Empirical Studies of Hydrophobicity. 2. Distribution of the Hydrophobic, Hydrophilic, Neutral, and Ambivalent Amino Acids in the Interior and Exterior Layers of Native Proteins[1a]

H. Meirovitch and H. A. Scheraga*[1b]

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853.
Received November 16, 1979

ABSTRACT: Previously, the 20 naturally occurring amino acids were classified into three groups: hydrophobic—Cys, Phe, Ile, Leu, Met, Val, and Trp; hydrophilic—Asp, Glu, Lys, Asn, Pro, Gln, Arg, Ser, and Thr; and a group of the neutral and ambivalent amino acids—Ala, His, Gly, and Tyr. In this paper, the fractions of occurrence of residues in these groups are determined for spherical layers (whose radii are expressed in units of the root-mean-square radius of gyration, R_g) around the centers of mass of 19 proteins. Deviations of these fractions from a random distribution constitute a measure of the effect of water on the structure of a native protein. Two sets of results are presented. In one, the residue is represented by its C^α atom and in the other by a remote side-chain atom. The main results are as follows: (1) In the larger proteins, a sphere of radius $0.75R_g$ is found to be radially homogeneous in the sense that there is no significant variation in the fractions of the three groups of amino acids in smaller layers of radii $\leq 0.75R_g$. The fraction of the hydrophobic residues in this sphere is higher than their average fraction in the whole protein (for C^α: 0.43 vs. 0.29, respectively). Their fraction decreases gradually toward the outermost layer, $\geq 1.3R_g$ (0.11 for C^α). The distribution of the hydrophilic residues also differs from a random one, but in the opposite direction (for C^α: in the interior 0.32 vs. 0.47; in the outermost layer, 0.66). The fractions of the neutral–ambivalent residues fluctuate slightly around their average fraction in the whole protein. Qualitatively, this behavior is found to exist also for the individual proteins as well as for the whole set. (2) The smaller proteins behave very much the same as the larger ones with only a few differences: (a) In their inner layers of radii $\leq 0.75R_g$, they are homogeneous with respect to C^α but not with respect to the side-chain atoms. (b) The fraction of hydrophobic residues in this sphere is significantly higher than for the larger proteins (for C^α: 0.51 vs. 0.43, respectively). This is accounted for by geometric effects and hydrophobic interactions. (c) The fractions of hydrophobic residues decrease, and those of hydrophilic residues increase, gradually from the inner to the outer layer for all proteins except BPTI and HIPIP. The above results are compared to those of Krigbaum and Komoriya, who used a similar methodology. (3) Using results for the fractions of residues buried in the proteins, obtained by Chothia and by Wertz and Scheraga (WS), and adopting the criterion of WS for a residue to be buried, we are able to define three layers (not necessarily spherical). Two of them (I and II) contain the buried residues and the third (III) the exposed ones. Application of our classification of amino acids to these results shows that the buried region is not uniform in composition, the innermost layer (I) being more hydrophobic and less hydrophilic than layer II. Layer II itself is more hydrophobic and less hydrophilic than the exposed layer (III). Comparison of these results to those obtained by our method for spherical layers indicates that the proteins in the sample do not deviate much from spherical shape; it also leads to the conclusion that the buried region of these proteins corresponds approximately to a sphere of radius R_g around the center of mass.

Introduction

Water plays an important role in protein folding. One consequence of hydration is that hydrophobic amino acid residues tend to be buried in the interior of the protein molecule to avoid unfavorable contacts with water, whereas hydrophilic residues tend to lie preferentially on the surface in contact with water.[2] In the first paper of this series,[3] the 20 naturally occurring amino acids were classified into four groups: hydrophobic, hydrophilic, neutral, and ambivalent. This classification is based on the values of two parameters,[4] calculated from the known structures of 19 proteins. The first parameter is the average distance of an amino acid from the center of mass of the protein and the second is the average angle between the two vectors, centers-of-mass-to-C^α and C^α-to-a-remote-side-chain-atom. By identifying smaller values for the average distance with hydrophobicity and larger ones with hydrophilicity, it was obvious that, on the average, the interior of the proteins is more hydrophobic and less hydrophilic than the exterior.

In the present paper, this classification is used to determine the distributions of the hydrophobic, hydrophilic, and the group of neutral and ambivalent residues as a function of distance from the center of mass of the protein, and the deviations of these distributions from a random one. The magnitudes of these deviations constitute a measure of the effect of water on the structure of the native protein. For this purpose, we define concentric spherical layers around the center of mass and the fractions of residues in the three groups of amino acids in each layer are evaluated. In order to scale the results obtained from a sample of proteins of different molecular weights, the radii of these layers are defined in units of the root-mean-square radius of gyration R_g. In many cases, the decision as to whether a residue is located within the limits of a given layer is not straightforward since different parts of the residue can pertain to different layers (especially for the larger side chains in the smaller proteins, where R_g is relatively small). For simplicity, we therefore represent a residue either by its C^α atom or, for comparison, by a remote side-chain atom.[3-5] Obviously, one would expect to find larger differences between the two sets of results for the smaller proteins, where the size of the side chain is relatively large as compared to R_g.

In view of the important role played by the hydrophobic interaction in the folding of a protein,[2] we focus attention on the concentration of hydrophobic residues in the interior layers. Our analysis is applied separately to the group of larger proteins and to that of the smaller ones. From results obtained in ref 3, we expect to find differences in the concentrations of the hydrophobic residues in the interiors of the two groups, due to both geometric effects and hydrophobic interactions.

We also analyze the *individual* proteins of the sample in order to examine their behavior and their fluctuations from the average properties of the whole protein sample.

