PROTEINS: Structure, Function, and Genetics 9:37–55 (1991)

Distribution and Complementarity of Hydropathy in Multisubunit Proteins

Alex P. Korn and Roger M. Burnett
The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104

ABSTRACT **A survey of 40 multisubunit proteins and 2 protein–protein complexes was performed to assay quantitatively the distribution of hydropathy among the exterior surface, interior, contact surface, and noncontact exterior surface of the isolated subunits. We suggest a useful way to present this distribution by using a "hydropathy level diagram." Additionally, we have devised a function called "hydropathy complementarity" to quantitate the degree to which interacting surfaces have matching hydropathy distributions. Our survey revealed the following patterns: (1) The difference in hydropathy between the interior and exterior of subunits is a fairly invariant quantity. (2) On average, the hydropathy of the contact surface is higher than that of the exterior surface, but is not greater than that of the protein as a whole. There was variation, however, among the proteins. In some instances, the contact surface was more hydrophilic than the noncontact exterior, and in a few cases the contact surface was as hydrophobic as the protein interior. (3) The average interface manifests significant hydropathy complementarity, signifying that proteins interact by placing hydrophobic centers of one surface against hydrophobic centers of the other surface, and by similarly matching hydrophilic centers. As a measure of recognition and specificity, hydropathy complementarity could be a useful tool for predicting correct docking of interacting proteins. We suggest that high hydropathy complementarity is associated with static inflexible interactions. (4) We have found that some subunits that bind predominantly through hydrophilic forces, such as hydrogen bonds, ionic pairs, and water and metal bridges, are involved in dynamic quaternary organization and allostery.**

Key words: protein–protein interactions, intersubunit binding, hydropathy, hydropathy complementarity, protein interfaces

INTRODUCTION

Our investigation of the distribution of hydrophobic and hydrophilic amino acids in proteins, and of the role of the hydrophobic force in determining and maintaining protein–protein interactions, was prompted by our desire to solve a particular problem in our laboratory. Having solved the crystallographic structure for hexon,[1] the trimeric protein which is the major component of the adenovirus capsid, we wanted to understand how its polypeptide chain is adapted to the various types of interprotein interactions responsible for the capsid's architecture. However, the current level of understanding of the general principles of protein–protein interaction was insufficient to answer some specific questions. We could not say, for example, that the hexon subunits were bound to each other optimally in terms of overall hydrophobic interactions. We also wanted to know if hydrophilicity-based forces, arising from hydrogen bonds, electrostatic interactions, and dipole–dipole pairing, cooperate with and are optimally balanced with hydrophobicity-based forces. Clearly, these are issues relevant not only to hexon but also to the general problem of protein–protein interactions. It was suggested recently[2] that drugs designed to interfere with the dimerization of the HIV protease could provide anti-AIDS therapy. Clearly, understanding the nature of the protease interface will be a prerequisite for drug design. Neither the methodology necessary for answering these questions nor a comprehensive survey examining these properties in other oligomeric systems was available at the outset of this investigation.

We propose a method that is readily conceptualized, easily derived, and easily applied. It obtains quantitative evaluations of the hydropathies of interacting and noninteracting surfaces in oligomeric proteins and complexes. By evaluating quantitative differences in hydropathy between interacting and noninteracting exterior surfaces, we hope to reveal any special nature of the binding surface. In addition, we have calculated the hydropathy difference between the exterior and interior of the subunits of oligomeric proteins and suggest that this quantity may be characteristic of optimally folded structures. We have devised a novel way, using a "hydropathy

Received September 21, 1989; revision accepted June 6, 1990.
Address reprint requests to Dr. Roger M. Burnett, The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104.

level diagram," to present the hydropathy values of the component parts of an interacting protein thereby making possible a facile characterization of the protein in terms of how its hydropathy is distributed. Further, we have devised a quantitative measure of the extent of cooperation between hydrophobic and hydrophilic binding in interacting surfaces using a function we call "hydropathy complementarity." What follows are our findings from a survey of 40 multisubunit proteins and 2 protein–protein complexes whose structures were obtained from the April 1988 Brookhaven Protein Data Bank.[3] During the course of this work other investigators[4,5] have also sought to make generalizations concerning protein interfaces, and these have been valuable in evaluating our own results.

METHODS AND DEFINITIONS

Choosing Examples From the Brookhaven Data Bank

Most entries in the April 1988 Brookhaven Data Bank[3] for oligomeric proteins or protein complexes were used for this survey. There were 19 homodimers, 7 heterodimers, 2 homotrimers, 6 homotetramers, 6 heterotetramers, 1 heterohexamer, and 1 homooctamer. Among the heterodimers were the only nonmultisubunit proteins in the survey: the complexes of a Fc fragment with protein A fragment (Brookhaven Data Bank code: 1FC2) and a lysozyme molecule with a Fab fragment (2HFL). For some Brookhaven entries that supplied only part of the complex it was necessary to generate the missing subunits by coordinate transformations on the supplied subunits. The resulting completed oligomer was checked visually on the graphics screen.

Some structures in the Data Bank were omitted in order to avoid oversampling with related molecules. For example, only human hemoglobins were analyzed: deoxyhemoglobin (1HHB), oxyhemoglobin (1HHO), fetal deoxyhemoglobin (1FDH), and sickle cell deoxyhemoglobin (1HBS). The following immunoglobulins, representing interfaces between light and heavy chains, were chosen: the Fab fragment of human myeloma IgG (3FAB); Fab fragment of murine, galactan-binding IgG (1FBJ); intact human Kol immunoglobulin (1IG2); the Fab fragment of phosphocholine binding murine myeloma immunoglobulin (1MCP); and a human Fc fragment (1FC1). However, human myeloma (Kol) Fab fragment (1FB4) was not included because it is part of the whole IgG structure (1IG2), which is already in the list. Other structures offered additional varieties of immunologically relevant interfaces: the two homodimers of the Bence–Jones variable region fragments (1REI and 2RHE), the IgG–lysozyme complex (2HFL), and the complex of protein A fragment with the Fc fragment (1FC2).

The highest resolution structure was used when there were several entries for the same, or similar, molecule. Data Bank entries with explicit warnings in the "REMARKS" section about the reliability of the amino acid sequences were rejected. Insulin, glucagon, and mellitin were not included because their polypeptide chains are too small to define proper exterior and interior atoms.

Viral capsid proteins were taken as presented in the Data Bank. If the repeat unit was a trimer (the majority of capsids have three subunits in the icosahedral asymmetric unit), the trimeric complex was analyzed but not the trimer–trimer contacts. The capsids of the human viruses, mengovirus (1MEV) and rhinovirus (4RHV), are formed from a repeating heterotetramer of VP1, VP2, VP3, and VP4. The VP4 molecules were entered only for their contact surfaces because their nonglobular, extended chain structures do not allow the definition of exterior and interior subsets.

Defining the Subsets of a Protein Subunit: Exterior Surface, Interior Bulk, Contact Surface, and Noncontact Exterior Surface

Each subunit of a multisubunit protein was divided up into subsets, each of which comprised a number of individual whole atoms (Fig. 1). The subset comprising the exterior surface of a subunit was defined as the set of atoms accessible to a probe sphere representing a solvent water molecule. This was implemented using Connolly's program MS[6,7] adapted for us by Dr. Kim Sharp. After a systematic investigation of the effects of the program's parameters, they were fixed at their optimum values: the probe sphere's radius was set at 1.7 Å and the dot density at 14 $Å^{-2}$. Solvent accessibility determinations were performed after ligands such as NAD, heme, saccharides, and metals were removed. The interior subset is the set of atoms remaining after the solvent-accessible atoms have been removed. The contact surface was defined as all the subunit's atoms positioned within 4.5 Å of any atom, or atoms, belonging to (a) neighboring subunit(s). The noncontact exterior surface is the set of atoms of the exterior solvent-accessible set that was not in contact with any other subunit(s).

Calculating Hydropathies of the Subsets

In the absence of an accurate atom-based hydropathy scale (see Discussion), we used Eisenberg's normalized "consensus" hydropathy scale[8] for the amino acids. This was constructed by merging four scales proposed by other authors. The hydropathy of a whole subunit is simply the sum of the hydropathies of the amino acids in the polypeptide chain divided by the number of amino acids. The subsets, however, consist of lists of individual atoms, and hence fractional amino acids, while the hydropathy scale is based on whole amino acids. Thus, Φ_{set}, the hydropathy of any one of these atom sets, is expressed as

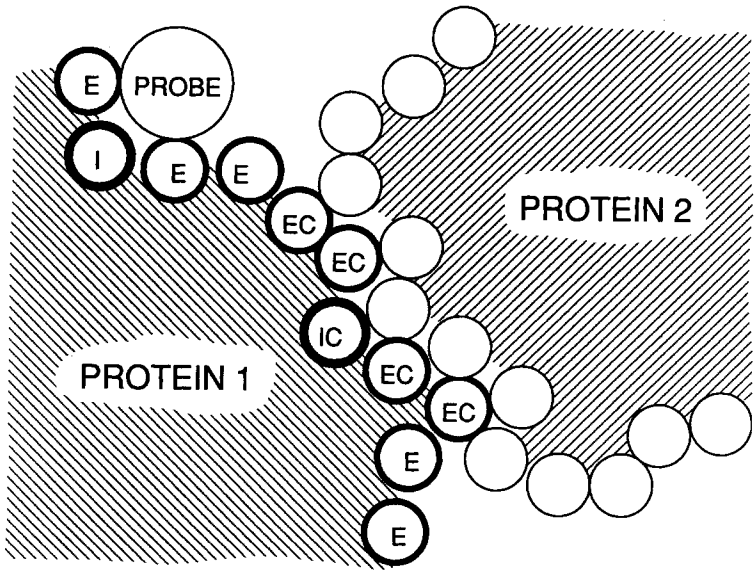

Fig. 1. Schematic diagram illustrating the various categories of atoms in a protein. These are solvent accessible atoms (E), interior atoms (I), contact atoms belonging to the external, solvent accessible subset (EC), and contact atoms belonging to the internal subset (IC). The probe sphere radius cannot be smaller than 1.7 Å because spheres of smaller radii can penetrate to the deep interior of the protein. As a result, atoms of protein 2 can penetrate more deeply into crevices on protein 1's surface than the probe sphere representing a solvent water molecule.

$$\Phi_{set} = \frac{\sum_A f_A \cdot \Phi_A}{\sum_A f_A} \quad (1)$$

The summations are over each amino acid, A, in the set. The term f_A is the fraction of the amino acid, A, present in the set, and Φ_A is the hydropathy of the whole amino acid. Equation (1) is equivalent to the mean of the hydropathies of the fractional amino acids in the set. It is implicit in this formula that each amino acid's hydropathy is apportioned equally to its constituent atoms.

It is more convenient to perform the summation over atoms rather than amino acids. The term, f_A, is replaced by $n_{a,A}/n_A$ where $n_{a,A}$ is the number of atoms in the fractional amino acid, A; and n_A is the total number of atoms in the whole amino acid, A. Additionally, the atom-based hydropathy, ϕ_a, which is equal to Φ_A/n_A, is used. Summing over the atoms, a, Eq. (1) becomes

$$\Phi_{set} = \frac{\sum_a \phi_a}{\sum_a 1/n_A} \quad (2)$$

The denominator normalizes the tally of Φ_{set} in terms of the amino acid hydropathy scale. Had we desired to retain it in the atom-based scale, a straight averaging of ϕ_a would have sufficed. Since the total hydropathy values are expressed in the amino acid-based scale, we could compare the hydropathy of a subset to those for the whole individual amino acids.

