Proc. Natl. Acad. Sci. USA
Vol. 77, No. 8, pp. 4643–4647, August 1980
Biochemistry

Hydrophobic basis of packing in globular proteins

(protein folding/secondary structure/nucleation/hydrophobic domains)

GEORGE D. ROSE* AND SIDDHARTHA ROY

Department of Chemistry, University of Delaware, Newark, Delaware 19711

Communicated by Frederic M. Richards, May 19, 1980

ABSTRACT **The self-assembly of globular proteins is often portrayed as a nucleation process in which the hydrogen bonding in segments of secondary structure is the precondition for further folding. We show here that this concept is unlikely because both the buried interior regions and the peptide chain turns of the folded protein (i.e., inside and outside) are predicted solely by the hydrophobicity of the residues, taken in sequential order along the chain. The helices and strands span the protein, and this observed secondary structure is seen to coincide with the regions predicted to be buried from hydrophobicity considerations alone. Our evidence suggests that linear chain regions rich in hydrophobic residues serve as small clusters that fold against each other, with concomitant or even later fixation of secondary structure. A helix or strand would arise in this folding process as one of a few energetically favorable alternatives for a given cluster, followed by a shift in the equilibrium between secondary structure conformers upon cluster association. The linear chain hydrophobicity alternates between locally maximal and minimal values, and these extrema partition the polypeptide chain into structural segments. This partitioning is seen in the x-ray structure as isodirectional segments bracketed between peptide chain-turns, with the segments expressed most often as helices and strands. The segment interactions define the geometry of the molecular interior and the chain-turns describe the predominant features of the molecular coastline. The segmentation of the molecule by linear chain hydrophobicity imposes a major geometric constraint upon possible folding events.**

Since Kauzmann's report (1), the nature of the hydrophobic core in globular proteins has become a critical focus in studies of structure and folding. Anfinsen (2) showed that folding events leading to the spontaneous emergence of the aggregated hydrophobic core must be a consequence of the linear sequence, and Kuntz (3) recognized that the core region is not well described by a spheroid of interior residues surrounded by surface residues but consists instead of hydrophobic channels that permeate the molecule.

The solvent interface of the protein can be quantitatively characterized with respect to a water-sized probe by measuring those atoms that are accessible to the probe and those that are not (4–7). Richards (8, 9), and Chothia (10, 11) have used this method to derive a series of elegant results.

In this paper, we introduce a method to describe quantitatively the atomic packing within the protein by measuring the protein/protein contact density. Related treatments have been reported (12–14).

Using this measure of packing, we show that the chain hydrophobicity and the packing density are well correlated. We also compare these results with temperature factors from x-ray refinement to show that minima in both packing density and hydrophobicity correspond to regions of greatest conformational flexibility.

The variation in linear chain hydrophobicity induces a partitioning of the molecule into structural segments that interact to form the hydrophobic core of the protein. Segmentation of the molecule imposes significant geometric constraints upon possible steps in the folding pathway, and we suggest a strategy for using this information in a computer program.

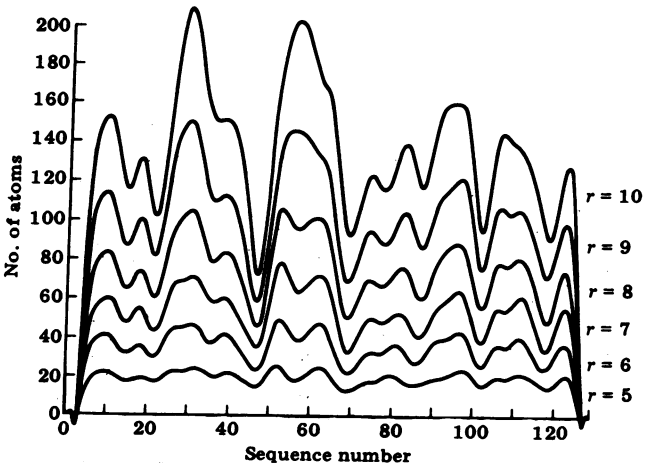

FIG. 1. Family of packing profiles for lysozyme. Each profile represents the number of atoms within a sphere of radius r about each residue's α-carbon. The resultant set of discrete values is then smoothed according to a procedure described in the text. This averaging technique gives rise to chain-end artifacts for the initial and final three-residue interval in each chain. Profiles are displayed for spheres with radii ranging from 5 to 10 Å. Peak positions in the packing profile correspond to linear chain sites buried in the protein's interior.