0024-9297/80/2213-1406$01.00/0

Table I
Average Fractions of the Hydrophobic, Hydrophilic, and Neutral-Ambivalent Residues in Spherical Layers around the Centers of Mass of 19 Proteins

spherical layers in units of R_g	C^α atoms				typical side-chain atoms[a]			
	hydrophobic	hydrophilic	neutral and ambivalent	no. of C^α's in layer	hydrophobic	hydrophilic	neutral and ambivalent	no. of typical side-chain atoms in layer
				Group A[b]				
global[c]	0.28	0.48	0.24	1148	0.28	0.48	0.24	1143[d]
0-0.50	0.55	0.24	0.21	62	0.72	0.13	0.15	87
0.50-0.65	0.49	0.27	0.24	97	0.63	0.17	0.20	93
0.65-0.75	0.50	0.34	0.16	105	0.52	0.24	0.24	59
0-0.75	0.51	0.29	0.20	264	0.64	0.17	0.19	239
0.75-0.95	0.33	0.46	0.21	298	0.36	0.39	0.25	247
0.95-1.15	0.19	0.56	0.24	308	0.16	0.57	0.27	274
≥1.15	0.10	0.60	0.30	278	0.09	0.68	0.24	383
1.15-1.3	0.13	0.61	0.26	167	0.11	0.65	0.24	179
≥1.3	0.06	0.58	0.36	111	0.07	0.70	0.23	204
				Group B[b]				
global	0.29	0.47	0.24	2096	0.29	0.47	0.24	2090[d]
0-0.50	0.44	0.32	0.24	173	0.44	0.31	0.25	169
0.50-0.65	0.43	0.24	0.33	189	0.46	0.26	0.28	203
0.65-0.75	0.43	0.34	0.23	166	0.50	0.27	0.23	168
0-0.75	0.43	0.30	0.27	528	0.47	0.28	0.25	540
0.75-0.95	0.35	0.44	0.21	512	0.38	0.37	0.25	455
0.95-1.15	0.21	0.52	0.26	581	0.23	0.50	0.27	467
≥1.15	0.15	0.62	0.23	475	0.11	0.67	0.22	628
1.15-1.3	0.18	0.59	0.23	271	0.13	0.63	0.24	303
≥1.3	0.11	0.66	0.23	204	0.10	0.70	0.20	325

[a] See Table II of ref 3 for identification of the side-chain atoms used in the computations. [b] See Table I of ref 3 for a listing of the proteins in each of these groups. [c] This refers to the whole protein, and the entries for C^α atoms must therefore be the same as those for the side-chain atoms. [d] The positions of several C^α and side-chain atoms were not well-defined in the X-ray data, and these residues were not included. This is why the total numbers of C^α's and side-chain atoms are not the same.

The results of this analysis provide information as to whether the general "rule" (that hydrophobic residues concentrate in the interior of a protein and hydrophilic residues on the outside) is simply an *average* property or applies individually to each protein. If the latter were true, then this general "rule" would be a much stronger one.

Our analysis is similar to one carried out by Krigbaum and Komoriya (KK),[6,7] who introduced their own scale of polarity (for comparison, see Table IV of ref 3). Most of our results agree well with theirs, but some do not. We also account for the results in a different way. The work of KK will be considered further in the Results and Discussion section.

It should be pointed out that the spherical distributions of the three groups of amino acids considered in the present paper are physically significant only for globular proteins, i.e., those which do not deviate much from spherical shape (for such globular proteins, the buried and exposed regions can be expressed in terms of spherical layers). In order to test the validity of the assumption of globularity, we therefore compare our results to those for the fractions of amino acids buried, obtained by Chothia[8] and by Wertz and Scheraga,[9] who used methods different from ours (and not based on the assumption of spherical shape). Such a comparison can also provide more information about the structure of proteins.

Finally, it should be emphasized that the results of the present paper not only provide information as to how a globular protein is organized but may also be incorporated in protein-folding algorithms.

Methods

As in ref 3, we use a sample of 19 proteins, which is divided into two groups: the 11 smaller proteins (group A) and the 8 larger ones (group B); for details, see Table I of ref 3. For each protein, the center of mass and the root-mean-square radius of gyration, R_g, are calculated. According to the classification of ref 3, the amino acids are divided here into three groups: the seven hydrophobic amino acids, Cys, Phe, Ile, Leu, Met, Val, and Trp; the nine hydrophilic ones, Asp, Glu, Lys, Asn, Pro, Gln, Arg, Ser, and Thr; and a group which includes both the neutral amino acid His and the three ambivalent amino acids Ala, Gly, and Tyr.

The fraction of hydrophobic residues in a given layer is determined by dividing the total number of C^α atoms of hydrophobic residues by the total number of all C^α atoms in the layer; similarly for the other two groups. The same types of calculations were carried out for a remote side-chain atom rather than C^α. The identification of the side-chain atom for each residue is given in Table II of ref 3. For Gly, the C^α atom was used in both sets of calculations.

Results and Discussion

In Table I, results are given for the average fractions of the hydrophobic, hydrophilic, and the group of neutral and ambivalent residues in spherical layers around the centers of mass of 19 proteins. In Table II, similar data are given separately for each protein of the set. These results are first discussed for the larger and smaller proteins and then compared with those of Krigbaum and Komoriya[6] and others.

Results for the Larger Proteins. We discuss first the results in Table I for the larger proteins (group B). As expected, the hydrophobic residues occur more frequently in the inner sphere of radius $0.75R_g$ (0.43 for C^α; 0.47 for the remote side-chain atoms) than in the whole proteins (0.29). As the distance from the center of mass increases, their frequency of occurrence decreases gradually to ~0.10 in the outermost layer. The opposite behavior is seen for the hydrophilic residues. In the sphere of radius $0.75R_g$

Table II
Average Fractions of the Hydrophobic, Hydrophilic, and Neutral-Ambivalent Residues in Spherical Layers around the Center of Mass of Each Protein of the Set of 19