Negatively charged oxygens at the C-terminus (designated by "OXT" in Brookhaven nomenclature) of polypeptide chains were not included in the tallies. Also excluded from the tally were modified or nonstandard amino acids for which no hydropathy value was available.

Averaging Hydropathies

After subset hydropathies, Φ_{set}, were calculated for each example in our survey of the Data Bank, averages for each of the subsets over the entire survey of 40 multisubunit and 2 protein–protein complexes were calculated. Both unweighted and weighted averages for each Φ_{set} were obtained. The unweighted average is a straight averaging of each chemically unique subunit. The weighted average, on the other hand, weights each hydropathy both in proportion to the number of fractional residues in the subset and to the number of copies of the subunit in the complex. Expressed mathematically, this is

$$<\Phi_{set}>_{weighted} = \frac{\sum N \cdot \Phi_{set}}{\sum N} \quad (3)$$

where N = (number of fractional amino acids in set) × (number of subunits in complex). The distinction between the two types of averages is that a weighted average for the hydropathy of the contact surfaces, for example, is the hydropathy of the average con-

tact surface; whereas, the unweighted average is the hydropathy of the contact surface of the average unique subunit.

This treatment is straightforward for isologous homooligomers, that is, for oligomers of chemically identical chains which are conformationally identical and which make contacts that are identical in each partner.[9] However, for heterologous homooligomers a different treatment was required. These are cases in which chemically and conformationally identical subunits are orientationally asymmetric thereby making different contacts, or in which the subunits are conformationally nonidentical resulting in different populations of exterior and interior atoms as well as of contact atoms (unless the conformational differences are far from the contact surface). An average of the differing subset hydropathies was calculated and was entered as one entry into the unweighted average. For the weighted average, the differing subset hydropathies for the individual subunits could be entered directly without weighting for the number of copies.

The human oxy- and deoxyhemoglobins posed a special problem as they differ essentially in the quaternary arrangements of the four subunits though with some small conformational differences. We decided that the whole, exterior surface and interior hydropathy values for both states should be represented only once in the unweighted average. Because of the small differences in exterior and interior atom populations between oxy and deoxy states, average hydropathies for the two states were entered. However, contact surface and noncontact exterior surface hydropathies were entered once for each state because the respective contact surfaces were taken to be unique sets. For the weighted average, however, oxy- and deoxyhemoglobin were treated as unique structures in their entirety, and each subset hydropathy for both forms was entered once each. As necessary, values for each of the α- and β-subunits were weighted by a factor of two for the weighted average calculation. Sickle cell hemoglobin, being treated as a distinct protein species, comprised a unique entry for all subsets. Standard deviations were derived from the unweighted averaging so that they would be a reflection of the fluctuations from protein to protein.

Hydropathy Complementarity

We have defined the unnormalized hydropathy complementarity (HC) of a protein–protein interface as the average of the absolute values of the differences in atom-based hydropathies of the interacting pairs of atoms making up the interface:

$$\mathrm{HC} = \frac{\sum_{i,j} |\phi_j - \phi_i|}{N_{i,j}} \qquad (4)$$

where ϕ_i and ϕ_j are the atom-based hydropathies of the two interacting atoms i and j, and $N_{i,j}$ is the number of interacting pairs. Because the summation is over all atom pairs, atoms in contact with two or three neighboring protein atoms will appear in the list two or three times, respectively.

We then carried out a normalization. The same calculation was repeated but with the atoms of one contact surface randomly paired with those of the other using a random number generator. Because the randomization mismatches atoms, random HC values generally will be higher than that for the actual pairing. The HC calculation for random pairings was repeated 50 times to obtain an average random HC and a standard deviation, σ. The difference between the average random HC and the actual, unnormalized HC was divided by the standard deviation to give the final expression of hydropathy complementarity in multiples of σ:

$$\mathrm{HC}_{\mathrm{normalized}} = \frac{\mathrm{HC}_{\mathrm{random}} - \mathrm{HC}_{\mathrm{unnormalized}}}{\sigma} \qquad (5)$$

Application and Averaging of Hydropathy Complementarity

Our convention for defining a subunit's interface has been to determine the contacts of its atoms with *all* its neighbors, as distinct from a convention which is restricted to a particular subunit–subunit pairing. Both definitions yield the same interface for dimers. Hydropathy complementarities were calculated for each unique subunit and for each unique interface. Thus, an isologous tetrahedral homotetramer yields one unique interface. However, the heterologous tetrahedral homotetramer, superoxide dismutase (2SOD), provides four distinct values because each of the four chemically identical, but conformationally nonequivalent, subunits yields a unique set of contacts with its three neighbors. The VP4 subunits of rhinovirus and mengovirus capsid proteins were included in the hydropathy complementarity calculations.

Both weighted and unweighted HC averages were calculated. The unweighted average was a straight averaging of HC values of all chemically unique subunits. The unique HC value for an isologous tetrahedral homotetramer comprised one entry. The average over the four different interfaces for a heterologous tetrahedral homotetramer comprised one entry, too. The weighted average summed the unique HC values in proportion to (1) the number of atom pairs comprising the interface, (2) the number of times each unique interface appears in the multisubunit complex, and (3) the geometry of association. For example, the numbers of interfaces in a tetrahedral, square planar, and linear homotetramer are 6, 4, and 3, respectively. For an isologous tetrahedral tetramer a weighting of 6/3 had to be

TABLE I. Hydropathy Averages for 42 Protein Structures

	Weighted average	Unweighted average	Standard deviation*
Whole	0.06	0.06	0.07
Interior	0.24	0.23	0.08
Total exterior surface	−0.14	−0.13	0.09
Contact surface	−0.02	0.00	0.21
Noncontact surface	−0.16	−0.14	0.12
$\Phi_I - \Phi_E$†	0.38	0.37	0.10
$\Phi_C - \Phi_{NC}$‡	0.13	0.14	0.24
Complementarity (HC)	4.7σ	2.3σ	4.2σ

*Standard deviations were derived from the unweighted average to indicate the expected spread of values from one individual subunit to another.
†Difference between the hydropathies of exterior and interior surface atoms.
‡Hydropathy difference between contact and noncontact exterior surface atoms.

used: there is a division by 3 because each subunit has 3 identical pairwise interfaces, and there is a multiplication by 6 because there are 6 such pairings. The standard deviation was calculated from the unweighted average.

Visual Examination

Molecular structures were displayed on Silicon Graphics IRIS 3020 or Stellar GS1000 raster graphics terminals using the graphics software HYDRA and Quanta (Polygen Corp., Waltham, Mass.), respectively.

RESULTS

General Trends in Hydropathy Distributions

We present the overall results of our survey of the distribution of hydropathies in the subunits of oligomeric proteins and in the proteins of protein–protein complexes in Table I and Figure 2. The latter is a "hydropathy level diagram" that visually presents some of the data in Table I in a conceptually useful way. The diagram shows the weighted averages of the hydropathies for whole subunits, total exterior surfaces (i.e. solvent-accessible, including contact and noncontact atoms), interiors (non-solvent-accessible), contact surfaces, and noncontact exterior surfaces. The vertical scale is marked in terms of the hydropathy scale. Indicated immediately to the right of the scale are the amino acids within this range. The span of hydropathies relevant to our survey is small relative to the full range encountered in individual amino acids. The values encountered stretched from a level between that of histidine and serine to somewhat above cysteine. Data for all the proteins in the survey are tabulated in Table II.

The subunits sequester their total hydropathy into exterior and interior subsets in a consistent fashion. The interior hydropathy is 0.37 ± 0.10 hydropathy units more than that of the exterior. Structures demonstrating smaller splittings tend to be cleavage products of complete chains, such as Fab fragment heavy chains and the VP subunits of mengo- and rhinovirus capsid proteins, or deviate from globularity, as in the cases of U-shaped wheat germ agglutinin or the rod-shaped HA2 subunit of hemagglutinin. Figure 3 shows pictorially how exterior and interior hydropathies are sequestered in a subunit of lactate dehydrogenase (4LDH). It is a cross section through the subunit in which the hydrophobic amino acids have been colored in shades of yellow and the hydrophilic amino acids in shades of blue. The atoms have been represented by flat discs. It is clear from this photograph that the blue periphery and the increasing intensity of yellow toward the center of the cross section dramatically exemplifies the sequestering of hydropathies between interior and exterior.

On average, the contact surfaces are more hydrophobic than exterior surfaces, and therefore also more hydrophobic than the noncontact (exterior) surfaces. However, the variation in the values for the hydropathy difference between contact and noncontact surfaces, $\Phi_C - \Phi_{NC}$, is remarkably high, as indicated by its large standard deviation. Table II shows that some contact surfaces in the survey were actually more hydrophilic than the noncontact surface. While protein–protein interactions are on average driven by a net hydrophobic force, there are some that are driven by a net hydrophilicity-based attraction. If we classify a hydrophobicity-driven interaction as simply one for which the contact surface hydropathy is greater than that of the noncontact surface (i.e., $\Phi_C - \Phi_{NC} > 0$), and vice versa for hydrophilicity-driven interactions, we find that there are about twice as many hydrophobicity-driven associations. Also, the distribution is skewed in favour of net hydrophobicity. The (unweighted) average value for $\Phi_C - \Phi_{NC}$ among the hydrophobicity-driven complexes is 0.28 ± 0.16 and −0.12 ± 0.09 for the hydrophilicity-driven complexes. It may be worth noting that there were very few subunits whose contact surfaces were as hydrophobic as their interiors. This means that contact surfaces generally do not resemble protein interiors in terms of hydropathy.

As an example, the contact surface for catabolite gene activator protein (3GAP) is shown in Figure 4 with the corresponding hydropathy level diagram in Figure 5. Figure 4a indicates the atoms of one subunit that are in contact with the other, which is not shown. The whole subunit is blue except for its contact atoms, which are red. In Figure 4b, the subunit is shown from the same viewing angle but has been colored in the hydropathy scale. The yellow-orange hydrophobic ridge corresponds with the contact region.

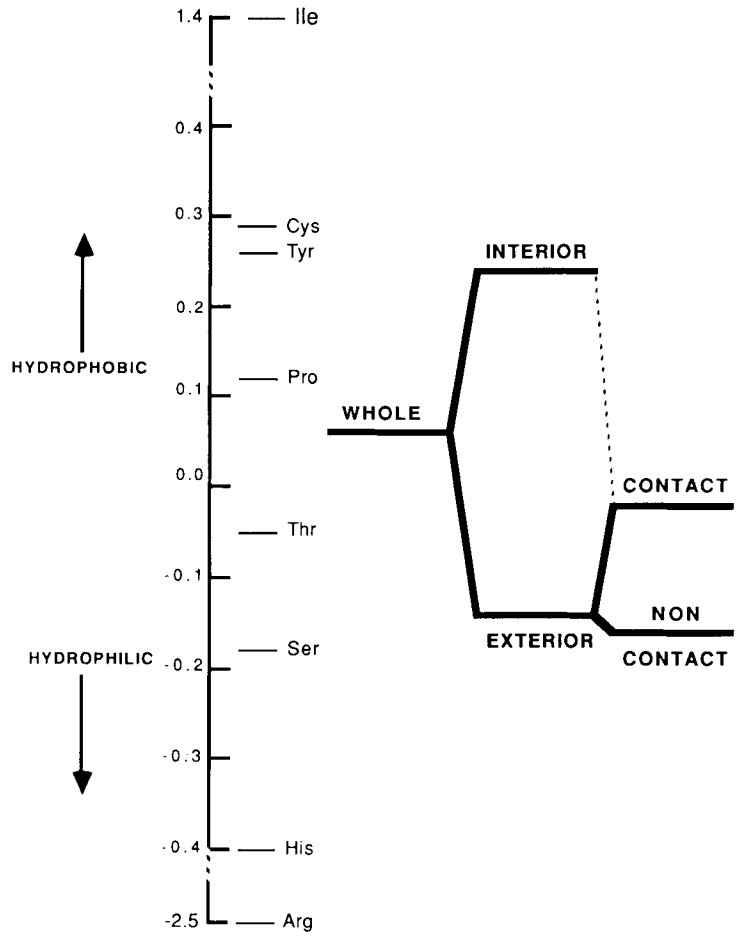

Fig. 2. Hydropathy level diagram indicating the sequestering of total hydropathy of the weighted average subunit into the various categories of protein atoms. The whole subunit divides into total solvent-accessible exterior surface and solvent-inaccessible interior subsets. The subset of exterior atoms splits further into contact surface and noncontact exterior surface. The minor component of the solvent-inaccessible interior subset contributing to the contact subset is indicated by the dotted line. The vertical scale indicates the range of hydropathy values occupied by the average values and the extremes of the scale. Representative amino acids are listed. The normalized consensus scale[8] has values: Ile 1.4, Phe 1.2, Val 1.1, Leu 1.1, Trp 0.81, Met 0.64, Ala 0.62, Gly 0.48, Cys 0.29, Tyr 0.26, Pro 0.12, Thr −0.05, Ser −0.18, His −0.40, Glu −0.74, Asn −0.78, Gln −0.85, Asp −0.90, Lys −1.5, Arg −2.5.

