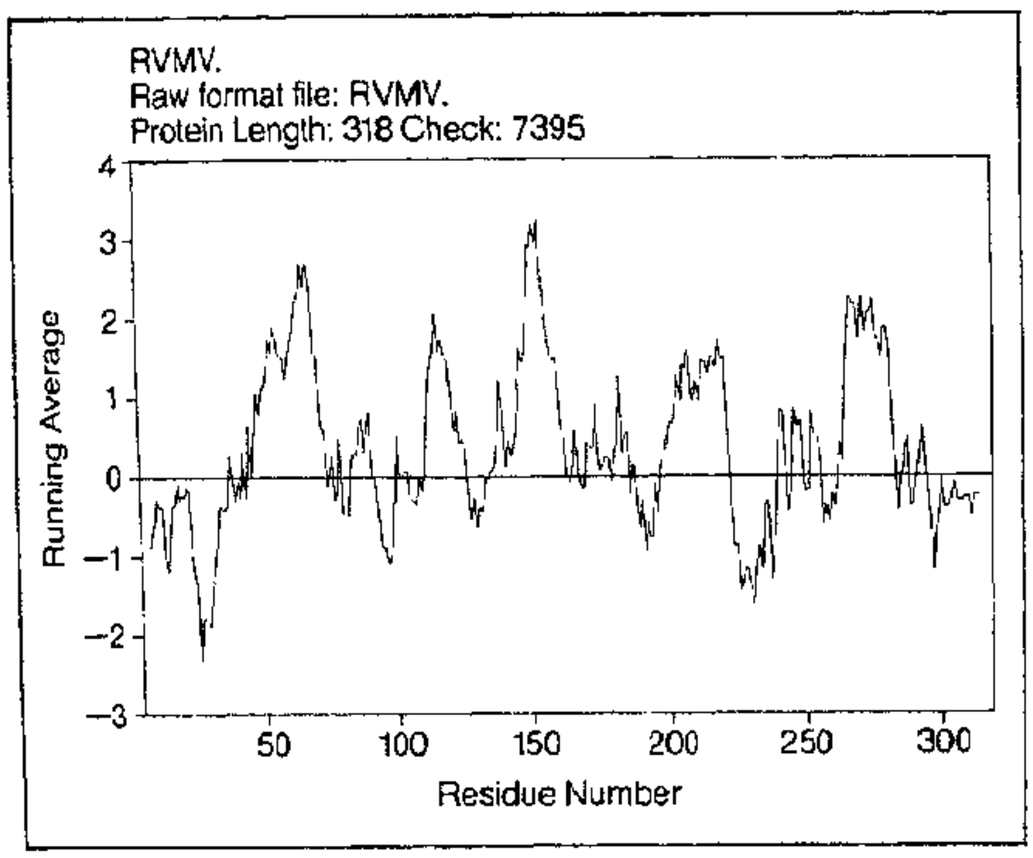

Figure 7. Kyte–Doolittle hydrophobicity plot ($w = 9$) of the M-chain of *Rsp viridis* (ref. 43).