spherical layers in units of R_g	C^α atoms				remote side-chain atoms[a]			
	hydrophobic	hydrophilic	neutral and ambivalent	no. of residues	hydrophobic	hydrophilic	neutral and ambivalent	no. of residues
				Group A[b]				
				Rubredoxin				
global	0.29	0.54	0.17	54	0.29	0.54	0.17	54
0-0.85	0.39	0.39	0.22	18	0.54	0.23	0.23	13
0.85-1.10	0.24	0.67	0.09	21	0.37	0.44	0.19	16
⩾1.10	0.27	0.53	0.20	15	0.09	0.78	0.18	23
				Ferredoxin				
global	0.35	0.40	0.24	54	0.35	0.40	0.24	54
0-0.85	0.44	0.25	0.31	16	0.71	0.07	0.21	14
0.85-1.10	0.32	0.41	0.27	22	0.25	0.35	0.40	20
⩾1.10	0.31	0.56	0.13	16	0.20	0.70	0.10	20
				Bovine Pancreatic Trypsin Inhibitor (BPTI)				
global	0.28	0.45	0.27	58	0.28	0.45	0.27	58
0-0.85	0.24	0.52	0.24	25	0.53	0.40	0.07	15
0.85-1.10	0.57	0.29	0.14	14	0.10	0.58	0.32	19
⩾1.10	0.11	0.47	0.42	19	0.25	0.37	0.38	24
				Oxidized Chromatium High-Potential Iron Protein (HIPIP)				
global	0.23	0.45	0.32	85	0.23	0.45	0.32	85
0-0.85	0.47	0.29	0.24	34	0.50	0.23	0.27	26
0.85-1.10	0.10	0.71	0.19	21	0.14	0.62	0.24	21
⩾1.10	0.07	0.43	0.50	30	0.11	0.49	0.40	35
				Tuna Cytochrome *c* (Oxidized)				
global	0.24	0.50	0.26	103	0.24	0.50	0.26	103
0-0.85	0.51	0.31	0.17	35	0.57	0.25	0.18	28
0.85-1.10	0.15	0.58	0.27	40	0.16	0.44	0.40	32
⩾1.10	0.04	0.60	0.36	28	0.09	0.70	0.21	43
				Calcium-Binding Parvalbumin				
global	0.28	0.45	0.27	108	0.28	0.45	0.27	108
0-0.85	0.62	0.26	0.12	34	0.79	0.07	0.14	28
0.85-1.10	0.20	0.51	0.29	41	0.17	0.45	0.38	29
⩾1.10	0.03	0.58	0.39	33	0.06	0.67	0.27	51
				Ribonuclease S				
global	0.23	0.57	0.20	124	0.23	0.57	0.20	124
0-0.75	0.48	0.39	0.13	31	0.50	0.32	0.18	28
0.75-0.95	0.23	0.57	0.20	30	0.28	0.53	0.19	32
0.95-1.15	0.18	0.63	0.19	27	0.17	0.74	0.09	23
⩾1.15	0.05	0.67	0.28	36	0.05	0.66	0.29	41
				Hen Egg White Lysozyme				
global	0.30	0.48	0.22	129	0.30	0.48	0.22	129
0-0.75	0.50	0.28	0.22	36	0.57	0.18	0.25	28
0.75-0.95	0.34	0.41	0.25	32	0.42	0.35	0.23	26
0.95-1.15	0.21	0.62	0.18	34	0.17	0.57	0.26	35
⩾1.15	0.11	0.67	0.22	27	0.15	0.70	0.15	40
				Flavodoxin (Oxidized)				
global	0.35	0.48	0.17	138	0.35	0.48	0.17	138
0-0.75	0.57	0.25	0.18	28	0.79	0.09	0.12	32
0.75-0.95	0.50	0.38	0.12	34	0.50	0.33	0.17	24
0.95-1.15	0.27	0.60	0.13	48	0.26	0.51	0.23	35
⩾1.15	0.11	0.61	0.28	28	0.04	0.81	0.15	46
				Staphylococcal Nuclease				
global	0.23	0.54	0.23	142	0.23	0.54	0.23	142
0-0.75	0.44	0.31	0.25	36	0.48	0.26	0.26	38
0.75-0.95	0.20	0.54	0.26	39	0.30	0.46	0.24	33
0.95-1.15	0.17	0.58	0.25	36	0.04	0.68	0.28	25
⩾1.15	0.10	0.74	0.16	31	0.09	0.74	0.17	46
				Sperm Whale Myoglobin				
global	0.29	0.43	0.28	153	0.29	0.43	0.28	153
0-0.75	0.50	0.28	0.22	36	0.68	0.09	0.23	35
0.75-0.95	0.42	0.34	0.24	38	0.43	0.30	0.27	33
0.95-1.15	0.15	0.50	0.35	40	0.12	0.56	0.32	34
⩾1.15	0.13	0.56	0.31	39	0.06	0.65	0.29	51

Table II (*Continued*)