General Trends in Hydropathy Complementarities

Table I shows that the weighted and unweighted averages of hydropathies are similar, but that the corresponding averages for hydropathy complementarity differ. The average interprotein interface, as characterized by the weighted average, has a HC value of 4.7σ. On the other hand, the interface of the average subunit, as given by the unweighted average, has a value of only 2.3σ. HC values have a large spread: the standard deviation (from the unweighted averaging) is 4.2σ. This standard deviation is a measure of the spread of values from one unique subunit interface to another. Nevertheless, these results successfully demonstrate that there is generally a coordinated cooperation between the hydrophobic and hydrophilic interactions. The fact that the unweighted average HC is smaller than the weighted average HC indicates that there were several proteins that had small or no complementarity and that their surface areas were small. When unweighted averages were calculated, these small surface area, low complementarity interfaces were given equal weighting with large surface area, high complementarity interfaces, and hence brought down the average HC value. The weighted average corrects this imbalance by giving proportionately greater weighting to large surfaces.

Figure 6 pictorially demonstrates hydropathy complementarity in beef liver catalase (7CAT). Fig-

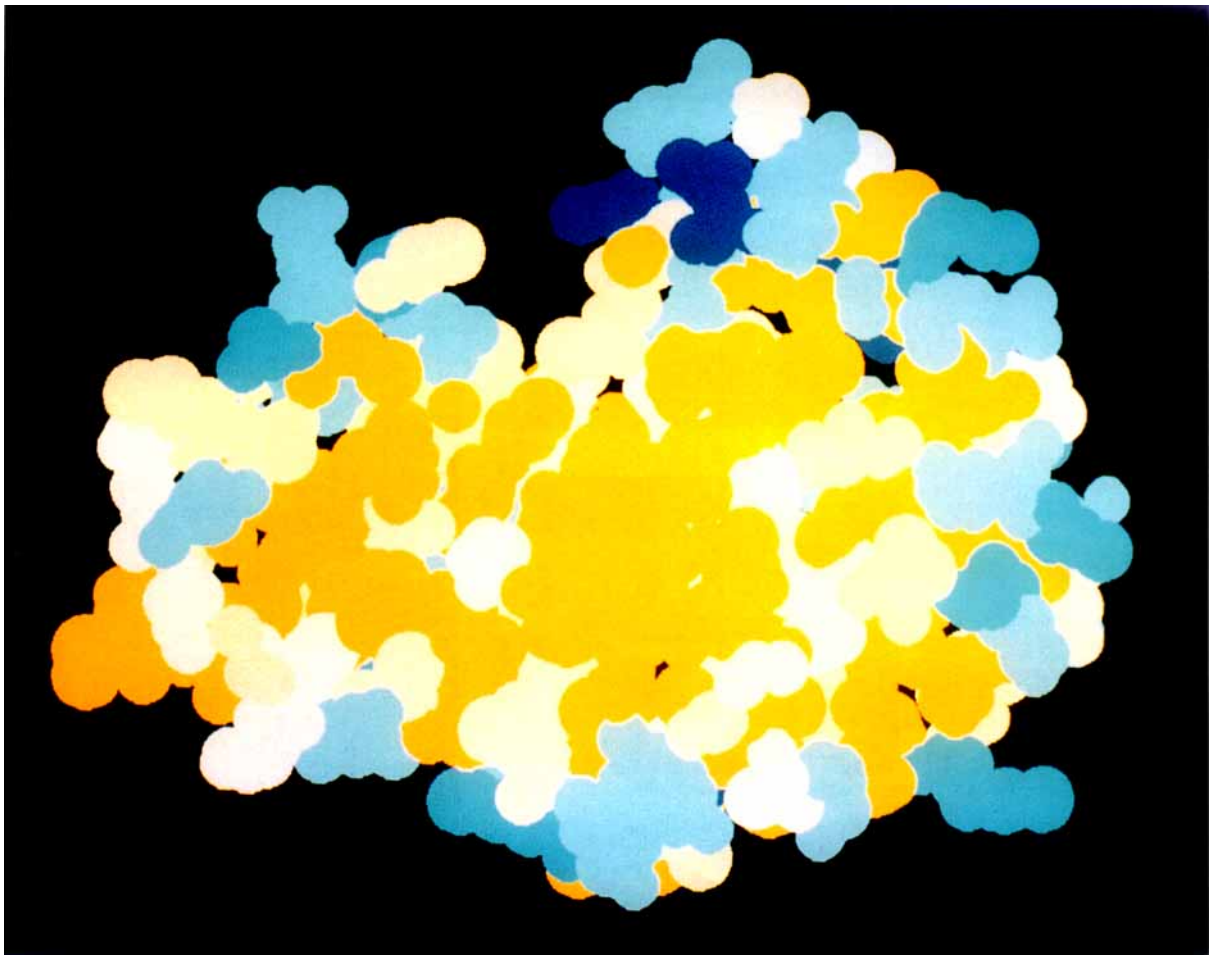

Fig. 3. Cross section of a subunit of lactate dehydrogenase (4LDH) showing the internal distribution of hydropathy. Represented as discs, the atoms of the hydrophobic amino acids are colored in shades of yellow and those of hydrophilic amino acids in shades of blue. This visually demonstrates the quantitated sequestering of hydropathy, $\Phi_I - \Phi_E$, for lactate dehydrogenase given in Table II.

ure 6a shows one of the four subunits colored homogeneously in white except for its contact atoms which are colored in the hydropathy scale. Figure 6b shows the same subunit totally colored in white, with superimposed atoms from the three neighbors with which it is in contact colored in the hydropathy scale. Careful comparison of these photographs should convince the reader that, in accordance with the quantitation, hydrophobic centers (yellow-orange) of one subunit generally associate with hydrophobic centers of neighboring subunits; and hydrophilic (blue) centers associate with hydrophilic centers.

Individual Cases

A major concern of ours was whether or not our quantitations represent true characterizations of protein–protein interaction surfaces. This is essential if the protocols are to be meaningful and useful in future studies. We compared our numerical results with all the qualitative descriptions made by the authors of the individual crystallographic structures. In large part there is unambiguous agreement.

Cytochrome c'

This dimer (2CCY) has an interface hydropathy that is 0.10 units greater than that of the noncontact surface ($\Phi_C - \Phi_{NC} = 0.10$), and the contact surface is described by Finzel et al.[10] as being "characterized by extensive hydrophobic interactions."

Citrate synthase

We measured citrate synthase (2CTS) as having a very hydrophobic contact surface ($\Phi_C - \Phi_{NC} = 0.43$) and large hydropathy complementarity (7.2σ). In agreement with these measurements, the authors Remington et al.[11] write that "most residues [*of the contact surface*] are hydrophobic. The residues that

TABLE II. Hydropathy Differences and Complementarities for Unique Protein Subunits

Code*	n†	Protein description	Chain	$\Phi_t - \Phi_w$	$\Phi_c - \Phi_{NC}$	HC‡ (σ)	$N_{i,i}$§	Ref.
6ADH	2	Horse alcohol dehydrogenase		0.49	0.46	7.1	546	42
7CAT	4	Beef liver catalase		0.32	0.19	9.8	3650	43
1CC5	2	Cytochrome c_5		0.20	0.44	3.5	43	44
2CCY	2	Cytochrome c'		0.37	0.11	0.2	210	10
5CHA	2	Bovine α-chymotrypsin		0.43	0.29	−1.8	274	45
1CN1	2	Concanavalin A		0.42	−0.05	2.4	465	46
2CTS	2	Citrate synthase		0.44	0.43	7.2	1806	11
3FAB	2	Human myeloma IgG Fab (New) fragment	Light chain	0.36	0.09	6.3	462	21
			Heavy chain	0.21	0.26			
1FBJ	2	Mouse galactan-binding IgG Fab fragment	Light chain	0.32	0.40	1.0	523	47
			Heavy chain	0.32	0.47			
1FC1	2	Human IgG Fc fragment		0.42	−0.08	−2.1	304	23
1FC2	2	Complex of protein A fragment with Fc fragment	Protein A	0.34	0.27	0.8	196	23
			Fc fragment	0.37	0.41			
1FDH	4	Human fetal deoxyhemoglobin	α	0.45	−0.20	2.6	519	48
			γ	0.52	0.05	0.3	445	
3GAP	2	Catabolite gene activator		0.47	0.64	1.1	279	12
1GD1	4	Glyceraldehyde dehydrogenase *(Bacillus)*		0.46	0.21	1.6	1159	13
2GN5	2	DNA binding protein		0.14	0.38	0.9	123	14
1GPD	2	Glyceraldehyde dehydrogenase (lobster)		0.39	−0.08	0.0	254	49
1GP1	2	Glutathione peroxidase		0.48	−0.05	3.6	208	15
3GRS	2	Glutathione reductase		0.43	0.27	9.6	1000	16
1HBS	4	Sickle cell hemoglobin	α	0.44	−0.08	0.3	480	50
			β	0.50	0.01	−1.9	456	
3HHB	4	Human deoxyhemoglobin	α	0.43	−0.25	3.0	513	51
			β	0.56	−0.03	0.4	453	
1HHO	4	Human oxyhemoglobin	α	0.43	−0.06	−3.0	555	52
			β	0.56	−0.07	−1.4	440	
2HFL	2	Complex of Fab fragment with chicken lysozyme	Fab fragment	0.36	−0.11	2.5	269	17
			lysozyme	0.63	−0.27			
1HMQ	8	Hemerythrin		0.43	−0.16	4.4	431	53
1HMG	6	Influenza virus hemagglutinin	HA1	0.37	0.15	8.4	1456	54
			HA2	0.18	0.24	9.9	1994	
1IG2	2	Human intact myeloma (Kol) immunoglobulin	Light chain	0.34	0.28	8.4	561	55
			Heavy chain	0.28	0.38			
4LDH	4	Lactate dehydrogenase		0.49	0.22	6.8	1636	56
1MCP	2	Mouse phosphocholine binding Fab	Light chain	0.34	0.47	2.4	522	22
			Heavy chain	0.23	0.42			
1MEV	4	Mengovirus capsid protein	VP1	0.23	−0.05	3.7	1192	57
			VP2	0.40	0.14	0.5	981	
			VP3	0.28	−0.03	0.5	658	
			VP4			4.5	223	
2PAB	4	Prealbumin		0.38	0.49	1.4	497	58
3PGM	4	Phosphoglycerate mutase		0.35	−0.12	−7.0	563	28
1PP2	2	Phospholipase A_2		0.32	−0.07	−2.4	335	30
1PYP	2	Inorganic pyrophosphatase		0.32	0.23	6.6	339	59
1REI	2	Bence–Jones dimer, V_L region, κ type		0.30	0.38	−2.2	171	60
2RHE	2	Bence–Jones dimer, V_L region, λ-type		0.28	0.18	2.3	225	61
4RHV	4	Rhinovirus capsid protein	VP1	0.29	0.15	3.6	2370	62
			VP2	0.34	0.05	−0.1	1399	
			VP3	0.35	−0.02	3.2	2018	
			VP4			4.0	733	
4SBV	3	Southern bean mosaic virus capsid protein		0.29	−0.27	−0.9	574	24
2SOD	4	Superoxide dismutase		0.37	0.55	−2.9	209	63
2SSI	2	Subtilisin inhibitor		0.26	−0.07	−2.5	167	64
2TBV	3	Tomato bushy stunt virus capsid protein		0.22	−0.36	0.5	448	65
1TIM	2	Triose phosphate isomerase		0.45	0.15	1.5	657	26
3WGA	2	Wheat germ agglutinin		0.16	0.05	19.	733	27
2WRP	2	Trp repressor		0.37	0.32	0.4	507	66

*Brookhaven Data Bank code.