PROFILE MAPS

The Packing Profile. A profile map of a protein graphs some physical quantity of interest as a function of the amino acid sequence. The *packing profile* is a graph of the protein/protein contact density about each residue in the sequence. Contact density is measured by computing the number of protein atoms other than hydrogen within a sphere of radius r about each α-carbon but excluding intraresidue atoms. Fig. 1 shows a family of packing profiles for lysozyme as r increases from 5 to 10Å. Some residues are completely surrounded by other lysozyme atoms and appear as peaks in density; other residues are partially surrounded by solvent. Because solvent is not tabulated, these loci are less densely packed and are seen as local minima in the profile. For surface residues, this procedure and the method of Lee and Richards (4) provide related information, but the packing profile reveals variation in contact density for interior residues as well.

The packing profiles shown in Fig. 1 are highly smoothed by least-squares fitting of the raw data to a quadratic polynomial using a seven-point moving window. This procedure is repeated three times, after which a cubic spline is interpolated through the fitted values. This simple method gives a smooth, differentiable curve that systematically removes dispersion without affecting the positions of dominant local extrema. The technique gives rise to chain-end artifacts for the initial and final three-residue interval in each chain.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: RMS, root mean square.

* Present address: Dept. of Biological Chemistry, Hershey Medical Center, Pennsylvania State University, Hershey, PA 17033.