spherical layers in units of R_g	C^α atoms				remote side-chain atoms[a]			
	hydrophobic	hydrophilic	neutral and ambivalent	no. of residues	hydrophobic	hydrophilic	neutral and ambivalent	no. of residues
				Group B				
				Adenylate Kinase				
global	0.29	0.52	0.19	194	0.29	0.52	0.19	194
0–0.50	0.43	0.36	0.21	14	0.36	0.43	0.21	14
0.50–0.65	0.29	0.21	0.50	14	0.50	0.18	0.32	22
0.65–0.75	0.44	0.44	0.12	16	0.33	0.33	0.33	12
0–0.75	0.39	0.34	0.27	44	0.42	0.29	0.29	48
0.75–0.95	0.35	0.47	0.18	45	0.43	0.37	0.20	40
0.95–1.15	0.33	0.50	0.17	54	0.25	0.61	0.14	51
≥1.15	0.12	0.74	0.14	51	0.13	0.74	0.13	54
				Papain				
global	0.27	0.43	0.30	212	0.27	0.43	0.30	212
0–0.50	0.40	0.35	0.25	20	0.42	0.35	0.23	17
0.50–0.65	0.50	0.21	0.29	14	0.44	0.33	0.23	18
0.65–0.75	0.50	0.44	0.06	16	0.58	0.33	0.09	12
0–0.75	0.46	0.34	0.20	50	0.47	0.34	0.19	47
0.75–0.95	0.30	0.42	0.28	57	0.39	0.31	0.29	51
0.95–1.15	0.17	0.45	0.38	58	0.17	0.41	0.41	46
≥1.15	0.15	0.53	0.32	47	0.10	0.60	0.29	68
				Concanavalin A				
global	0.28	0.52	0.20	237	0.28	0.52	0.20	237
0–0.50	0.57	0.14	0.29	14	0.70	0.05	0.25	20
0.50–0.65	0.50	0.19	0.31	26	0.52	0.24	0.24	25
0.65–0.75	0.44	0.44	0.12	18	0.47	0.42	0.11	19
0–0.75	0.50	0.26	0.24	58	0.56	0.24	0.20	64
0.75–0.95	0.31	0.53	0.16	70	0.34	0.48	0.18	50
0.95–1.15	0.16	0.58	0.26	61	0.16	0.59	0.25	51
≥1.15	0.10	0.75	0.15	48	0.07	0.75	0.18	72
				Chymotrypsinogen A				
global	0.34	0.45	0.21	245	0.34	0.45	0.21	245
0–0.50	0.35	0.30	0.35	20	0.36	0.29	0.35	17
0.50–0.65	0.53	0.29	0.18	17	0.52	0.24	0.24	21
0.65–0.75	0.68	0.21	0.11	19	0.69	0.19	0.12	16
0–0.75	0.52	0.27	0.21	56	0.52	0.24	0.24	54
0.75–0.95	0.40	0.30	0.30	47	0.45	0.25	0.29	51
0.95–1.15	0.26	0.58	0.17	66	0.36	0.48	0.16	44
≥1.15	0.19	0.61	0.19	57	0.12	0.71	0.16	73
				Subtilisin BPN'				
global	0.25	0.44	0.31	275	0.25	0.44	0.31	275
0–0.50	0.42	0.26	0.32	19	0.44	0.28	0.28	18
0.50–0.65	0.39	0.13	0.48	23	0.42	0.20	0.38	26
0.65–0.75	0.21	0.16	0.63	19	0.39	0.09	0.52	23
0–0.75	0.34	0.18	0.48	61	0.42	0.18	0.40	67
0.75–0.95	0.43	0.37	0.20	68	0.41	0.33	0.26	57
0.95–1.15	0.16	0.52	0.32	89	0.19	0.44	0.37	64
≥1.15	0.09	0.67	0.24	57	0.07	0.70	0.23	87
				Carboxypeptidase A				
global	0.29	0.48	0.23	307	0.29	0.48	0.23	307
0–0.50	0.52	0.33	0.14	21	0.48	0.39	0.13	23
0.50–0.65	0.50	0.41	0.09	22	0.58	0.32	0.10	19
0.65–0.75	0.42	0.21	0.37	24	0.46	0.23	0.31	26
0–0.75	0.48	0.31	0.21	67	0.50	0.31	0.19	68
0.75–0.95	0.38	0.44	0.18	71	0.44	0.35	0.21	57
0.95–1.15	0.21	0.47	0.32	95	0.20	0.51	0.29	83
≥1.15	0.12	0.69	0.19	74	0.11	0.66	0.23	98
				Thermolysin				
global	0.22	0.46	0.32	316	0.22	0.46	0.32	316
0–0.50	0.41	0.33	0.26	27	0.28	0.38	0.34	26
0.50–0.65	0.31	0.19	0.50	26	0.33	0.31	0.36	33
0.65–0.75	0.22	0.52	0.26	31	0.50	0.29	0.21	24
0–0.75	0.31	0.36	0.33	84	0.36	0.33	0.31	83
0.75–0.95	0.25	0.44	0.31	72	0.19	0.45	0.36	67
0.95–1.15	0.16	0.57	0.27	77	0.23	0.47	0.30	70
≥1.15	0.18	0.47	0.35	83	0.13	0.57	0.30	96

Table II (*Continued*)

spherical layers in units of R_g	C^α atoms				remote side-chain atoms[a]			
	hydrophobic	hydrophilic	neutral and ambivalent	no. of residues	hydrophobic	hydrophilic	neutral and ambivalent	no. of residues
				Lactate Dehydrogenase Apoenzyme M4				
global	0.37	0.45	0.19	329	0.37	0.45	0.19	329
0-0.50	0.47	0.37	0.16	38	0.50	0.29	0.21	34
0.50-0.65	0.42	0.26	0.32	47	0.44	0.28	0.28	39
0.65-0.75	0.65	0.26	0.09	23	0.56	0.33	0.11	36
0-0.75	0.49	0.30	0.21	108	0.50	0.30	0.20	109
0.75-0.95	0.38	0.48	0.15	82	0.44	0.38	0.18	82
0.95-1.15	0.28	0.53	0.19	81	0.28	0.55	0.17	58
≥1.15	0.24	0.57	0.19	58	0.19	0.64	0.18	80

[a] See Table II of ref 3 for identification of the side-chain atoms used in the computations. [b] For the smallest proteins in this group, there are only three layers indicated because of the small number of residues per layer, and the first layer was taken as 0-0.85 instead of 0-0.75.

they occur much less frequently (~0.29) than in the whole proteins (0.47). Their fraction increases gradually in the outer layers up to 0.66 for C^α and 0.70 for the remote side-chain atoms (in layer ≥1.3). The frequency of occurrence of the neutral–ambivalent group remains almost constant in the various layers and fluctuates only slightly around the value 0.24, the fraction of those residues in the whole proteins.

It can be seen that, inside the sphere of radius $0.75R_g$ (i.e., in concentric layers with radii $\leq 0.75R_g$), the fractions of the three types of residues fluctuate slightly (in a random way) around their fractions in the sphere of radius $0.75R_g$. The largest fluctuations (0.6) in these fractions occur for the C^α atoms of the hydrophilic and neutral groups. In this sense (i.e., that no radial preference in the distributions is detected), we may designate this spherical core as "homogeneous". In fact, in the next paper,[10] we show (as has already been demonstrated by KK[6] in their Figure 5) that a nonradial type of order (viz., the existence of hydrophobic and hydrophilic clusters) does occur in this sphere. It should be emphasized that, even though the hydrophobic residues occur more frequently in the inner sphere (radius $0.75R_g$) than in the outer layers, they still constitute less than 50% of the total number of residues in this inner sphere. Therefore, the common notion that the core is hydrophobic is not an adequate characterization.[11] It is also of interest to point out that the outer layer, ≥1.15, is not homogeneous, since the fraction of the hydrophobic residues decreases and the fraction of the hydrophilic residues increases in going from sublayer, 1.15–1.30, to ≥1.30.

The above conclusions are based on the average behavior of residues in the sample of the larger proteins. It is of interest to examine whether these conclusions also hold for *each* of the proteins of group B and to check the deviations from the average behavior. Results for the individual proteins are given in Table II. We first examine the compositions of the whole proteins with respect to the three types of amino acids (in the lines denoted "global"); our primary interest in these data lies in the fractions of the hydrophobic amino acids, which are thought to play an essential role in protein folding. It should be noted that the fractions of the hydrophilic residues for each of the proteins of group B do not deviate much from 0.47, the average value which appears in Table I. These values fall in the range between 0.52 (found for adenylate kinase and concanavalin A) and 0.43 (papain). The deviations of the fractions of hydrophobic residues from their average occurrence (0.29; Table I) are larger. The highest value (0.37) occurs for lactate dehydrogenase; the lowest (0.22, found for thermolysin) still provides sufficient hydrophobic interactions for the stabilization of this protein structure. These fluctuations in fractions of hydrophobic residues are compensated mainly by the corresponding low and high fractions of the group of neutral and ambivalent residues (0.19 and 0.32, respectively).