†Number of subunits in the complex.

‡HC values have two-digit accuracy, but for comparison are listed only to one decimal place.

§Interface size expressed in number of atom pairs between the subunit and all neighbors. The average value is given for the heterologous homooligomers 2SOD, 2TBV, and 4SBV.

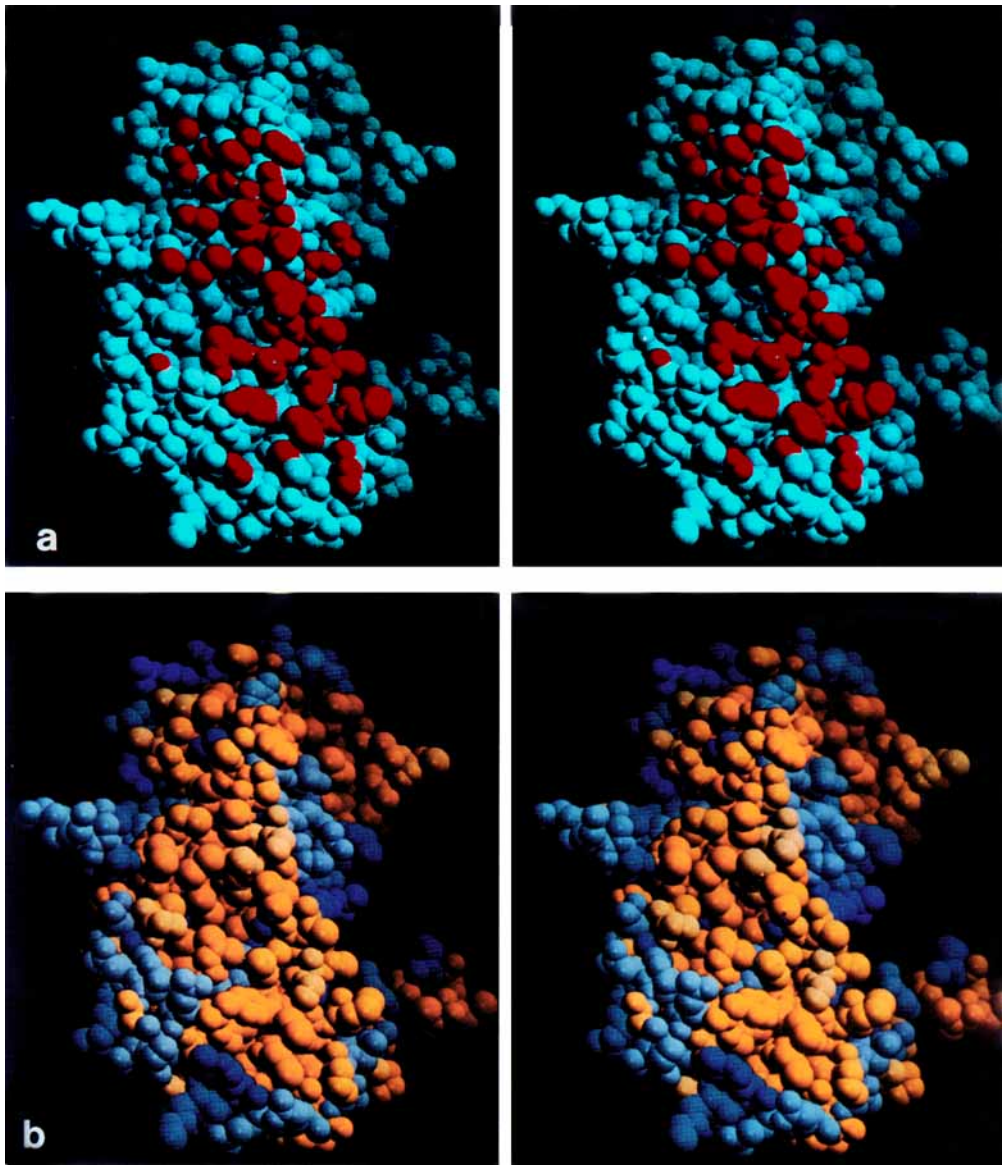

Fig. 4. Stereoscopic photographs of one subunit of the homodimer, catabolite gene activator protein (3GAP) showing the contact surface. **(a)** The contact atoms are indicated in red on a blue background of noncontact atoms. **(b)** All the subunit's atoms are colored in the hydropathy scale with those belonging to hydrophobic amino acids in shades of yellow-orange and those of hydrophilic amino acids in shades of blue. The hydrophobic (yellow-orange) ridge corresponds fairly well with the contact surface.

are hydrophilic make, almost without exception hydrogen bonds and salt-bridges between the two subunits."

Catabolite gene activator protein

This protein (3GAP) has an interface hydropathy of 0.64 units greater than that of the noncontact surface (see Figs. 4 and 5), and it was observed by Weber and Steitz[12] that "most of the interactions are hydrophobic with only two symmetrical hydrogen bonds between the subunits." The low complementarity of 1.1σ indicates that there is little cooperation between hydrophobic and hydrophilic forces in agreement with the paucity of hydrophilicity-based interactions.

Glyceraldehyde dehydrogenase

Bacillus stearothermophilus glyceraldehyde dehydrogenase (1GD1) is described[13] as having a "large number of hydrophobic contacts." The authors also tabulate an extensive number of hydrogen bonds, salt bridges, and water bridges. Our evaluation indicated a very hydrophobic surface ($\Phi_C - \Phi_{NC} = 0.21$) but only a moderate hydropathy complementarity of 1.6σ.

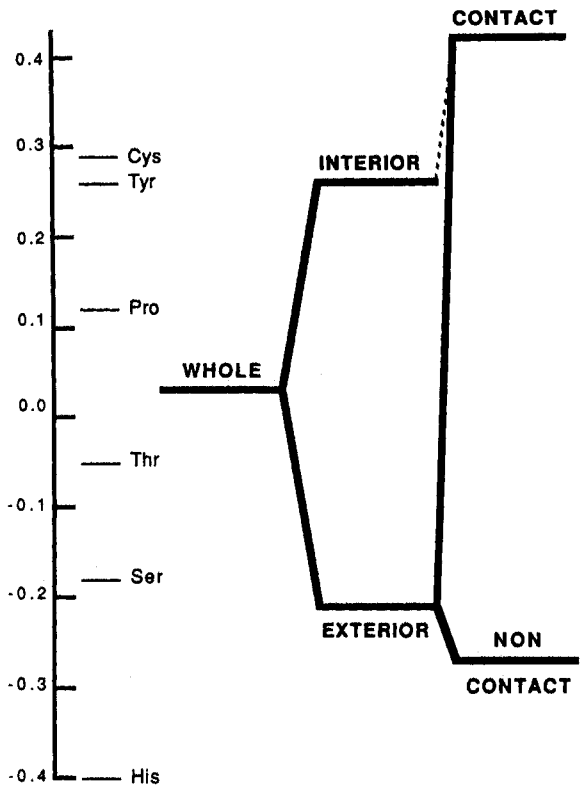

Fig. 5. Hydropathy level diagram for the catabolite gene activator protein subunit shown in Figure 4. It quantitatively presents the allotment of hydropathy in the subunit and, importantly, the high hydrophobicity of the contact surface that is shown in the photograph of Figure 4b.

DNA binding protein

DNA binding protein (2GN5) dimer association is described by Brayer and McPherson[14] as "primarily hydrophobic, and it is known that the dimer is stable under conditions of high ionic strength, extreme pH values, dilution and at elevated temperatures." In tabulating the interacting groups they found at least 20 hydrophobic side chains, at least 10 electrostatic interactions, and one symmetrical H-bond pair. Our quantitation indicates that $\Phi_C - \Phi_{NC}$ is equal to 0.38, in agreement with this description. However, in conflict with this description was our determination of HC as being only 0.91σ.

Glutathione peroxidase

Our results with the dimeric glutathione peroxidase (1GP1) indicated no net hydrophobicity or net hydrophilicity for its contact surface. However, there was a high complementarity of hydropathy of 3.6σ. We noted many ionic interactions as well as aromatic stacking of tyrosine rings in this protein. The authors, Epp et al.,[15] observed that "both polar and apolar groups are found among the contact residues" and "in addition to the hydrophobic contacts, four H-bonds can be formed across the twofold axis." Thus, the cooperation and nearly equal balance of hydrophobic and hydrophilic forces are reflected in our quantitation.

Glutathione reductase

This protein (3GRS) dimerizes with a net hydrophobic attraction, and a high degree of participation of hydrophilic bonds in concert with the hydrophobic bonds, as indicated by its HC value of 9.6σ. The authors, Karplus and Schultz,[16] have described the dimer interface in agreement with this quantitation. The interface is 43–58% nonpolar, but boasts many hydrogen-bond interactions, a salt bridge, and three water bridges. Interestingly, there is some spatial segregation of hydrophilic amino acids: "in all cases the charged atoms are located at the periphery of the interface."

Fab fragment–lysozyme complex

The antigen–antibody complex, Fab fragment–lysozyme (2HFL), is overall hydrophilic at its binding interface with $\Phi_C - \Phi_{NC} = -0.11$, but with hydrophobic and hydrophilic residues collaborating in a coordinated fashion to maintain the contact (HC = 2.5σ). The authors of this structure, Sheriff et al.,[17] found in their detailed examination 74 van der Waals contacts, 3 salt bridges and 10 hydrogen bonds. Comparing this lysozyme–Fab complex with another lysozyme–Fab complex specific to a different epitope, D1.3-lysozyme,[18] Sheriff et al. found that the latter had no electrostatic interactions, and its association constant was almost two orders of magnitude weaker. Unfortunately, we did not have access to these coordinates, but we would predict that it would have a smaller hydropathy complementarity.