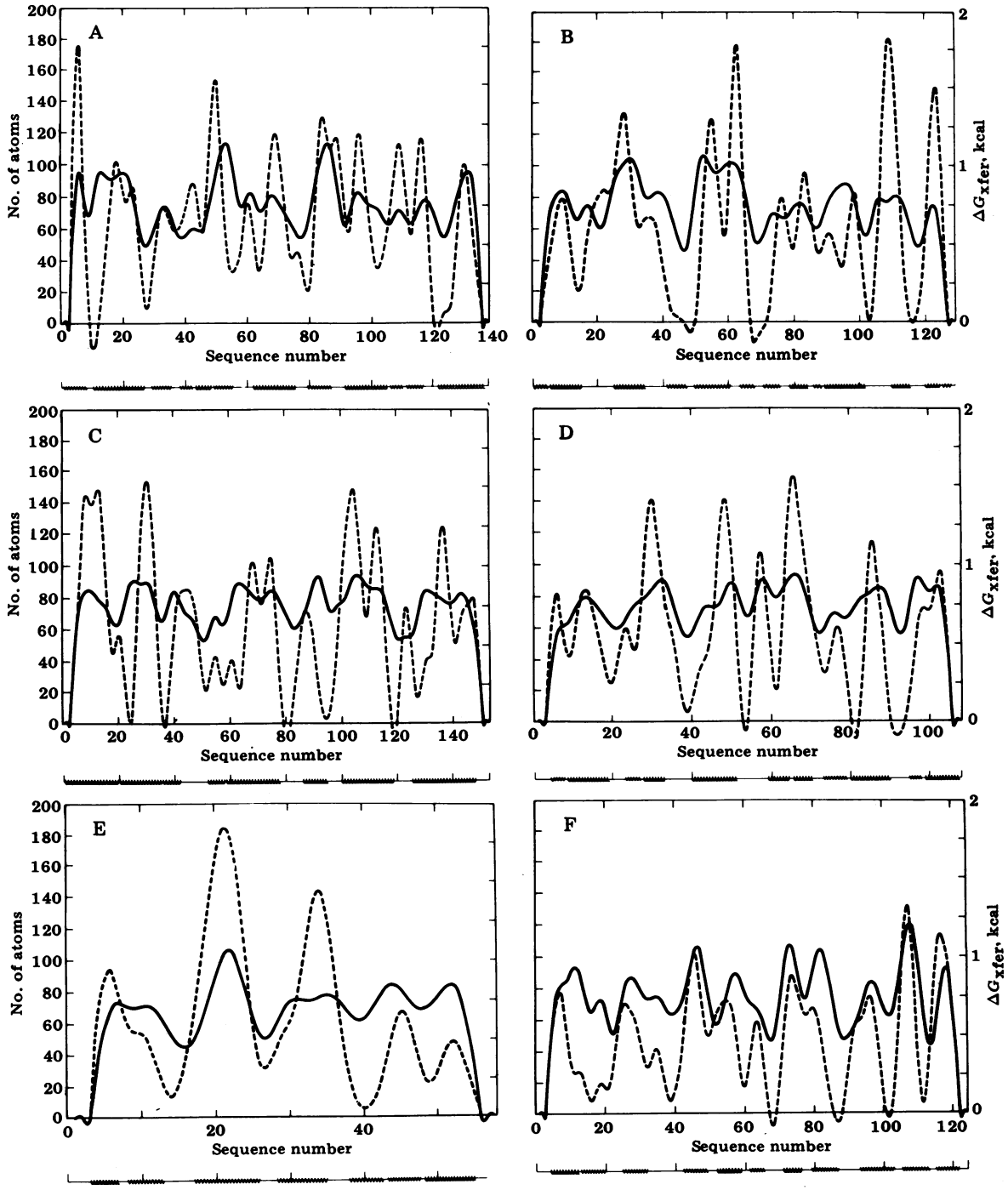

FIG. 2. Comparison between the packing profile and the hydrophobicity profile for a small but diverse set of proteins. The hydrophobicity profile uses only linear sequence information; it plots the free energy of transfer from aqueous to organic solvent (15) against the sequence number. ---, Smoothed hydrophobicity profile; —, smoothed packing profile. The close agreement between peak and valley positions suggests that the chain sites of locally maximal hydrophobicity fold together to establish the closely packed interior of the protein, thereby bringing intervening sites of lesser hydrophobicity to the solvent accessible surface. (*A*) Flavodoxin; (*B*) lysozyme; (*C*) myoglobin; (*D*) parvalbumin; (*E*) pancreatic trypsin inhibitor; and (*F*) RNase A. Observed helices (▬) and strands (〰) are shown running beneath each protein's profile map. Segments of secondary structure usually span the interior of protein molecules, and they are seen here to coincide well with the peaks in packing and hydrophobicity.

Packing profiles about other representative points, such as the side-chain centroid, have been tested, and they do not turn out to be significantly different from the patterns shown in Fig. 1. A family of packing profiles, as depicted, is equivalent to a radial distribution function for each residue in the sequence.

It is evident in Fig. 1 that the peak and valley positions in the packing profile are not very sensitive to the sphere radius over the range 6 to 9 Å. The packing profiles used in the ensuing section have a radius of 8 Å.

Predicting the Hydrophobic Core. The hydrophobicity profile of a protein is a graph of each residue's free energy of transfer from aqueous to organic solvent (15) plotted against the sequence number. The values used for ΔG_{xfer}, expressed in kcal/mol (1 kcal = 4.184 kJ) are: tryptophan, 3.4; phenylalanine, 2.5; tyrosine, 2.3; leucine, 1.8; isoleucine, 1.8; valine, 1.5; methionine, 1.3; alanine, 0.5; cysteine, 0.5; for the rest of the amino acids and proline, 0.

It has been shown (16) that local minima in the hydropho-

bicity profile correspond to the peptide chain turns and solvent-exposed parts of helices. This fact was interpreted to mean that chain sites corresponding to local maxima in hydrophobicity fold to form the hydrophobic core, whereas intervening sites of lesser hydrophobicity are disposed to the solvent-accessible surface of the molecule.

We now present direct evidence for this point of view. In Fig. 2, the hydrophobicity profile is superimposed on the packing profile for a series of proteins. The correspondence between the positions of the local peaks and valleys in these two types of profiles is readily apparent. This agreement between peaks shows that the chain sites of locally maximal hydrophobicity are the parts that pack together in the three-dimensional structure of the protein.