The main conclusions drawn from Table I hold also for each of the proteins of group B; i.e., the fraction of the hydrophobic residues is the highest in the inner sphere of radius $0.75R_g$ and decreases toward the outside (small deviations from this behavior are detected for adenylate kinase, subtilisin BPN′, and thermolysin). The hydrophilic group behaves in an opposite manner, except for the outermost layer in thermolysin, and the fraction of the neutral and ambivalent residues in the various layers does not change in any particular direction but rather fluctuates randomly (sometimes substantially) around the fraction in the sequence. It should be noted that the concentration of hydrophobic residues in the inner sphere of radius $0.75R_g$, even in the extreme case (concanavalin A), does not exceed the value 0.56 (for side-chain atoms); for some of the proteins it is much smaller (0.31 for C^α in thermolysin). This means that the common notion that the core is hydrophobic is also not an adequate characterization for each of the proteins of group B. On the other hand, the outer layers of these proteins (except for thermolysin) are highly hydrophilic. The fraction of the hydrophilic residues in the sphere of radius $0.75R_g$ is around 0.30 and is substantially lower (0.18, for both C^α and side-chain atoms) only for subtilisin BPN′. It is important to point out that the fractions of the three classes of amino acids in the sublayers of the inner sphere of radius $0.75R_g$, in most cases, fluctuate randomly around their fractions in this sphere. These fluctuations are substantially higher than those detected for group B in Table I; this behavior stems from the much smaller data base provided by the individual proteins. A gradual radial decrease and increase (from the center of mass outward) in the fractions of the hydrophobic and hydrophilic residues, respectively, is detected only for concanavalin A and for the C^α atoms in thermolysin. The homogeneity (i.e., absence of radial preferences in the distributions) discussed previously, therefore, exists for most of the proteins of group B. Obviously, this homogeneity can be accounted for in terms of hydrophobic interactions: when a nonpolar residue becomes buried, its hydrophobic free energy has decreased, and there is no additional driving force to move it *further* toward the center.[12] Therefore, the distributions of the hydrophobic, and, hence, of the hydrophilic, residues in the inner buried region of the protein (region I; see below) should be homogeneous. Indeed, we show in a later section that the buried region of the proteins studied here is approximately a sphere of radius R_g around the center of mass. The inhomogeneities found in concanavalin A and

thermolysin may be due to departure from spherical shape (in this case, the sphere of radius $0.75R_g$ may be partially exposed).

Results for the Smaller Proteins. We now examine the results in Table I for the smaller proteins (group A). Qualitatively, the three types of amino acids show almost the same behavior as in group B, but quantitatively there are some differences. The fraction of the hydrophobic residues in the inner spheres (radius $0.75R_g$) is larger in group A than in group B (0.51 vs. 0.43 for C^α; 0.64 vs. 0.47 for the side-chain atoms) and their fraction in the outermost layer of group A is slightly lower than in group B (0.10 vs. 0.15 for C^α for layer ≥ 1.15). For the hydrophilic residues, the fractions of the C^α and side-chain atoms, respectively, in layer ≥ 1.15 are almost the same in the two samples of proteins; there is discordance, however, for the side-chain atoms in the layer 0.95–1.15 and more so in the inner sphere (radius $0.75R_g$), where the value for group A is significantly lower than that for group B (0.17 vs. 0.28, respectively). It should be noted that the fractions of the three types of amino acids are almost the same in groups A and B (compare the data in the lines denoted "global" in Table I). This means that the differences found above for the fractions of hydrophobic residues in groups A and B (and differences in the fractions of the neutral–ambivalent residues which will be discussed later) stem from differences in the radial distributions of residues for smaller and larger proteins. This has already been pointed out by KK.[6]

These differences in the fractions of the hydrophobic residues for groups A and B stem from geometric and energetic effects.[3] In the smaller proteins (group A), a larger fraction of the residues is accessible to the surrounding water than in group B. Therefore, the hydrophobic residues of group A bury themselves in the interior more strongly than those of group B in order to avoid unfavorable contacts with water. For the same reason, the hydrophobic side chains orient toward the center of mass more strongly in group A than in group B (see ref 3). This tendency of the hydrophobic side chains highly increases their fraction in the inner sphere of group A, since the sizes of the side chains are relatively large as compared to R_g in these proteins. On the other hand, no significant difference in behavior of the hydrophilic amino acids in the two samples of proteins was detected in ref 3. The relatively low fraction of hydrophilic side-chain atoms in the inner sphere of group A (as compared to the fractions of C^α and the corresponding fractions of group B) is a result of the tendency of these side chains to orient toward the outside of the proteins[3,4] and, again, the larger ratio between the size of the side chain and R_g. It should be noted that the neutral and ambivalent residues occur less frequently in the inner sphere and more frequently toward the outer layer (for C^α these fractions change gradually from 0.20 to 0.36, respectively). This is because Ala and Gly behave as hydrophilic amino acids in the smaller proteins (see ref 3), i.e., are concentrated more in the periphery than in the interior. Since these two amino acids constitute the majority of residues in the neutral–ambivalent group, their individual distribution also affects the distribution of the entire group significantly.

Table I reveals that the homogeneity found in the inner sphere (radius $0.75R_g$) for group B exists also for the C^α atoms of the smaller proteins. For the side-chain atoms, on the other hand, the fractions of the three groups of amino acids change systematically in the three concentric layers of radii $\leq 0.75R_g$. The frequencies of occurrence of the hydrophobic residues show the largest inhomogeneity, decreasing from 0.72 to 0.52 in going from the innermost sphere to the outer layer ($0.65–0.75R_g$). The opposite behavior is detected for the hydrophilic and for the neutral–ambivalent group. This inhomogeneity probably arises from the fact that, in the smaller proteins, the sphere of radius $0.75R_g$ is partially exposed whereas, in the larger ones, it is buried (see the discussion in the section Comparison with Other Methods and in footnote 13).