Hemoglobin

Jaenicke and Helmreich[19] have stated, based on an atom by atom comparison of intersubunit interactions of oxy (1HHO) and deoxy (3HHB) forms of human hemoglobin, that "the deoxy tetramer is more strongly cross-linked with more salt-bridges and constrained while the oxy state is loose" and "dissociates easily." This description correlates with our quantitation. The dramatically increased contact surface hydrophilicities and HCs for the deoxy structure is due to this increased hydrophilic–hydrophilic pairing of the salt bridges. Note the negative HC values for the subunits of the oxy state and the positive HC values for the deoxy state in Table II. Negative HC values correspond to anticomplementarity of the interface. The greater HC for the deoxy form manifests the greater specificity in the quaternary positioning of the subunits. The increased number of salt bridges is reflected in the greater hydrophilicity of the α-subunit's contact surface in the deoxy state.

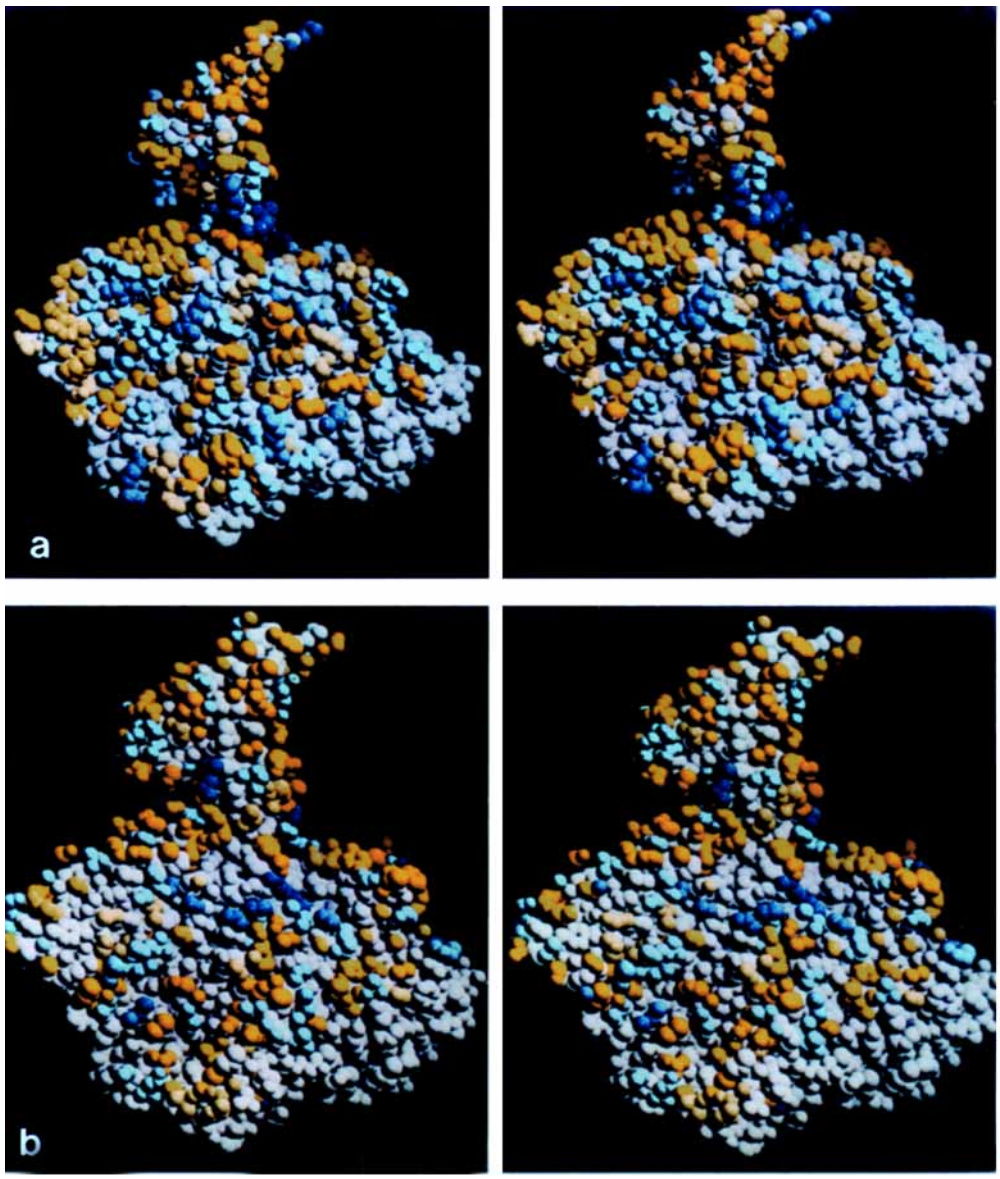

Fig. 6. Hydropathy complementarity is demonstrated visually in the interface of a subunit of tetrameric beef liver catalase (7CAT) with its three neighbors. **(a)** The subunit alone, with its contact atoms colored in the hydropathy scale and the others white. **(b)** The entire subunit plus the atoms belonging to the other three subunits with which it is in contact. The entire subunit is white with the contact atoms belonging to the neighboring subunits colored in the hydropathy scale. In general, regional clustering of hydrophobic (yellow-orange) atoms in **a** are matched with similar clustering seen in **b.** This holds true for hydrophilic (blue) atom clustering.

Hemerythrin

Hemerythrin's (1HMQ) eight subunits are strongly bonded together by more hydrophilic bonds than hydrophobic ones ($\Phi_C - \Phi_{NC} = -0.16$) (see Fig. 7) and they cooperate with each other (HC = 4.4σ). Stenkamp et al.[20] tabulated van der Waals contacts, hydrogen bonds, and water bridges between the subunits suggesting, in accord with the HC value, that a combination of hydrophobic and hydrophilic forces is at play. Additionally, the authors write that it is "specific interactions that orient the subunits for proper docking." The HC value may be an expression of this specificity.

Immunoglobulin IgG

The interfaces in the IgG immunoglobulin structure are divisible into two categories: light chain–heavy chain interaction in the Fab region and the isologous heavy chain–heavy chain interaction in the Fc region. It was fortunate to have several ex-

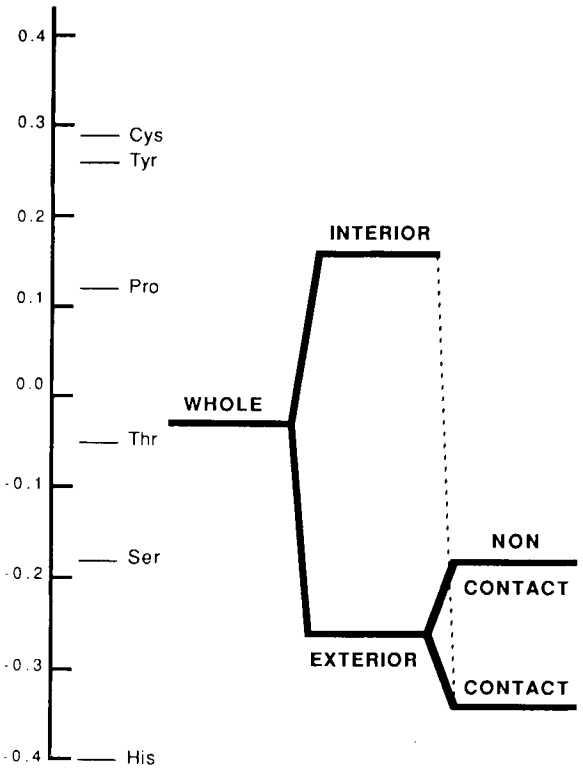

Fig. 7. Hydropathy level diagram for a hemerythrin subunit (1HMQ) showing that the binding forces in this octameric protein are predominantly hydrophilic.

amples of IgG molecules in order to check for generalities within a single class of molecules. For all of the human and mouse Fab fragments (3FAB, 1FBJ, 1IG2, and 1MCP) the binding between light and heavy chains is driven by a large net hydrophobic component, and is well coordinated between hydrophilic and hydrophobic bonds. In contrast, the heavy chain–heavy chain interaction in the Fc region (1FC1) is slightly hydrophilic and shows no complementarity of hydropathy. The available descriptions of the contacts in the Fab region[21,22] indicate they are compact and very tight with extensive van der Waals contacts, numerous H-bonds, and an ion pair. Diesenhoffer[23] studied the contact types in the Fc fragment and tabulated H-bonds and charged pairs but also some hydrophobic contacts. One interaction surface in the Fc region, the CH_2–CH_2, which is proximal to the hinge region, is mediated by carbohydrates. We suggest, in the light of potential flexibility in the Fc fragment as suggested by the dominance of hydrogen bonds and ion pairs (see section on functional correlations), that the two disulfide bonds in this region may be there to counteract the trend to flexibility and to maintain immobility in interchain interaction in the Fc region.

Southern bean mosaic virus capsid protein

This (4SBV) is a trimeric protomer held together primarily through hydrophilic residues ($\Phi_C - \Phi_{NC} = -0.27$). This is also borne out by the detailed tabulation by Rossmann et al.[24] The atomic structure revealed Ca^{2+} ions chelated by glutamates at the quasi-3-fold axis, numerous putative hydrogen bonds, as well as clustering of charged interactions at the quasi-3-fold axis.

Tomato bushy stunt virus capsid protein

The trimeric protomer of tomato bushy stunt virus capsid protein (2TBV) is characterized by remarkably hydrophilic contact surfaces between the subunits ($\Phi_C - \Phi_{NC} = -0.36$). This is in full accord with the detailed analysis of the contact surfaces given by Harrison.[25]

Triose phosphate isomerase

The authors' description of triose phosphate isomerase (1TIM)[26] indicates spatial segregation of polar and nonpolar contacts. At the dimer's 2-fold axis the subunits are held by salt bridges, main and side chain hydrogen bonds. However, further out (particularly at loop 70–80 and Met 14) the interactions are hydrophobic. Our quantitation reflects this segregation moderately well ($HC = 1.5\sigma$).

Wheat germ agglutinin

We measured a remarkably large hydropathy complementarity ($HC = 19\sigma$) for this lectin (3WGA). The author of this structure, Wright,[27] tabulated the contacts and found that over one third of the interactions are between identical residues, thus giving rise to many atom pairs with perfect hydropathy matching ($\Phi_j - \Phi_i = 0$). Furthermore, many of the other contacts are conservative, meaning that the interactions, while not between identical residues, preserve hydropathy matching. The contacting surfaces were described as consisting of a combination of hydrophobic and hydrophilic contacts, the latter consisting of hydrogen bonds, a symmetrical ion pair, and some water-bridged interactions.

Exceptional Cases

There were two proteins that unambiguously did not agree with our quantitation:

Phosphoglycerate mutase

Yeast phosphoglycerate mutase (3PGM), which has a net hydrophilic attraction between subunits, demonstrated the largest anticomplementarity ($HC = -7.0\sigma$). The structure was checked on the graphics screen for errors in using the transformation matrices that generated the other subunits, but the image matched the published figures for this enzyme's structure.[28] We also noticed the unusual occurrence of hydrophobic patches on the external, noncontact surface of the tetramer.

Subtilisin inhibitor

A discrepancy between our results and the literature account was found for subtilisin inhibitor (2SSI). The contact surface is only slightly more hydrophilic than the noncontact surface and no coordination of hydrophobics and hydrophilics exists (HC = -2.5σ). However, Mitsui et al.[29] found that all the 48 interactions under 3.4 Å, except for two possible hydrogen bonds, are nonpolar. This discrepancy may very well be due to the "noise" inherent in using a hydropathy scale that is spatially coarse due to its being based on whole amino acid hydropathies, rather than using a scale that is more finely resolved by being based on individual atoms. This "noise" would be especially apparent on relatively small contact surfaces, which is the case in subtilisin inhibitor. If the fine details of the interactions, which are lost in our current methods, has been included in our study, it might then have succeeded in showing complementarity. For example, the authors of this structure noted that the straight (hydrophobic) chain of Arg-90, which is involved in the two symmetrical hydrogen bonds effected by the distal (hydrophilic!) guanidinium group, is "held tightly by a (hydrophobic) pocket formed by the Thr 15, Ala 22, Pro 27 and Leu 79 of the opposite subunit." This complementarity is lost in our current method because all of the atoms of arginine were assigned the same hydropathy value.