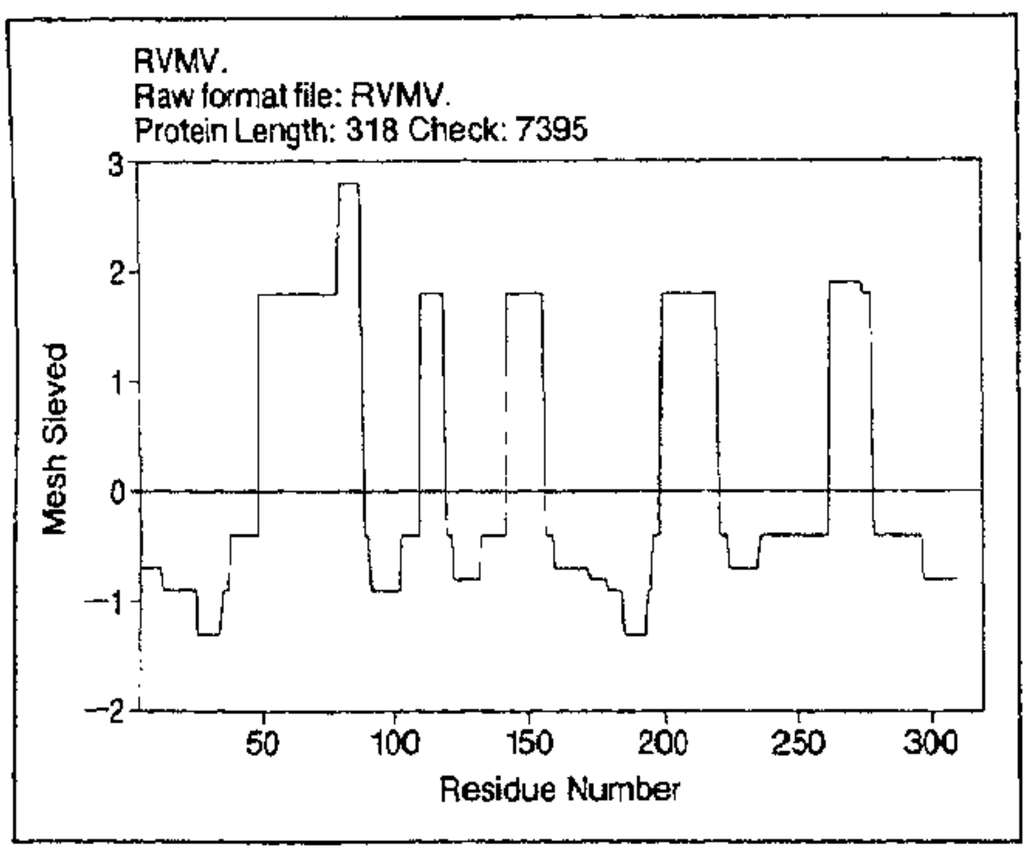

Figure 8. The Sieved–Kyte–Doolittle hydrophobicity plot ($w = 9$) of the M-chain of *Rsp viridis* (ref. 43).