The hydrophobicity profiles shown in Fig. 2 were smoothed by the same procedure used for the packing profiles, yielding a curve that is more highly averaged than curves shown in an earlier publication (16). This additional smoothing supresses a small number of partially buried turns and the solvent-accessible parts of helices, allowing the major solvent-exposed turning points to stand out distinctly.

Visually, the packing and hydrophobicity profiles appear to be strongly correlated in each of the proteins shown in Fig. 2. We have compared profiles by computing the root mean square difference between their extrema. These comparison statistics are discussed in the *Appendix*.

Fig. 3 is included to show a graphical comparison between different proteins. In this control, the hydrophobicity profile for parvalbumin is compared to the packing profile computed for the first 108 residues of lysozyme; it is apparent that the peaks and valleys are not well correlated.

The helices and strands run back and forth across a protein (17), segmenting the molecule (18). These secondary structures, shown on the profile maps in Fig. 2, coincide well with the densely packed regions in the packing profile.

Packing modes between segments of secondary structure are known to be quite specific (19), and it is likely that a peak in hydrophobicity can experience a minor shift in position when accommodating to the particulars of the hydrogen-bonded secondary structure format. These packing particulars may explain why peaks in the packing profile in Fig. 2 are often slightly displaced from peaks in the hydrophobicity profile. However, the near coincidence of the peaks in each case indicates that the interaction centers giving rise to the tertiary structure are largely determined without regard to the specific secondary structure format that is promoted in each case.

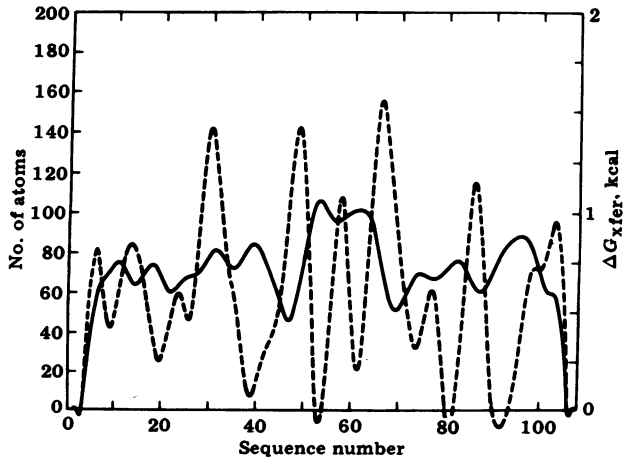

FIG. 3. Control case showing the packing profile for the first 108 residues of lysozyme (—) and the hydrophobicity profile for parvalbumin (- - -). The close agreement in Fig. 2 between peak and valley positions in the two profiles is not evident here. All possible between-protein controls are listed in Table 1.

CONFORMATIONAL FLEXIBILITY

The temperature factors derived from refinement of x-ray coordinates provide useful information about molecular motion (20). These factors have been published in the form of discrete parameters, $\overline{U}^2$, for the main chain atoms of hen egg lysozyme.

These $\overline{U}^2$ values have been smoothed according to the procedure described earlier in this paper. The smoothed $\overline{U}^2$ profile is shown superimposed on the lysozyme hydrophobicity profile in Fig. 4. The five large solvent-exposed loops of the lysozyme molecule, around residues 17, 47, 67, 102, and 117, are easily identified as the dominant valleys in the hydrophobicity profile, and these are seen to correspond especially well to five large peaks in molecular flexibility, as assayed by thermal motion.

IMPLICATIONS FOR PROTEIN FOLDING

Extending the work of Ptitsyn and Rashin (21), Richmond and Richards (22) treated myoglobin as a system of rigid helical segments and Cohen *et al.* (23) showed that only 121 configurations of packed helices are possible after imposition of two innocuous geometric constraints. Choosing allowed configurations by geometric self-consistency is also the basis for the distance geometry approach of Crippen and coworkers (24–26).

It has been shown how linear chain hydrophobicity partitions the protein into structural segments between turns (18), similar to Cohen and coworkers' segmentation of myoglobin (23). In the general case, however, the length of each segment is not known at the start, but the segment is bounded: with all residues in a helix, the length is at a minimum; with all residues in strand, length is maximal; and in other configurations, the length is bounded between these two extremes.