Practical Applications

As seen from the examples discussed in this section, our quantitation method for the most part agrees with the detailed, qualitative descriptions of the contact surfaces of individual structures given by the authors. We began to feel confident, therefore, that the findings of our survey might form a basis for predicting correct interacting partners. The following supports our expectation. Before visually checking the coordinate transformation for the second subunit of dimeric inorganic pyrophosphatase (1PYP), an attempt was made to predict the binding surface and the geometry of binding. First, the subunit surface was examined for any concentrations of hydrophobic residues. As we have seen in the survey, hydrophobically bound oligomeric proteins are in the majority, so this check had a reasonable chance of success. Luckily, one region became immediately obvious. Then, in a second image of the subunit was generated. By concentrating on the physical topology of the putative binding surface, the second subunit was manipulated until a snug fit was made with the first subunit. The docking was then checked by generating an image of the true partner. It was pleasing to discover that the predicted and true positions of the second subunit were qualitatively very close. Both the correct binding site and the general orientation were predicted correctly.

A second example demonstrates the diagnostic value of both hydropathy and hydropathy complementarity. Wheat germ agglutinin (3WGA) was initially analyzed using the coordinates in the Brookhaven Data Bank entry for a complete dimer. Not only were overall "contact surface" hydropathies (Φ_C) very different for each subunit, but the interface hydropathy complementarity was negative. It turned out that we had analyzed a crystallographic dimer rather than the physiological one. Repeating the analysis on the correct pair of subunits, we found identical contact surface hydropathies for the two subunits, as expected of an isologous homodimer, and the highest hydropathy complementarity in the survey.

DISCUSSION

Functional Correlations

While the average subunit–subunit bonding is predominantly hydrophobic, a significant number of subunits have contact surfaces that are more hydrophilic than even those parts of the external surface adjacent to solvent. Proteins require a certain concentration of charged and hydrogen bonding groups on their outer surface for maintaining solubility in an aqueous environment. Therefore, when a contact surface is more hydrophilic than the solvent-facing surface, a special purpose must be present. For instance, all the hemoglobin interfaces are low in hydropathy and some are very hydrophilic. An outstanding feature of hemoglobin's function is the switching of its quaternary organization subsequent to oxygen binding. Does this functional aspect correlate with our characterization of its interfaces?

Several examples indicate that flexibility and conformational dynamism are accomplished through hydrophilic bonds. In the capsid organization of tomato bushy stunt virus (2TBV) there are two types of 2-fold symmetric interactions between trimeric groups of subunits. The problem of maintaining chemically homologous contacts at interfaces that are not symmetrically equivalent has been discussed by Harrison.[25] The subunits are positioned differently in the two distinct 2-fold interactions, however, the same residues are involved in both interfaces. The flexibility required to accommodate the different orientations is achieved by changing the side chain conformations of the hydrophilic, hydrogen bonding, amino acids to access alternate hydrogen bonding partners in the two quasi-equivalent interfaces. Harrison suggested that alternative sets of hydrogen bonding and salt bridging patterns explain the shifting quaternary organization responsible for the disk-to-helix transition in tobacco mosaic virus.

Among the many hydrophilic interactions binding the subunits of the trimeric capsid protomers of tomato bushy stunt virus (2TBV) and southern bean mosaic virus (4SBV) are bridging calcium ions[25]

chelated by negatively charged residues at their quasi-3-fold axes. These capsids swell in the presence of EDTA. This dynamic structural behavior was attributed by Harrison to the removal of Ca^{2+} ions from these sites.[25]

Phospholipase A2 (1PP2) is a dimeric, allosteric enzyme whose activity is enhanced by Ca^{2+} binding. Keith et al.[30] and Brunie et al.[31] have examined the intersubunit interface and found an ionic interaction, which might compete with Ca^{2+} binding, and several other polar interactions including bridging water molecules. It was postulated that the enhancement of enzymatic activity consequent to calcium chelation is effectuated by the disruption of the intersubunit ion pairs. The Ca^{2+} binding event is somehow transmitted to the nearby catalytic sites. Here also, hydrophilic bonds at an interface seem associated with dynamism in subunit–subunit interaction.

Some multisubunit proteins may require a stable and highly accurate interface, rather than one that is dynamic. Glutathione reductase (3GRS) is a good example as it is a dimer in which the interface is an integral part of the enzyme's catalytic site. Glutathione is bound between the subunits and the catalytic side chains come from both subunits.[16] As isolated subunits cannot be active, it is very important that the dimeric state remain intact with subunits accurately positioned relative to each other. The glutathione reductase interface has an extremely high hydropathy complementarity (HC = 9.6σ) suggesting that high complementarity is a mechanism for achieving accurate and static positioning of bound proteins. It is noteworthy that low HC values, which would favor flexible association, were found in the previous examples of hydrophilic interfaces.

Methodology

Choice of hydropathy scale

Several hydropathy scales for the amino acids have been devised and have been reviewed by Eisenberg.[8] Most scales are based on experimentally determined free energies of transfer of the appropriate amino acid analogues from water to a hydrophobic liquid such as *n*-octanol. Other scales are based on the frequency of appearance of the amino acids on the solvent accessible surface of folded proteins. Eisenberg[8] merged four scales, two based on free energies of transfer and two based on surface accessibility, and normalized the result so that the average was zero and the standard deviation was 1.0 (Fig. 2). We chose this normalized "consensus" scale for our survey placing faith in Eisenberg's argument that any errors present in any one original scale due to peculiarities of experimental technique or method of calculation will be minimized by merging with other scales.

A possible criticism of this approach derives from a common belief that proteins sequester amino acid side chains so that their charged or polar tips lie on the exterior surface while their aliphatic carbon atoms are buried in the interior. If this is true, apportioning hydropathy equally among all atoms would be incorrect for charged residues spanning the interior–exterior boundary, and differences between interior and exterior subsets would be underestimated.

We tested the validity of our method by examining all charged residues in the horse alcohol dehydrogenase dimer (6ADH). The solvent-accessibility program, MS, first was used to identify all external atoms, and then the "exposure" of side chains to solvent was measured. Exposure was defined as the number of side chain atoms in contact with solvent divided by the total number of atoms in the side chain. We found a high percentage exposure (73–85%) for all charged residues confirming crystallographic observations that long side chains, such as those for arginines and lysines, often extend into solvent. The results indicate that differences in the total hydropathy obtained using residue-based or atom-based scales should be slight. Assume that an atom-based scale was available, and that it was normalized so that the sum of the atom-based hydropathies would equal the consensus scale value. As most of the charged side chain atoms are exposed to solvent, their contribution to the protein's exterior hydropathy comes not only from the charge-bearing atoms at the tip but also from most of the side chain's aliphatic atoms. In the limit, when the total amino acid is exposed, the two scales give identical values. The hydropathy complementarity function in particular should in practice not be overly sensitive to the actual scale used, as it merely compares whether residues of similar nature are neighbors. A slight "smearing" of the hydropathy across all atoms of a residue could even compensate for experimental errors present in the individual data sets.

Defining the subsets of protein subunits

The categories of whole, total exterior surface, interior, contact surface, and noncontact exterior surface of a folded polypeptide chain define both volume and area. In order to assign hydropathy values to these categories we placed them on a common basis by defining each as a set of whole atoms. Thus, the exterior surface was defined as the set of atoms that could be touched from the outside by a probe sphere representing a water molecule. This definition follows the original suggestion of Richards[32] and was implemented using Connolly's program MS.[7] Two parameters in MS are critical: probe sphere radius and dot density. The algorithm lays a grid of latitude and longitude lines on the van der Waals sphere of each protein atom. Next, a probe sphere is placed on the atom's van der Waals surface at each point of intersection of the latitude and longitude lines. For each of these positions a check is made for

collisions with other atoms. If there are no collisions, then the point is regarded as part of the solvent-accessible surface, and the protein atom to which the point belongs is defined by our criterion as being solvent-accessible and exterior. The grid lining spacing determines the density of the intersection points, and is determined by the program parameter, "dot density."

The radius of the probe sphere was chosen after visualizing the results for the α-subunit of human deoxyhemoglobin (3HHB) on the molecular graphics terminal. With fixed dot-density, the effect of the probe sphere size was examined. Increasing its radius resulted in increasing failure to pick out surface atoms. As the shallower crevices on the surface of the molecule become less and less accessible to the expanding sphere, the atoms lying in these crevices can no longer be touched by the probing sphere. Conversely, with diminishing radii the probe sphere can enter the crevices, and with even smaller radii the sphere becomes even smaller than the spaces between protein surface atoms. At this point the probe penetrates the protein's surface. When the probe radius was less than 1.7 Å, atoms clearly residing deep in the protein's interior were defined as "solvent-accessible." A value of 1.7 Å was chosen because it generated a shell of closely packed atoms with a minimum of intervening spaces, and without mistakenly selecting any interior atoms. It is noteworthy that a commonly used value, the van der Waals radius of oxygen (1.4 Å), was not ideal. The optimal value of 1.7 Å most probably compensates for the hydrogen atoms present in a real molecule, but absent in our data sets.

With the probe radius set at 1.7 Å, the dot density was varied systematically from 1 to 15 $Å^{-2}$ to determine how the number of atoms picked as solvent accessible changes with this parameter. As dot density increased, the number of "exterior" atoms in the α-subunit increased smoothly and asymptomatically from 347 to 488 out of the total of 1,112 atoms. A smaller number of "exterior" atoms arises from deficient sampling of the atomic surfaces by the probe sphere when the dot density is low. As an extreme example, if there is only one collision-free position for the probe sphere on a protein atom, and there is no intersection of latitude and longitude lines at that position, then that atom will be missed. Increasing the number of grid lines by raising the dot density increases the chance of placing a grid line intersection on the collision-free point. The asymptotic value for the hemoglobin α-subunit test was reached at 14 $Å^{-2}$ and this value was used in all subsequent work.

We defined a contact surface of a subunit as the set of all its atoms within 4.5 Å of any atom, or atoms, belonging to any neighbor. The choice of cut-off distance was based on visual examination of the distribution of contact atoms belonging to the α-subunit of human deoxyhemoglobin. Each set of "contact" atoms chosen at varying cut-off distances was displayed as space-filling spheres in a different color than the background of the rest of the subunit. Distances of 4.0 Å or less picked out too few atoms, failing to define a continuous contact surface of closely packed atoms and missing atoms that were obviously interacting. At the other extreme, distances of 5.0Å or greater were inappropriate because they selected pairs that were clearly too distant for interaction, and, in some cases, were even physically separated by an intervening atom. The midpoint of the range between 4.0 and 5.0 Å was chosen as a compromise.

In our care to define the contact surface correctly, it paradoxically turned out that the contact surface does not comprise in strict terms a subset of the exterior surface. While the atoms of a contact surface are necessarily "external," they include more than those that are just solvent accessible. Nonhydrogen protein atoms can be as small as 1.4 Å in radius, while the probe sphere used to define the solvent-accessible atoms has a radius of 1.7 Å. The atoms of a protein can probe deeper into the crevices of an interacting protein than can the larger probe sphere representing solvent water. This is illustrated in Figure 1. Thus, a search for contact pairs exclusively among solvent-accessible atoms will miss a number of bona fide interacting atoms. As a result of our definition, the atoms of the contact surfaces are more densely packed than if they had been chosen from the solvent accessible surfaces. Also, the additional atoms reside more deeply in the protein. We used our structure of trimeric adenovirus hexon[1] to test the quantitative difference between our definition of contact surface and one that is limited to the solvent-accessible atoms. We found that 17% fewer atoms were selected as the "contact" surface when it was derived only from the solvent-accessible atoms than when it was derived from all atoms in the trimer. The hydropathy complementarity was about 30% lower. However, the hydropathy of the contact surface was only 0.02 units less.