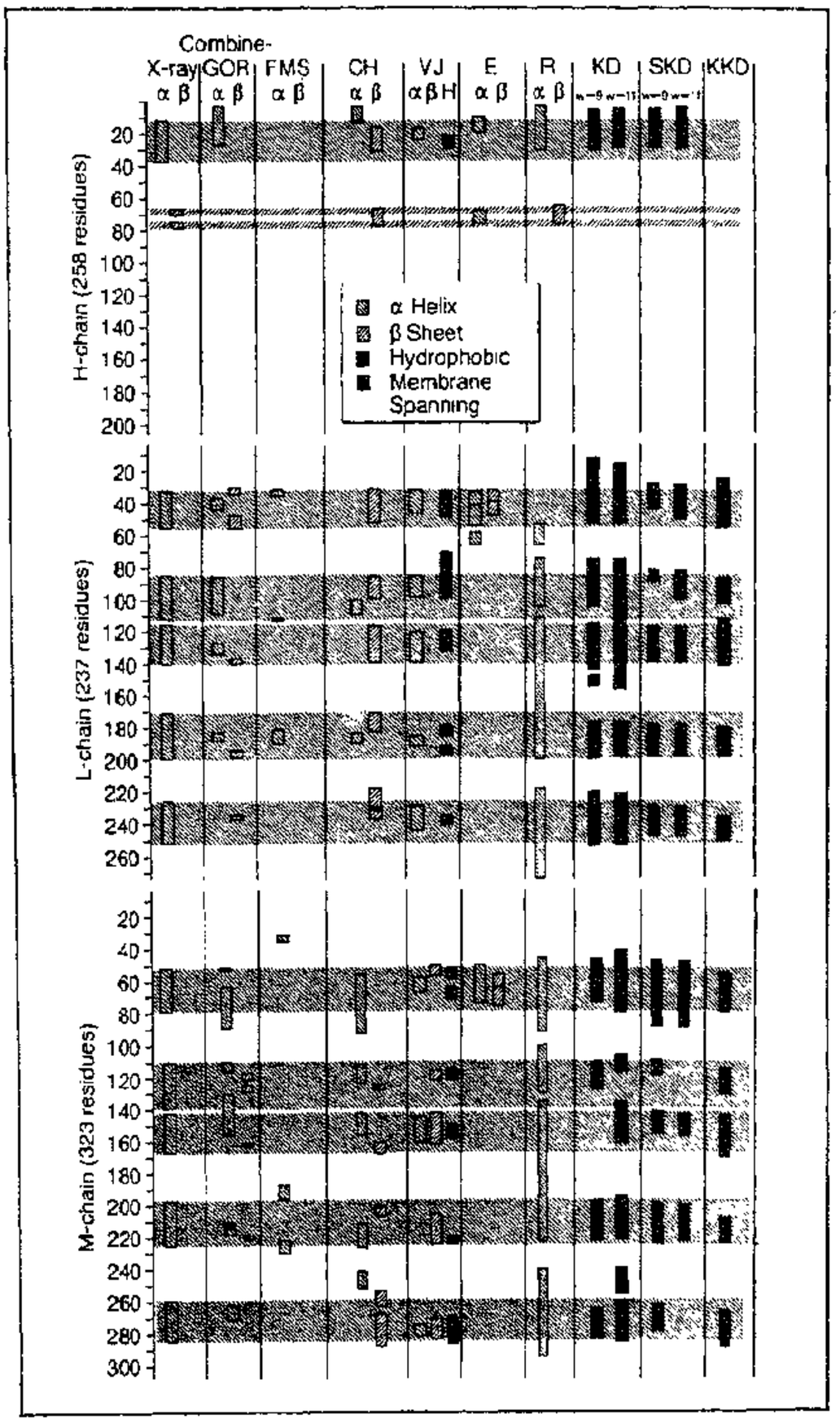

Figure 9. Diagrammatic representation of the transmembrane sequences and their conformations as determined by X-ray diffraction studies. The various predicted sequences by various methods are shown (ref. 43).

peripheral sequences, and excellent agreement was found. However no statement is made regarding their conformation.

There is one additional scheme that should be stressed. It is called sieving[35]. It is similar to a Kyte–Doolittle plot, however, instead of averaging, one takes the running median value with a specified window size, the plot so obtained (Figure 8) yields a much clearer plot than does the Kyte–Doolittle plot.

Several other algorithms were also tested for their accuracy of prediction of the transmembrane regions. In Figure 9 is seen the results for all these predictive schemes. The methods tested are GOR (Biou *et al.*[36]; Garnier *et al.*[37]); FMS (Finer-Moore and Stroud[38]); CH (Chou and Fasman[15]); VJ (Vogel *et al.*[39]); E (Eisenberg *et al.*[40]); R (Rose *et al.*[41]); KD (Kyte and Doolittle[33]); SKD (Bangham[35]); KKD (Klein *et al.*[34]). With the exception of the three methods discussed above (KKD, KD and SKD) the other methods do not

yield accurate results. The only method which predicted the small β-sheet regions was the Chou–Fasman method.

In summary, it has been demonstrated that despite the excellent data that X-ray diffraction studies yield, it often can be misleading concerning the biological significance of that structure. It has been shown that there is a pragmatic use of prediction. However, there is a tendency to blindly accept such results without caution. As there are estimated to be several million different proteins in our universe, and as X-ray crystallography still takes time and crystals, the field of prediction of protein structure still has a future.