The network of interactions between segments in the native structure is precisely a description of the geometry of the molecular interior. The chain-turns that bracket each segment together with the solvent-exposed parts of helices form the protein/solvent boundary. Thus, the linear chain hydrophobicity is at once a thermodynamic basis for both the secondary structure segmentation and the tertiary structure interactions within the molecule.

In our model for protein folding by *hierarchic condensation* (27, 28), these hydrophobically determined segments interact with each other to form small folding modules of low stability. These modules further coalesce in stepwise fashion, giving rise to a population of equilibrium intermediates that is ultimately pulled in the direction of the successfully folding transition states. A similar pattern of assembly has been described by Crippen (29).

We now suggest that the chain sites corresponding to local maxima in hydrophobicity serve as the folding primitives in such a process.

DISCUSSION

Historically, distinguishing between the inside and the outside of protein molecules has been an elusive task. Of course, any globular molecule will have an inside and an outside in some overall geometric sense; but the molecular coastline of protein molecules is highly irregular, and this fact complicates any quantitative procedure for deciding whether a given component is on the inside or the outside. For example, our intuitive expectation that the center of mass lies deep within the interior is confounded by instances in which this locus is found at or near the exposed active site cleft, between lobes of similar size.

What we have shown is how the amino acid sequence partitions a protein into its inside and outside. By definition, the inside regions are the densely packed chain sites where the

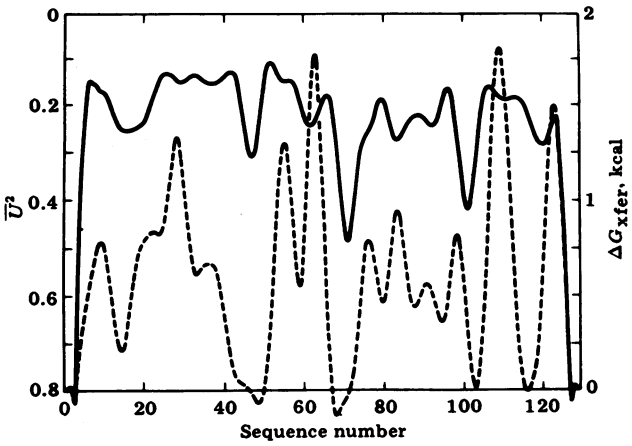

FIG. 4. The hydrophobicity profile (---) compared with main-chain atom temperature factors (—) for lysozyme. Temperature factors, shown here as $\overline{U}^2$, are an experimental measure of conformational flexibility; $\overline{U}^2$ increases with greater flexibility. The $\overline{U}^2$ values shown here, taken from Sternberg *et al.* (20), were smoothed according to the procedure described in the text. Values for $\overline{U}^2$ are displayed so as to increase in the direction of the abscissa in order to show the positive correlation between a large value for $\overline{U}^2$ and a minimum value in the hydrophobicity profile. The major solvent-exposed loops in lysozyme, around residues 17, 47, 67, 102, and 117, correspond to five pronounced local maxima in conformational flexibility.

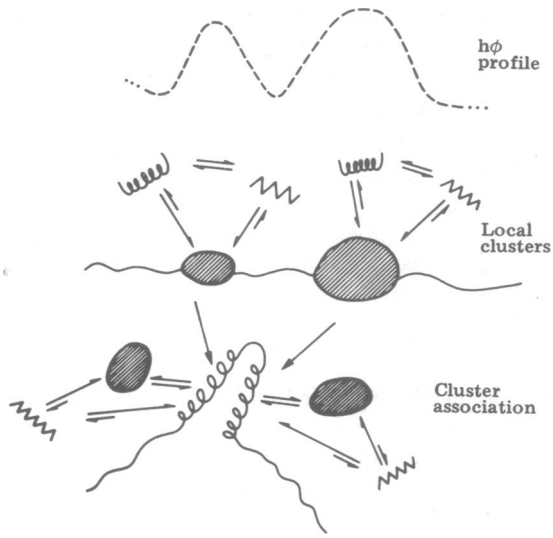

FIG. 5. Formation of hydrogen-bonded segments of secondary structure at sites where the linear chain hydrophobicity is in a local maximum. A few energetically favorable conformers, particularly helix and strand, are thought to be in equilibrium with each other at these sites. The equilibrium is then shifted in one direction or another upon cluster association. In this diagram, a hydrophobicity profile for two hydrophobic chain sites is shown above a schematic representation of the chain itself. Upon cluster association, equilibrium is shifted. In this example, the shift is toward helix.