The procedure used by other workers[5,33] to define the contact surface is to determine the solvent-accessible surface of the isolated subunits that are buried in the oligomer. A hydropathy value for this "buried accessible surface" is obtained from the difference in hydropathy of the solvent-accessible surface of the complex and that of the isolated subunits. While computationally straightforward, this procedure will miss certain contact atoms and falsely include others. First, depending on the probe radius used, there may be bona fide contact atoms that are solvent inaccessible (IC in Fig. 1) and therefore undetectable by this procedure. Second, there are solvent-accessible atoms that might line an internal cavity, or bubble, which can be created when two concave surfaces come together. These atoms be-

come "buried" upon complexation, but they are not necessarily involved in contacts because they may be too far apart at opposite sides of the internalized cavity.

Hydropathy complementarity

The difference in hydropathy between the contact and noncontact surfaces can tell us whether there is a net hydrophobic or net hydrophilic component to intersubunit attraction. However, this characterization does not provide a complete description of the interacting surfaces. If the contact surface is overall neither hydrophobic nor hydrophilic, hydropathic attraction is not precluded as both types of interaction may be operating equally. The new function identifies and quantifies interfaces by the extent to which both hydrophobic and hydrophilic binding forces collaborate. It determines whether hydrophobic centers of one protein surface interact with hydrophobic centers on the neighboring surface(s), and simultaneously if hydrophilic centers interact with neighboring surface hydrophilic centers. The latter interactions include ionic pairs, hydrogen bonds, and dipole–dipole interactions, which all involve mainly hydrophilic amino acids. The individual interactions are not explicitly determined by the function, which thus measures the overall complementarity of hydropathy for an interface.

A convenient way to measure hydropathy complementarity in a pair of interacting atoms is to calculate the differences in their atom-based hydropathies. Good matching, whether the interaction be hydrophobic–hydrophobic or hydrophilic–hydrophilic, means that the difference between their hydropathy values is small, and vice versa. The hydropathy, HC, for a set of interacting pairs comprising an entire interface is the average of the absolute values of the differences in atomic hydropathies taken over all pairs, as given in Eq. (4). This number is then normalized in order to take into account the range of hydropathies in the interface, to compare results from different oligomeric proteins, and to judge the statistical significance of the findings. If the hydropathy complementarity for a pair of interacting surfaces is statistically significant, it is reasonable to conclude that the surfaces have been brought together by a purposeful match of amino acids with similar hydropathies.

These points may be understood pictorially if the two interacting surfaces are imagined to be sheets covered with intricate, nonregular, blue (hydrophilic) and yellow (hydrophobic) patches that match perfectly when brought together. If both surfaces are totally blue, a perfect match will also be obtained. But this kind of matching cannot show a statistically significant (normalized) complementarity because any orientation of the sheets would have brought about perfect color matching. For an interface composed of only hydrophilic residues, the range of hydropathy values will be limited, and the chance of a bad match is small. For such an interface, a straight tally of HC without normalization against random pairings would give a wrong impression of high complementarity. The normalization procedure, which evaluates HC as a multiple of the standard deviation for random pairings, ensures a meaningful evaluation of complementarity. Most importantly, the normalized hydropathy complementarity can be considered to be a test for and a quantitative measure of the specificity of binding.

Future Work

We look forward to increasing the sophistication of our quantitation method. We believe that the most serious error in our current protocol is the use of equal apportioning of amino acid-based hydropathies to the amino acids' constituent atoms. Not assigning unique hydropathy values to different atom types results in lowering the resolution of the hydropathy distribution. As we have seen with subtilisin inhibitor, the coarseness in the distribution resulted in missing interaction detail. Including that detail probably would have increased the calculated hydropathy complementarity. The results from small surface areas are particularly subject to error because of poor sampling. A hydropathy scale based on individual atoms is, therefore, greatly needed. Eisenberg and McLachlan[33] have suggested a scale consisting of only four values, not including the rare sulfur atom. We refrained from exploiting this scale for fear that it could not reflect the realistic span of hydropathies in proteins. The efforts of workers such as Abraham and Leo,[34] and Roseman[35] to assign unique hydropathy values for each individual atom, distinguishing, for example, aromatic from aliphatic carbons and carbonyl oxygens from hydroxyl oxygens, will generate a finely tuned and realistic scale that should avoid the hazards of an amino acid-based scale.

Another degree of sophistication would be to consider the actual surface of exposure of each atom as suggested by Eisenberg and McLachlan.[33] For this approach, the atomic hydropathy scale must be normalized on a per Å^2 basis. This means that the hydropathy of an exterior surface is calculated by weighting each atom's hydropathy by the actual area of the atom's van der Waals surface that is accessible to solvent. This area is proportional to the number of contact points a solvent probe sphere can make with the protein atom. This method is readily applicable to solvent-accessible surface hydropathy determinations. However, how does one apply this approach to contact atoms which are defined by their being touched at only one point? Clearly, definitions must be carefully worked out for all other subsets for the sake of maintaining definitional compatibility and consistency. In future applications of this general approach one might also use another

normalization procedure which eliminates from the interior category those parts of amino acid atoms that are always inaccessible.[36] For example, α carbons in free amino acids are almost fully buried. These perpetually buried areas are therefore not indicators of protein conformation and therefore should be removed from the interior category.

CONCLUSIONS

Our findings are summarized as follows. (1) The hydropathy difference between protein interior and exterior, $\Phi_I - \Phi_E$, is an almost invariant quantity and may be a characteristic of all stably folded proteins. (2) The average interprotein interface is more hydrophobic than the rest of the protein's surface, but not more so than the protein as a whole. (3) Most interprotein interfaces manifest specificity, as measured by the function of hydropathy complementarity. It is reasonable to assume that particularly large values for hydropathy complementarity are characteristic of interprotein associations that are inflexible and static. (4) Some protein–protein interactions that are dynamic and flexible in their quaternary organization are bound by hydrophilic forces. These were the capsid proteins of tomato bushy stunt virus and southern bean mosaic virus, hemoglobin, phospholipase A2, and tobacco mosaic virus. (5) Although less preferred than hydrophobic forces, hydrophilic bonds are capable of contributing significant binding force.

Our finding of a constant value for the splitting of a protein's hydropathy into interior and exterior atom subsets is in basic agreement with the observation of Janin et al.[5] on the relatively fixed fractions of nonpolar, charged and polar groups that contribute to the accessible surfaces of proteins. Our survey indicated that the average difference between inside and exterior hydropathies was 0.37, with a relatively small standard deviation of 0.10 among individuals. The hydropathy difference between interior and exterior for monomeric proteins may turn out to be larger than that for the subunits of oligomeric proteins because the latter generally possess external hydrophobic binding surfaces that would be absent in the former. In 1959, Kauzmann suggested that proteins must fold in such a way as to achieve an optimum differential in hydropathy between interior and exterior.[37] Our survey of oligomeric proteins of known structure provides an empirical estimate for that optimum differential. If true, then future investigators may be able to use this value, or a more accurate one, as a guide in predicting the folding of proteins of unknown structure. It has been suggested that such an approach can be ancillary to an energy-based predictive scheme[38] or even replace it.[39]

The energy criterion is also the final arbiter for predicting protein–protein binding. However, when free-energy calculations are difficult, empirical methods can be helpful alternatives. Toward this goal, we have characterized contact surfaces in terms of hydropathy. We showed that contact surfaces are on average more hydrophobic than noninteracting internal surfaces, and that in two out of three cases the binding is overall hydrophobic. Our success in predicting the contact region of a subunit of inorganic pyrophosphatase by visually scanning the surface for hydrophobic clustering justifies some optimism for this approach. At the same time, however, one cannot conclude that binding energy derives only from hydrophobic contacts. The contribution of hydrophilic interactions to binding strength was well demonstrated in our comparisons of two lysozyme–Fab fragment complexes and of the oxy and deoxy states of hemoglobin. Statistically, it seems that hydrophobicity-based interactions are the preferred means of generating binding energy. This contradicts the suggestion[40,41] that hydrophilic bonds have the restricted role of just establishing specificity, while the entropy-based hydrophobic forces determine the free energy of association.

Contact interfaces have also been investigated by Argos[4] and Janin et al.[5] Argos found, in agreement with our observations, that the hydropathy of the average interface lies between those of the exterior and interior. Also valuable is his observation of the predilection of aromatic amino acids for an interface, a fact in concert with their large hydropathy. Janin et al., contradicting us and Argos, concluded that the contact interface is more hydrophobic than even the interior. However, their classification of atoms simply as either nonpolar, polar, or charged, may be an oversimplification.

While hydrophobicity is the main driving force in the binding of multisubunit proteins, there is no guarantee that this is so for other types of protein–protein interactions. The variety of schemes for binding proteins together might be greater than that revealed in our survey, as this was almost totally restricted to multisubunit complexes. Therefore, a number of different criteria for evaluating correct protein docking will be useful. We believe that when the amino acids at a putative interface between two proteins match each other in their hydropathies, there is a very convincing case for predicting that they can bind to each other. The hydropathy complementarity function is a useful empirically based criterion for evaluating the match. As a test for recognition and specificity, it might be most useful in late stages of docking experiments to determine exact positioning once the general binding area has been recognized by less exact means, such as hydrophobic clustering. Other chemical properties, besides hydropathy, could also be used to construct similar additional complementarity tests.

We are confident that as interest in the fundamental characteristics of protein interfaces mounts and

as more investigations are pursued, accurate appraisals of the role of the hydrophobic force will lead to better understanding of protein interactions and their successful prediction.

ACKNOWLEDGMENTS

We are indebted to Dr. Kim Sharp for his indispensible aid in applying the solvent-accessibility program and to Dr. Wayne A. Hendrickson, who generously made available to us his computer molecular graphics facilities. We are grateful to Dr. Phoebe S. Hexem for her insightful criticism of the manuscript. The investigation was supported by a grant from the National Institute of Allergy and Infectious Diseases (AI 17270).