We next examine the behavior of the individual proteins of group A (Table II) and concentrate first on the fractions of the three types of amino acids in the whole proteins. The fractions for the hydrophobic residues fall in the range 0.23–0.35 (the average in group A is 0.28) and those for the neutral–ambivalent residues in the range 0.17–0.32 (the average is 0.24). These values are comparable to the corresponding ranges found for group B. The hydrophilic residues, on the other hand, fluctuate much more here than in group B, from 0.40 to 0.57 (the average is 0.48).

Because of the small number of residues per layer, the results for the six smallest proteins are given only for three layers. The results for cytochrome *c* and the six proteins following it in Table II lead to the same picture obtained from Table I: the hydrophobic residues occur with high frequency in the inner sphere and the frequency decreases gradually toward the outside; the hydrophilic residues show an opposite behavior. The residues of the neutral–ambivalent group for most of the proteins (HIPIP, cytochrome *c*; parvalbumin; ribonuclease S; and myoglobin) are concentrated more in the periphery than in the interior as was found in Table I for the average behavior. It should be recalled that the fractions of hydrophobic residues in the inner spheres of the proteins of group A are generally larger than those of group B (as is also reflected in the results of Table I). These fractions are especially high in calcium-binding parvalbumin and flavodoxin (0.79 for the side-chain atoms in the innermost layer of both proteins) and, for these proteins, one can identify a "hydrophobic core". Deviations from the general distribution of the hydrophobic and hydrophilic residues are detected only in some of the smallest proteins of group A. For example, in HIPIP, the fraction of hydrophilic residues in the second layer is much higher than in the third one (for C^α, 0.71 vs. 0.43, respectively). In the inner sphere of BPTI, there is a lower fraction of C^α atoms of hydrophobic residues than in the second layer (0.24 vs. 0.57, respectively). This is compensated by the higher fraction of hydrophilic C^α atoms in the inner layer than in the second one (0.52 vs. 0.29, respectively). This behavior of these small proteins may also possibly be due to their deviations from spherical shape.

Comparison with the Results of Krigbaum and Komoriya.[6,7] For each of the side chains, Krigbaum and Komoriya[6] determined an interaction parameter ξ, which is related to polarity and is argued to originate from van der Waals (VW) forces rather than from hydrophobic interactions, hydrogen bonding, etc. They also divided the amino acids into three groups, nonpolar, indifferent, and polar, and calculated their fractions in spherical layers around the centers of mass of larger and smaller proteins, viz., those with >300 and <100 residues, respectively (see their Figure 11). Although our grouping (see footnote 32 of ref 3) differs from theirs (especially for the indifferent and polar amino acids—cf. the solid lines in column VI with the dashed lines of column II of Table IV of ref 3), the main trends of the results are similar. They also computed the average value of ξ, i.e., $\langle\xi\rangle$, for these same layers, and these results are described in Figure 4 of their paper. The following are their main conclusions: (1) For

the entire sample of 23 proteins, as well as for the larger and smaller ones, there is a gradual increase in $\langle\xi\rangle$ (increase in polarity) in going from the interior to the exterior parts of the proteins. Our results in Table I, where the fractions of hydrophobic and hydrophilic residues decreased and increased, respectively, in going from the innermost layer toward the outside (with essentially no change in the fraction of neutral–ambivalent residues) are in accord with theirs. (2) In the innermost layers, $\langle\xi\rangle$ is substantially lower in the smaller proteins than in the larger ones. Our results for groups A and B are in accord with this conclusion. (3) In the outer layers, $\langle\xi\rangle$ is much higher in the smaller proteins than in the larger ones. We did not observe significant differences in the results for the side-chain atoms between groups A and B, but our results for the C^α atoms for the two groups differed in the same direction as those for $\langle\xi\rangle$. (4) According to their Figure 4, $\langle\xi\rangle$ increases continuously, even in the innermost layers of the larger proteins; therefore, they found no radial homogeneity in this region, in contrast to our results. In this context, we judge the results in their Figure 11 for the three or four innermost layers (which correspond approximately to our innermost sphere of radius $0.75R_g$) as indicating homogeneity; i.e., they do not show any clear trend.

The difference cited above in point 4 is a significant one. The supposed absence of radial homogeneity is cited by KK as evidence that the structure in this region is determined by VW, rather than by hydrophobic, interactions. They argued that hydrophobic interactions should lead to a fairly isotropic composition of amino acids in the core. According to their view, the creation of clusters of polar and nonpolar residues is also due mainly to VW interactions rather than to hydrogen bonding, electrostatic interactions, etc. They also attributed the lower polarity in the core of the smaller proteins compared to the larger ones to the need for stabilization by means of favorable VW interactions: "Since the major contribution to stabilization arises from interactions within the core, and smaller molecules have proportionally fewer core residues, this must be offset by substituting residues of low polarity for those of medium polarity (e.g., Ile for Ala) in the core of smaller molecules". The basis for this argument is best summarized by the following quotation from the KK paper:[6] "hydrophobic interactions can only be effective during the initial phase of the refolding process. Once a nonpolar side chain becomes buried, the hydrophobic interaction criterion is fully satisfied and there is no driving force for further organization of the sheath–core structure. Such a buried side chain still interacts with its near-neighbors *through van der Waals forces*". As an argument against this view, we cite (1) our finding that the inner core is homogeneous, (2) the view of hydrophobic bonding expressed in footnote 4 of ref 3 (and the sentence in the Introduction where reference is made to this footnote). [We attribute the lower polarity in the core of the smaller proteins compared to the larger ones to both the hydrophobic interaction and geometric effects (i.e., larger surface-to-volume ratio, and larger side-chain-to-R_g ratio, in the smaller proteins compared to the larger ones).], and (3) that the radial gradation of the fractions of hydrophobic and hydrophilic residues outside the sphere of radius $0.75R_g$ (as well as their aggregation into clusters) can be accounted for as arising mainly from interresidue hydrogen bonding (to compensate for loss of hydrogen bonding to water in the denatured form); a detailed discussion of the roles of hydrogen bonding and hydrophobic interactions in cluster formation is given in ref 10.

Comparison with Other Methods. It is of interest to compare the results obtained by our method to those for the fractions of residues buried in the interior of the proteins obtained by Chothia[8] and by Wertz and Scheraga[9] (WS), using other methods. Such a comparison can provide more information about the structures of proteins and also will constitute a test of the adequacy of our description of the proteins studied in terms of spherical layers. As was pointed out in the Introduction, the use of spherical layers is physically significant only for proteins which do not deviate much from spherical shape. In this case, the inner spherical layers contain the buried residues and the outer layers the exposed ones. The methods used by Chothia and by WS actually define the protein surface (which is not defined in our method); therefore, agreement between their results and ours would indicate that the proteins are globular (i.e., close to spherical in shape).