hydrophobicity is observed to be at a local maximum, whereas the outside regions correspond to less densely packed sites, particularly the chain turns, where the hydrophobicity is at a local minimum.

Kauzmann's generalization that proteins have a hydrophobic core (1) has been qualified by findings by Richards showing that some hydrophobic residues are accessible to solvent (8, 9). Chothia (11) further showed that only a very weak correlation exists between the hydrophobicity of an individual residue and the probability of finding that residue buried in the interior.

We suggest that consideration of the average chain-sequential hydrophobicity serves to reconcile these paradoxical results. Local maxima represent the highest local concentrations of hydrophobic carbon atoms. These sites are removed from contact with the solvent by burying their surfaces against each other, thereby creating the molecule's hydrophobic core. A local minimum in chain hydrophobicity is frequently positive-valued (see Fig. 2) but is nevertheless disposed to the solvent-accessible surface as a consequence of burying a neighboring chain site where the hydrophobicity is at a local maximum.

The hydrophobic character of segment interactions becomes apparent when the core is defined as the sum of chain regions where the protein/protein contact density is maximal.

A protein is partitioned into its inside and outside as a consequence of its folding. The evidence presented here is consistent with a process wherein local sites of maximal hydrophobicity in the amino acid sequence fold together to establish the interior core, bringing intervening sites to the solvent.

The self-assembly of globular proteins is often portrayed as a nucleation process in which the fixation of hydrogen-bonded secondary structure directs further folding. That concept now seems unlikely because chain segmentation can be predicted without explicit knowledge of the helices and strands.

Our evidence suggests that a more likely first step in folding is the formation of hydrophobic clusters corresponding to the local maxima in chain hydrophobicity. Next, cluster association would occur, leading through a hierarchy of intermediates to the native conformation. In this stepwise process, a helix or a strand would arise as one of a few favorable conformational alternatives accessible to a given hydrophobic cluster, with the equilibrium between secondary structure conformers possibly shifted in one direction or another upon cluster association, as illustrated in Fig. 5.

Baldwin and coworkers (30–32) have new experimental evidence that bears on the order of hydrogen-bonding events during stepwise assembly, and Kanehisa and Tsong (33–35) have described a kinetics of self-assembly starting from identical hypothetical clusters and leading to the native state. Related treatments of the folding–unfolding transition have been presented by Creighton (36) and by Go and Taketomi (37–39); and Tanaka and Scheraga (40, 41) have also modeled the folding process in graduated stages.

APPENDIX

It is desirable to have a simple, single-number measure of the agreement between extrema in the packing and hydrophobicity profiles. We have compared these profiles by computing the root mean square (RMS) difference between their respective extrema, with the packing profile taken as a measured standard. For each peak or valley in the hydrophobicity profile we compute the distance along the sequence to the closest peak or valley in the packing profile. Then, the RMS difference between the two profiles is given by

$$\text{RMS} = \left(1/N \sum_{i=1}^{N} (r_i^{h\phi} - r_j^{\text{pack}})^2\right)^{1/2} \quad [1]$$

in which $r_i^{h\phi}$ is the sequence coordinate of the ith extrenum in the hydrophobicity profile, r_j^{pack} is the sequence coordinate of the nearest corresponding extrenum in the packing profile; and N is the number of extrema in the hydrophobicity profile.