REFERENCES

1. Roberts, M.M., White, J.L., Grütter, M.G., Burnett, R.M. Three-dimensional structure of the adenovirus major coat protein hexon. Science 232:1148–1151, 1986.
2. Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J., Kent, S.B.H. Conserved folding in retroviral proteases: Crystal structure of a synthetic HIV-1 protease. Science 245:616–621, 1989.
3. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
4. Argos, P. An investigation of protein subunit and domain interfaces. Protein Eng 2:101–113, 1988.
5. Janin, J., Miller, S., Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. J. Mol. Biol. 204: 155–164, 1988.
6. Connolly, M.L. Solvent-accessible surfaces of proteins and nucleic acids. Science 221:709–713, 1983.
7. Connolly, M.L. Analytical molecular-surface calculation. J. Appl. Crystallogr. 16:548–558, 1983.
8. Eisenberg, D. Three-dimensional structure of membrane and surface proteins. Annu. Rev. Biochem. 53:595–623, 1984.
9. Klotz, I.M., Darnall, D.W., Langerman, N.R. Ligand-promoted association or dissociation of proteins. In: "The Proteins," Vol. 1, 3rd ed. Neurath, H., Hill, T. (eds.). New York: Academic Press, 1975:293–411.
10. Finzel, B.C., Weber, P.C., Hardman, K.D., Salemme, F.R. Structure of ferricytochrome *c'* from *Rhodospirillum molischianum* at 1.67 Å resolution. J. Mol. Biol. 186:627–643, 1985.
11. Remington, S., Wiegand, G., Huber, R. Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. J. Mol. Biol. 158:111–152, 1982.
12. Weber, I.T., Steitz, T.A. Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. J. Mol. Biol. 198:311–326, 1987.
13. Skarzyński, T., Moody, P.C.E., Wonacott, A.J. Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å resolution. J. Mol. Biol. 193:171–187, 1987.
14. Brayer, G.D., McPherson, A. Refined structure of the gene 5 DNA binding protein from bacteriophage fd. J. Mol. Biol. 169:565–596, 1983.
15. Epp, O., Ladenstein, R., Wendel, A. The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. Eur. J. Biochem. 133:51–69, 1983.
16. Karplus, P.A., Schultz, G.E. Refined structure of glutathione reductase at 1.54 Å resolution. J. Mol. Biol. 195: 701–729, 1987.
17. Sheriff, S., Silverton, E.W., Padlan, E.A., Cohen, G.H., Smith-Gill, S.J., Finzel, B.C., Davies, D.R. Three-dimensional structure of an antibody-antigen complex. Proc. Natl. Acad. Sci. U.S.A. 84:8075–8079, 1987.
18. Amit, A.G., Mariuzza, R.A., Phillips, S.E.V., Poljak, R.J. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. Science 233:747–753, 1986.
19. Jaenicke, R., Helmreich, E. In: "Protein–Protein Interactions." New York: Springer-Verlag, 1972: 96.
20. Stenkamp, R.E., Sieker, L.C., Jensen, L.H. Adjustment of restraints in the refinement of methemerythrin and azidomethemerythrin at 2.0 Å resolution. Acta Crystallogr., Sect. B 39:697–703, 1983.
21. Saul, F.A., Amzel, L.M., Poljak, R.J. Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin New at 2.0 Å resolution. J. Biol. Chem. 253:585–597, 1978.
22. Satow, Y., Cohen, G.H., Padlan, E.A., Davies, D.R. Phosphocholine binding immunoglobulin Fab McPC603: An X-ray diffraction study at 2.7 Å. J. Mol. Biol. 190:593–604, 1987.
23. Deisenhofer, J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8-Å resolution. Biochemistry 20:2361–2370, 1981.
24. Rossmann, M.G., Abad-Zapatero, C., Hermodson, M.A., Erickson, J.W. Subunit interactions in southern bean mosaic virus. J. Mol. Biol. 166:37–83, 1983.
25. Harrison, S.C. Protein interfaces and intersubunit bonding: The case of tomato bushy stunt virus. Biophys. J. 32: 139–153, 1980.
26. Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D., Waley, S.G. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. Nature (London) 255:609–614, 1975.
27. Wright, C.S. Refinement of the crystal structure of wheat germ agglutinin isolectin 2 at 1.8 Å resolution. J. Mol. Biol. 194:501–529, 1987.
28. Campbell, J.W., Watson, H.C., Hodgson, G.I. Structure of yeast phosphoglycerate mutase. Nature (London) 250: 301–303, 1974.
29. Mitsui, Y., Satow, Y., Watanabe, Y., Iitaka, Y. Crystal structure of a bacterial protein proteinase inhibitor (*Streptomyces* subtilisin inhibitor) at 2.6 Å resolution. J. Mol. Biol. 131:697–724, 1979.
30. Keith, C., Feldman, D.S., Deganello, S., Glick, J., Ward, K.B., Jones, E.O., Sigler, P.B. The 2.5 Å crystal structure of dimeric phospholipase A_2 from the venom of *Crotalus atrox*. J. Biol. Chem. 256:8602–8607, 1981.
31. Brunie, S., Bolin, J., Gewirth, D., Sigler, P.B. The refined crystal structure of dimeric phospholipase A_2 at 2.5 Å. Access to a shielded catalytic center. J. Biol. Chem. 260: 9742–9749, 1985.
32. Richards, F.M. Areas, volumes, packing, and protein-structure. Annu. Rev. Biophys. Bioeng. 6:151–176, 1977.
33. Eisenberg, D., McLachlan, A.D. Solvation energy in protein folding and binding. Nature (London) 319:199–203, 1986.
34. Abraham, D.J., Leo, A.J. Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. Proteins 2:130–152, 1987.
35. Roseman, M.A. Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. J. Mol. Biol. 200:513–522, 1988.
36. Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H., Zehfus, M.H. Hydrophobicity of amino acid residues in globular proteins. Science 229:834–838, 1985.
37. Kauzmann, W. Some factors in the interpretation of protein denaturation. Adv. Protein Chem. 14:1–63, 1959.
38. Saitô, N., Shigaki, T., Kobayashi, Y., Yamamoto, M. Mechanism of protein folding: I. General considerations and refolding of myoglobin. Proteins 3:199–207, 1988.
39. Novotný, J., Rashin, A.A., Bruccoleri, R.E. Criteria that discriminate between native proteins and incorrectly folded models. Proteins 4:19–30, 1988.
40. Blundell, T.L. Protein–protein recognition and assembly. In: "Structural Aspects of Recognition and Assembly in Biological Macromolecules." Balaban, M., Sussman, J.L., Traub, W., Yonath, A. (eds.). Rehovot: Balaban International Science Services, 1981:281–286.

41. Chothia, C., Janin, J. Principles of protein–protein recognition. Nature (London) 256:705–708, 1975.
42. Colonna-Cesari, F., Perahia, D., Karplus, M., Eklund, H., Brändén, C.I., Tapia, O. Interdomain motion in liver alcohol dehydrogenase. Structural and energetic analysis of the hinge bending mode. J. Biol. Chem. 261:15273–15280, 1986.
43. Murthy, M.R.N., Reid, T.J. III, Sicignano, A., Tanaka, N., Rossmann, M.G. Structure of beef liver catalase. J. Mol. Biol. 152:465–499, 1981.
44. Carter, D.C., Melis, K.A., O'Donnell, S.E., Burgess, B.K., Furey, W.F. Jr., Wang, B.-C., Stout, C.D. Crystal structure of *Azotobacter* cytochrome c_5 at 2.5 Å resolution. J. Mol. Biol. 184:279–295, 1985.
45. Blevins, R.A., Tulinsky, A. The refinement and the structure of the dimer of α-chymotrypsin at 1.67-Å resolution. J. Biol. Chem. 260:4264–4275, 1985.
46. Shoham, M., Yonath, A., Sussman, J.L., Moult, J., Traub, W., Kalb (Gilboa), A.J. Crystal structure of demetallized concanavalin A: The metal-binding region. J. Mol. Biol. 131:137–155, 1979.
47. Suh, S.W., Bhat, T.N., Navia, M.A., Cohen, G.H., Rao, D.N., Rudikoff, S., Davies, D.R. The galactan-binding immunoglobulin Fab J539: An X-ray diffraction study at 2.6-Å resolution. Proteins 1:74–80, 1986.
48. Frier, J.A., Perutz, M.F. Structure of human foetal deoxyhaemoglobin. J. Mol. Biol. 112:97–112, 1977.
49. Moras, D., Olsen, K.W., Sabesan, M.N., Buehner, M., Ford, G.C., Rossmann, M.G. Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde 3-phosphate dehydrogenase. J. Biol. Chem. 250:9137–9162, 1975.
50. Padlan, E.A., Love, W.E. Refined crystal structure of deoxyhemoglobin S. I. Restrained least-squares refinement at 3.0-Å resolution. J. Biol. Chem. 260:8272–8279, 1985.
51. Fermi, G., Perutz, M.F., Shaanan, B., Forume, R. The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. J. Mol. Biol. 175:159–174, 1984.
52. Shaanan, B. Structure of human oxyhaemoglobin at 2.1 Å resolution. J. Mol. Biol. 171:31–59, 1983.
53. Stenkamp, R.E., Jensen, L.H. Subunit interactions in the metaquohemerythrin octamer. In: "Structural Aspects of Recognition and Assembly in Biological Macromolecules." Balaban, M., Sussman, J.L., Traub, W., Yonath, A. (eds.). Rehovot: Balaban International Science Services, 1981: 197–212.
54. Wilson, I.A., Skehel, J.J., Wiley, D.C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature (London) 289:366–373, 1981.
55. Huber, R., Deisenhofer, J., Colman, P.M., Matsushima, M., Palm, W. Crystallographic structure studies of an IgG molecule and an Fc fragment. Nature (London) 264:415–420, 1976.
56. Adams, M.J., Ford, G.C., Liljas, A., Rossmann, M.G. Atomic co-ordinates for dogfish M_4 apo-lactate dehydrogenase. Biochem. Biophys. Res. Commun. 53:46–51, 1973.
57. Rossmann, M.G., Arnold, E., Erickson, J.W., Frankenberger, E.A., Griffith, J.P., Hecht, H.-J., Johnson, J.E., Kamer, G., Luo, M., Mosser, A.G., Rueckert, R.R., Sherry, B., Vriend, G. Structure of a human common cold virus and functional relationship to other picornaviruses. Nature (London) 317:145–153, 1985.
58. Blake, C.C.F., Geisow, M.J., Oatley, S.J., Rérat, B., Rérat, C. Structure of prealbumin: Secondary, tertiary and quaternary interactions determined by Fourier refinement at 1.8 Å. J. Mol. Biol. 121:339–356, 1978.
59. Arutiunian, E.G., Terzian, S.S., Voronova, A.A., Kuranova, I.P., Smirnova, E.A., Vainstein, B.K., Hohne, W.E., Hansen, G. X-ray diffraction study of inorganic pyrophosphatase from bakers-yeast at 3-Å resolution. Dokl. Akad. Nauk. SSSR 258:1481–1485, 1981.
60. Epp, O., Lattman, E.E., Schiffer, M., Huber, R., Palm, W. The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2.0-Å resolution. Biochemistry 14:4943–4952, 1975.
61. Furey, W., Jr., Wang, B.C., Yoo, C.S., Sax, M. Structure of a novel Bence–Jones protein (Rhe) fragment at 1.6 Å resolution. J. Mol. Biol. 167:661–692, 1983.
62. Arnold, E., Luo, M., Vriend, G., Rossmann, M.G., Palmenberg, A.C., Parks, G.D., Nicklin, M.J.H., Wimmer, E. Implications of the picornavirus capsid structure for polyprotein processing. Proc. Natl. Acad. Sci. U.S.A. 84:21–25, 1987.
63. Tainer, J.A., Getzoff, E.D., Beem, K.M., Richardson, J.S., Richardson, D.C. Determination and analysis of the 2 Å structure of copper, zinc superoxide dismutase. J. Mol. Biol. 160:181–217, 1982.
64. Mitsui, Y., Satow, Y., Watanabe, Y., Hirono, S., Iitaka, Y. Protein–protein interfaces in the complex of subtilisin BPN′ with its protein inhibitor, *Streptomyces* subtilisin inhibitor. In: "Structural Aspects of Recognition and Assembly in Biological Macromolecules." Balaban, M., Sussman, J.L., Traub, W., Yonath, A. (eds.). Rehovot: Balaban International Science Services, 1981: 101–115.
65. Hopper, P., Harrison, S.C., Sauer, R.T. Structure of tomato bushy stunt virus. V. Coat protein sequence determination and its structural implications. J. Mol. Biol. 177:701–713, 1984.
66. Schevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L., Sigler, P.B. The three-dimensional structure of *trp* repressor. Nature (London) 317:782–786, 1985.