According to Chothia, a residue is considered buried if less than 5% of its surface area is accessible to water (he calculated the accessible area by the method of Lee and Richards[14]). This criterion is much more stringent than that used by WS, which required "that most of the atoms of the residue have little or no contact with water for the residue to be considered inside". This is the main reason the results obtained by the two methods[8,9] are significantly different. For example, according to Chothia (Table II of ref 8), the fraction of residues exposed to water is 0.72, whereas WS obtained only 0.46.[13,15] The fact that different samples of proteins were used in these calculations (Chothia used 12 proteins and WS used 20) does not seem to affect the results much since the two samples consist of both smaller and larger proteins. In the following discussion, we therefore assume that each of the two methods would lead to the same results for both samples and also for the sample of 19 proteins that we used.

Combination of the methods of Chothia and WS enables us to define three layers: an outer layer which contains the residues exposed to solvent, according to the criterion of WS (0.46 of the total number of residues), an inner layer consisting of the residues defined as buried by Chothia's criterion (0.28), and an intermediate layer consisting of the rest of the residues (0.26). It should be noted that these layers are not necessarily spherical; their shapes are determined by the shapes of the proteins. Using the classification of reference 3 and information about the number of occurrences of each amino acid in the samples and the fractions of residues buried,[8,9] we calculated for each layer the fractions of hydrophobic, hydrophilic, and neutral–ambivalent residues, and the results are presented in the last three lines of Table III, where these layers are denoted III, I, and II, respectively. Relying on recent work of Schrier and Schrier,[15] we adopt the criterion of WS for a residue to be buried. This means that the buried region consists of both layers I and II and the exposed part of layer III. Alternatively, we used our entire sample of 19 proteins and also calculated the fractions of the three types of amino acids in three *spherical* layers (0–0.80, 0.80–1.0, and >1.0 in units of R_g) chosen to contain, as closely as possible, the same fractions of the total number of residues as appear in layers I, II, and III, respectively (compare the percent occurrences for layers I, II, and III with those for the first three spherical layers in columns 5 and 9 of Table III). These results are shown in lines 2–4 of Table III. The following conclusions can be drawn from the table.

(1) According to the criterion of WS, the majority of the hydrophobic residues are buried (i.e., appear in layers I and II), whereas the majority of the hydrophilic residues are exposed to the solvent (appear in layer III). On the

Table III
Average Fractions of the Hydrophobic, Hydrophilic, and Neutral-Ambivalent Residues in Spherical Layers for a Sample of 19 Proteins[a] and for Layers (of Unspecified Shape) Obtained from the Results of Chothia[b] and Wertz and Scheraga[c]

	C^α atoms				typical side-chain atoms			
layers	hydrophobic	hydrophilic	neutral and ambivalent	%[d]	hydrophobic	hydrophilic	neutral and ambivalent	%[d]
			This Work (19 Proteins)					
global	0.29	0.47	0.24		0.29	0.47	0.24	
0-0.80[e]	0.45	0.31	0.24	30	0.50	0.25	0.25	28
0.80-1.0[e]	0.30	0.46	0.24	26	0.33	0.42	0.25	23
≥ 1.0[e]	0.16	0.59	0.25	44	0.13	0.63	0.25	49
≥ 1.3[e]	0.09	0.63	0.28	21	0.09	0.70	0.21	21
			Results Based on Ref 8 and 9[f]					
buried I	0.50	0.24	0.26	28	0.50	0.24	0.26	28
buried II	0.32	0.48	0.20	26	0.32	0.48	0.20	26
exposed III	0.13	0.62	0.25	46	0.13	0.62	0.25	46

[a] This work. [b] Reference 8. [c] Reference 9. [d] These are the ratios (in percent) of the numbers of residues in the layers to the total number of residues in the sample. [e] In units of R_g. [f] The last four columns are identical with the previous four because ref 8 and 9 did not distinguish between C^α and side-chain atoms.

other hand, Chothia's more stringent criterion for a residue to be buried leads to approximately equal numbers of hydrophobic residues in the interior (layer I) as in the exterior (layers II and III).

(2) The frequency of occurrence of the hydrophobic residues is highest in layer I and decreases gradually toward the outer layer III. An opposite behavior is shown by the hydrophilic residues. The fractions of the neutral–ambivalent residues fluctuate slightly around their fraction in our entire sample of 19 proteins. The criterion of WS (which we adopt here) implies that layer II, even though buried, is much more hydrophilic and less hydrophobic than the core, i.e., layer I.

(3) Our results (especially for the typical side-chain atoms) are close to those based on the works of Chothia[8] and WS[9] (compare lines 2–4 for the side-chain atoms with the last three lines). The largest discrepancy (0.07) occurs for the fractions of the hydrophilic residues (C^α) in the innermost layers. This agreement between the two sets of results indicates that the proteins studied do not deviate much from spherical shape, and therefore the use of spherical layers is justified. In fact, for our sample, the sphere of radius R_g around the center of mass coincides approximately with the unexposed parts of the proteins (layers I and II). This result pertains more to the larger proteins (group B) of our sample, since they contribute more residues than the smaller ones. Obviously, this result in not a general one, in that it does not apply to proteins that are either smaller or larger than those in group B because, as R_g increases, the radius of the buried sphere (in units of R_g) also increases.

(4) Table III also shows that layer I corresponds approximately to the sphere of radius $0.75R_g$, which was found to be homogeneous for the larger proteins. We also provide in Table III our results for the exterior spherical layer ≥ 1.3. In this layer, the frequency of occurrence of the hydrophobic residues is lower and that of the hydrophilic residues is higher than their fractions in layer III. This means that the exposed layer III is not homogeneous, as we have pointed out previously. It has to be emphasized that the conclusions deduced in point 3 are based on a comparison between averages taken over a large sample of proteins. Therefore, it still can happen, for certain proteins of the sample (whose shapes deviate significantly from spherical), that the buried and exposed regions cannot be expressed adequately in terms of spherical layers.