Table 1 lists the RMS differences between the hydrophobicity profile for each protein and the packing profiles for all proteins in the test set. The comparison of a protein's hydrophobicity profile with another protein's packing profile serves as a control. Since two different proteins usually differ in number of residues, the comparison is only made between the shorter of the two and an equivalent number of residues in the longer. In almost every case, the RMS difference between a protein's hydrophobicity profile and its own packing profile is significantly

Table 1. Root mean square (RMS) difference between extrema in two profiles

Packing profile	Hydrophobicity profile*					
	Fla	Lzm	Mb	Myo	PTI	RNase A
Fla	2.6	3.8	3.4	4.0	1.9	3.3
Lzm	3.5	2.4	3.5	4.0	3.9	3.0
Mb	3.6	3.5	2.7	4.0	3.7	3.1
Myo	2.2	3.6	2.3	1.3	2.4	3.0
PTI	2.6	1.8	2.4	2.7	0.9	2.6
RNase A	3.4	2.5	3.3	2.9	2.9	1.2

* All hydrophobicity profiles are compared to each protein's packing profile as a measured standard. Fla, flavodoxin; Lzm, lysozyme; Mb, myoglobin; Myo, parvalbumin; PTI, pancreatic trypsin inhibitor. Values along the diagonal are RMS differences between each protein's own packing and hydrophobicity profiles. Off-diagonal values along each row serve as controls and measure RMS differences between the packing profile for a protein and the hydrophobicity profile for every other protein. RMS values are related to the standard deviation, as discussed in the text. Two SDs can be compared by using the F test, a ratio of the squares of the SDs (ratio of variances). The F test measures the likelihood that the observed ratio would have occurred by chance alone. In almost all cases, the RMS value for a protein's packing profile and its own hydrophobicity profile is significantly less than that protein's packing profile and any other protein's hydrophobicity profile, as assessed by the F test. The case of flavodoxin versus pancreatic trypsin inhibitor appears to be an exception based upon RMS values alone, but it can be understood by referring to the structures themselves. Twenty-seven of the 30 controls have a level of significance exceeding the 90% confidence level—that is, better than would have been expected by chance alone 90% of the time. Twenty-two are above the 95% confidence level and 13 are above the 99% confidence level.

less than between that protein's hydrophobicity profile and every other protein's packing profile.

The RMS difference between two arbitrary profile maps is limited by the most likely difference between corresponding extrema. Globular proteins have an average of one chain turn every eight residues (42), so the expected mean difference between corresponding extrema is only about 4. For this reason, the F test is too stringent a measure, but we have done nothing to compensate for this fact. The worst control in Table 1 is the comparison between the pancreatic trypsin inhibitor hydrophobicity profile and the flavodoxin packing profile, and this requires explanation. In comparisons of this sort, the expected RMS difference increases as a function of the number of residues being compared. This tendency biases some of the controls (off-diagonal entries) in Table 1 toward seemingly better agreement but, apart from limiting our sample set to proteins of similar size, we have not attempted to correct for this fact since significance is already at a convincing level.

The RMS difference between the first 58 residues of the flavodoxin hydrophobicity and packing profiles is 2.0, and the pattern of segmentation is similar in the two proteins. The observed agreement is therefore to be expected.

Whenever extrema in the packing and hydrophobicity profiles are not in 1:1 correspondence, the RMS measure fails to detect correlations that are apparent in a graphical comparison. Throughout Fig. 2, for example, there are instances in which a single packing peak forms a tight envelope about two hydrophobic peaks. In such cases the RMS difference deteriorates although the presumed agreement is good. Although a more refined statistical test would undoubtedly show greater differentiation, the RMS index is an easily understood, single-number measure that distinguishes effectively between a protein and its controls, and we have adopted it for this reason.

Note Added in Proof. We have been informed of a new result by Kanehisa and Tsong (41) showing that hydrophobic domains are enriched in helices and β-sheets.

We wish to thank Drs. Christian Sander and Fred Richards for useful discussion and critical remarks and Dr. Michael Sternberg for providing us with numerical $\overline{U}^2$ values. Some of this work was completed while attending a 1979 Centre Européen de Calcul Atomique et Moléculaire summer workshop on protein structure in Orsay, France. The work was also supported by U.S. Public Health Service Grant GM 27370.