1. Anfinsen, C. B., Haber, E., Sela, M. and White, F. H., *Proc. Natl. Acad. Sci. USA*, 1961, **47**, 1309.
2. Anfinsen, C. B., *Science*, 1973, **181**, 233.
3. Levit, M. and Warshel, A., *Nature*, 1975, **253**, 694.
4. Nemethy, G. and Scheraga, H. A., *Q. Rev. Biophys.*, 1977, **3**, 239.
5. McCammon, J. A., Gelin, B. R. and Karplus, M., *Nature*, 1977, **267**, 585.
6. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. and Weiner, P., *J. Am. Chem. Soc.*, 1984, **106**, 765.
7. Levinthal, C., *J. Chem. Phys.*, 1968, **65**, 44.
8. DeSantis, P., Giglio, E., Liquori, A. M. and Ripamonti, A., *Nature*, 1965, **206**, 456.
9. Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V., *J. Mol. Biol.*, 1963, **7**, 95.
10. Pauling, L. and Corey, R. B., *Proc. Natl. Acad. Sci. USA*, 1951, **37**, 729.
11. Fasman, G. (ed.), in *Prediction of Protein Structure and the Principles of Protein Conformation*, Plenum Press, New York, 1989, pp. 193–316.
12. Fasman, G., *Biopolymers*, 1987, **26**, S59.
13. Davidson, B. and Fasman, G. D., *Biochemistry*, 1967, **6**, 1616.
14. Guzzo, A. V., *Biophys. J.*, 1965, **5**, 809.
15. Chou, P. Y. and Fasman, G. D., *Biochemistry*, 1974, **13**, 211.
16. Chou, P. Y., in *Prediction of Protein Structure and the Principles of Protein Conformation* (ed. Fasman, G.), Plenum Press, New York, 1989, pp. 549–586.
17. Venkatachalam, C. M., *Biopolymers*, 1968, **6**, 1425.
18. Chou, P. Y. and Fasman, G. D., *J. Mol. Biol.*, 1977, **115**, 135.
19. Chou, P. Y. and Fasman, G. D., *Biophys. J.*, 1979, **26**, 367.
20. Edelman, G. M., Cunningham, B. A., Reeke, G. N., Jr., Becker, J. W., Waxdall, M. J. and Wang, J. L., *Proc. Natl. Acad. Sci. USA*, 1972, **69**, 2580.
21. Hardman, K. D. and Ainsworth, C. F., *Biochemistry*, 1972, **11**, 4910.
22. McCubbin, W. D., Oikawa, K. and Kay, C. M., *Biochem. Biophys. Res. Commun.*, 1971, **43**, 666.
23. Gratzer, W. B., Bailey, E. and Beaven, G. H., *Biochem. Biophys. Res. Commun.*, 1967, **28**, 914.
24. Epand, R. M., *Can. J. Biochem.*, 1971, **49**, 166.
25. Moŗan, E., Chou, P. Y. and Fasman, G. D., *Biochem. Biophys. Res. Commun.*, 1977, **77**, 1300.
26. Sasaki, K., Dockerill, S., Adamiak, D. A., Tickle, I. J. and Blundell, T. L., *Nature*, 1975, **257**, 751.
27. Chou, P. Y. and Fasman, G. D., *Biochemistry*, 1975, **14**, 2536.
28. Hruby, V., Krstenansky, J., Gysin, B., Pelton, J. T., Trivedi, D. and McKee, R. L., *Biopolymers*, 1986, **25**, 5135.
29. Rosenblatt, M., Habener, J. F., Tyler, F. A., Shephard, G. L. and Potts, J. T., Jr., *J. Biol. Chem.*, 1979, **254**, 1414.
30. Briggs, M. S. and Gierasch, L. M., *Biochemistry*, 1984, **23**, 3111.
31. DeGrado, W. F., Kezdy, E. J. and Kaiser, E. T., *J. Am. Chem. Soc.*, 1981, **113**, 679.
32. Deisenhofer, J. and Michel, H., *Science*, 1989, **245**, 1463; Diesenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H., *Nature*, 1985, **318**, 618.
33. Kyte, J. and Doolittle, R. F., *J. Mol. Biol.*, 1982, **157**, 105.
34. Klein, P., Kanehisa, M. and DeLisi, C., *Biochim. Biophys. Acta*, 1985, **815**, 468.
35. Bangham, J. A., *Anal. Biochem.*, 1988, **174**, 142.
36. Biou, V., Gibrat, J.-F., Levin, J. M., Robson, B. and Garnier, J., *Protein Eng.*, 1988, **2**, 185.
37. Garnier, J. Osguthorpe, D. J. and Robson, B., *J. Mol. Biol.*, 1978, **120**, 97.
38. Finer-Moore, J. and Stroud, R. M., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 155.
39. Vogel, H., Wright, J. K. and Jahnig, F., *EMBO J.*, 1985, **4**, 3625.
40. Eisenberg, D., Weiss, R. M. and Terwilliger, T. C., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 140.
41. Rose, G. D., Gelowitz, A. R., Lesser, G. J., Lee, R. H. and Zehfus, M. H., *Science*, 1985, **229**, 834.
42. Chou, P. Y. and Fasman, G. D., *Adv. Enzymol.*, 1978, **47**, 45.
43. Fasman, G. D. and Gilbert, W. A., *Trends Biochem. Sci.*, 1990, **15**, 89.

ACKNOWLEDGEMENT. This research was supported in part by an NSF Grant (DMB-8713193).