Conclusions

It is of interest to point out that some of the proteins in the sample studied here are composed of relatively low fractions of hydrophobic amino acids, e.g., the low value of 0.22 for thermolysin compared to the highest value of 0.37 for lactate dehydrogenase. It is reasonable to expect that this low content of hydrophobic residues is compensated to some extent by an efficient arrangement of the amino acid residues in the protein which allows them to interact with each other to stabilize the structure by other interactions, such as disulfide bonds, electrostatic interactions, etc.

The distributions of hydrophobic and hydrophilic residues deviate considerably from a random one. For example, for the group of larger proteins, the inner sphere of radius $0.75R_g$ contains ~50% more hydrophobic residues than their fraction in these proteins. For the smaller proteins, this deviation appears to be much higher, ~75% for C^α and over 100% for the side-chain atoms (in the inner layer). The preference of the hydrophobic residues to be located in the interior and that of the hydrophilic residues to be in the exterior part of the protein constitutes a strong "rule", which has been found to hold not only on the average but also for most of the individual proteins studied. The commonly accepted dogma is thus supported, viz., that water plays an important role in protein organization and should be taken into account in any protein-folding algorithm.

The radii of the spherical layers are given in units of R_g. Therefore, the radial distributions are not independent of protein size. For example, in the group of larger proteins, we found a sphere of radius $0.75R_g$ to be homogeneous. In even larger globular proteins, one would expect to find a homogeneous sphere with a larger radius because of the smaller surface-to-volume ratio.

Summary

We have examined the fractions of hydrophobic, hydrophilic, and the group of neutral and ambivalent amino acids in spherical layers around the centers of mass of 19 proteins. The main conclusions are the following. (1) In the inner sphere of radius $0.75R_g$ and in the outermost layer (from $1.3R_g$ outward), the fractions of hydrophobic residues are substantially higher and lower, respectively, than in the whole protein. In the intermediate layers, there is a gradual decrease in the fraction of hydrophobic residues toward the outside. The hydrophilic residues show an opposite behavior. Significant deviations from this rule were detected only for the small proteins BPTI and HIPIP. (2) The average fraction of the hydrophobic residues in the inner sphere of the smaller proteins is significantly

higher than in the larger ones. This fraction was found to be especially high (above 0.6) in calcium-binding parvalbumin and flavodoxin. (3) In the larger proteins, the sphere of $0.75R_g$ was found to be approximately homogeneous in the sense that the fractions of the three groups of amino acids do not change significantly in concentric layers of radii $\leq 0.75R_g$. (4) In the larger proteins, the fractions of the neutral–ambivalent amino acids do not change much in the various layers. They fluctuate around their fraction in the whole protein. In the smaller proteins, their fraction is low in the inner sphere and increases gradually toward the outside. This is due to Gly and Ala which behave like hydrophilic amino acids in the smaller proteins; i.e., they are concentrated more in the periphery than in the interior. (5) Our method gives only the radial distribution of the three types of amino acids, without specifying the "boundaries" between regions; hence, it does not provide information about the exposed and buried regions of the proteins. In order to obtain information about these distributions with respect to the exposed and buried regions of the proteins, we combined the results of Chothia and WS, assumed the classification of amino acids according to ref 3, and adopted the less stringent criterion of WS for a residue to be buried.[15] The conclusion drawn was that the distributions of the hydrophobic and hydrophilic amino acids are inhomogeneous, even in the buried region. Three concentric layers (not necessarily spherical), I, II, and III, have been defined where the outermost layer (III) is exposed and the innermost (I) and the intermediate (II) layers are buried. The fraction of the hydrophobic residues is highest in layer I and decreases gradually toward layer III. The opposite behavior occurs for the hydrophilic residues, and fractions of the neutral–ambivalent residues fluctuate slightly around their fraction in our sample of 19 proteins. We also used our sample to calculate the fractions of the three types of amino acids in three *spherical* layers chosen to contain as closely as possible the same fractions of the total number of residues as appear in layers I, II, and III, respectively. The close agreement in the results for the fractions of the three types of amino acids in the two sets of layers indicates that the proteins studied do not deviate much from spherical shape and, therefore, our method, which uses spherical layers, is adequate. Comparison between the two sets of layers showed that, for the sample of proteins studied, the buried region is approximately a sphere of radius R_g around the center of mass of the protein.

Acknowledgment. We thank Dr. George Némethy for helpful discussions.

References and Notes

(1) (a) This work was supported by research grants from the National Institute of General Medical Sciences of the National Institutes of Health, U.S. Public Health Service (GM-14312), and from the National Science Foundation (PCM79-20279). (b) To whom requests for reprints should be addressed.
(2) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1.
(3) Meirovitch, H.; Rackovsky, S.; Scheraga, H. A. *Macromolecules* **1980**, *13*, 1398.
(4) Rackovsky, S.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5248.
(5) Krigbaum, W. R.; Rubin, B. H. *Biochim. Biophys. Acta* **1971**, *229*, 368.
(6) Krigbaum, W. R.; Komoriya, A. *Biochim. Biophys. Acta* **1979**, *576*, 204.
(7) After completion of the manuscript, we became aware of the paper by Krigbaum and Komoriya[6] which treated the same problem with a similar technique. We have, therefore, revised this paper to reflect this fact.
(8) Chothia, C. *J. Mol. Biol.* **1976**, *105*, 1.
(9) Wertz, D. H; Scheraga, H. A. *Macromolecules* **1978**, *11*, 9.
(10) Meirovitch, H.; Scheraga, H. A. *Macromolecules*, submitted.
(11) Some proteins contain water in internal crevices, presumably bound to internal polar groups. Our computed distributions reflect the occurrence of internal polar groups without regard to whether or not water is bound to them.
(12) See footnote 4 of ref 3.
(13) It should be pointed out that the results obtained by Chothia and by WS are not of general validity in that they apply only to proteins in the size range of their data set and would not apply to a larger protein. This is because the surface-to-volume ratio of a globular protein decreases with increasing size—hence most of the residues of a very large protein will be buried.
(14) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379.
(15) For bovine pancreatic ribonuclease A, in particular, Schrier and Schrier (personal communication) calculated 0.70 and 0.42 by the methods of Chothia and WS, respectively. Schrier and Schrier also measured the heats of transfer of (a) ribonuclease A and (b) its constituent amino acids from water to 10% ethanol and interpreted the ratio of these heats (0.48 ± 0.06) as the fraction of residues exposed in native ribonuclease. This value is close to the value of 0.42, computed by the method of WS.