1. Kauzmann, W. (1959) *Adv. Protein Chem.* **14**, 1–63.
2. Anfinsen, C. B. (1973) *Science* **181**, 223–230.
3. Kuntz, I. D. (1972) *J. Am. Chem. Soc.* **94**, 8568–8572.
4. Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55**, 379–400.
5. Finney, J. L. (1975) *J. Mol. Biol.* **96**, 721–732.
6. Finney, J. L. (1978) *J. Mol. Biol.* **119**, 415–441.
7. Shrake, A. & Rupley, J. A. (1973) *J. Mol. Biol.* **79**, 351–372.
8. Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.
9. Richards, F. M. (1974) *J. Mol. Biol.* **82**, 1–14
10. Chothia, C. (1975) *Nature (London)* **254**, 304–308.
11. Chothia, C. (1976) *J. Mol. Biol.* **105**, 1–44.
12. Kauzmann, W., Moore, K. & Schultz, D. (1974) *Nature (London)* **248**, 447–449.
13. Crippen, G. M. & Kuntz, I. D. (1978) *Int. J. Pept. Protein Res.* **12**, 47–56.
14. Kuntz, I. D. & Crippen, G. M. (1979) *Int. J. Pept. Protein Res.* **13**, 223–228.
15. Nozaki, T. & Tanford, C. (1971) *J. Biol. Chem.* **246**, 2211–2217.
16. Rose, G. D. (1978) *Nature (London)* **272**, 586–590.
17. Levitt, M. & Chothia, C. (1976) *Nature (London)* **261**, 552–558.
18. Rose, G. D. & Seltzer, J. P. (1977) *J. Mol. Biol.* **113**, 153–164.
19. Choothia, C., Levitt, M. & Richardson, D. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4130–4134.
20. Sternberg, M. J. E., Grace, D. E. P. & Phillips, D. C. (1979) *J. Mol. Biol.* **130**, 231–253.
21. Ptitsyn, O. B. & Rashin, A. A. (1975) *Biophys. Chem.* **3**, 1–20.
22. Richmond, T. J. & Richards, F. M. (1978) *J. Mol. Biol.* **119**, 537–555.
23. Cohen, F. E., Richmond, T. J. & Richards, F. M. (1979) *J. Mol. Biol.* **132**, 275–288.
24. Crippen, G. M. (1977) *J. Comp. Phys.* **24**, 96–107.
25. Crippen, G. M. & Havel, T. F. (1978) *Acta Crystallogr.* **A34**, 282–284.
26. Kuntz, I. D., Crippen, G. M. & Kollman, P. A. (1979) *Biopolymers* **18**, 939–958.
27. Rose, G. D. (1979) *J. Mol. Biol.* **134**, 447–470.
28. Rose, G. D. (1980) *Biophys. J.*, in press.
29. Crippen, G. M. (1978) *J. Mol. Biol.* **126**, 315–332.
30. Schmid, F. X. & Baldwin, R. L. (1979) *J. Mol. Biol.* **135**, 199–215.
31. Labhardt, A. M. & Baldwin, R. L. (1979) *J. Mol. Biol.* **135**, 231–244.
32. Labhardt, A. M. & Baldwin, R. L. (1979) *J. Mol. Biol.* **135**, 245–254.
33. Kanehisa, M. I. & Tsong, T. Y. (1978) *J. Mol. Biol.* **124**, 177–194.
34. Kanehisa, M. I. & Tsong, T. Y. (1979) *Biopolymers* **18**, 1375–1388.
35. Kanehisa, M. I. & Tsong, T. Y. (1979) *Biopolymers* **18**, 2913–2928.
36. Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol.* **33**, 231–297.
37. Go, N. & Taketomi, H. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 559–563.
38. Go, N. & Taketomi, H. (1978) *Int. J. Pept. Protein Res.* **13**, 235–252.
39. Go, N. & Taketomi, H. (1978) *Int. J. Pept. Protein Res.* **13**, 447–461.
40. Tanaka, S. & Scheraga, H. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1320–1323.
41. Kanehisa, M. I. & Tsong, T. Y. (1980) *Biopolymers* **19**, in press.
42. Tanaka, S. & Scheraga, H. A. (1977) *Macromolecules* **10**, 291–304.
43. Rose, G. D. & Wetlaufer, D. B. (1977) *Nature (London)* **268**, 769–770.