8. Kumar, M. & Monteith, J. L. in *Plants and the Daylight Spectrum* (ed. Smith, H.) 133-144 (Academic, London, 1982).
9. Daughtry, C. S. T., Galio, K. P. & Bauer, M. E. *Agron. J.* **75**, 527-531 (1983).
10. Hatfield, J. L., Asrar, G. & Kanemasu, E. T. *Remote Sens. Envir.* **14**, 65-76 (1984).
11. Wiegand, C. L. & Richardson, A. J. *Agron. J.* **76**, 543-548 (1984).
12. Asrar, G., Fuchs, M., Kanemasu, E. T. & Hatfield, J. L. *Agron. J.* **76**, 300-306 (1984).
13. Sellers, P. J. *Int. J. Remote Sens.* **6**, 1335-1372 (1985).
14. Keeling, C. D. *Proc. Carbon Dioxide Research Conference: Carbon Dioxide, Science and Consensus* 2.1-2.62 (Department of Energy Conf.-82097, 1982).
15. Komhyr, W. D. *et al. J. geophys. Res.* **90**, 5567-5596 (1985).
16. Machta, L., Hansen, K. & Keeling, C. D. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Anderson, N. R. & Malahoff, A.) 131-144 (Plenum, New York, 1977).
17. Woodwell, G. M., Houghton, R. A. & Tempel, N. R. *J. geophys. Res.* **78**, 932-940 (1973).
18. Fung, I. Y. in *Proc. 6th Oak Ridge natn. Lab. Life Sci. Symp. on the Global Carbon Cycle* (ed. Jacobs, V.) (US Department of Energy, Oak Ridge, in the press).
19. Heimann, M. M., Keeling, C. D. & Fung, I. Y. in *Proc. 6th Oak Ridge natn. Lab. Life Sci. Symp. on the Global Carbon Cycle* (ed. Jacobs, V.) (US Department of Energy, Oak Ridge, in the press).
20. Junge, C. E. & Czeplak, G. *Tellus* **20**, 422-434 (1968).
21. Machta, L. *Bull. Am. met. Soc.* **53**, 402-420 (1972).
22. Fung, I. Y., Prentice, K., Mathews, E., Lerner, J. & Russel, G. *J. geophys. Res.* **88**, 1281-1294 (1983).
23. Bolin, B. & Keeling, C. D. *J. geophys. Res.* **68**, 3899-3920 (1963).
24. Pearman, G. I. & Hyson, P. *J. geophys. Res.* **85**, 4457-4467 (1980).
25. Azevedo, A. E. thesis, Columbia University (1982).
26. Woodwell, G. M. *et al. Science* **222**, 1081-1086 (1984).
27. Woodwell, G. M. in *The Role of Terrestrial Vegetation in the Global Carbon Cycle: Measurement by Remote Sensing* (ed. Woodwell, G. M.) 1-17 (SCOPE 23, Wiley, New York, 1984).
28. Tucker, C. J., Townshend, J. R. G. & Goff, T. E. *Science* **227**, 369-375 (1985).
29. Justice, C. O., Townshend, J. R. G., Holben, B. N. & Tucker, C. J. *Int. J. Remote Sens.* **6**, 1278-1318 (1985).
30. Goward, S. N., Tucker, C. J. & Dye, D. G. *Vegetacio* (in the press).
31. Tucker, C. J., Vanpraet, C. L., Sharman, M. J. & Van Ittersum, G. *Remote Sens. Envir.* **17**, 233-249 (1985).
32. Tarpley, J. D., Schneider, S. R. & Money, R. L. *J. Climate appl. Met.* **23**, 491-493 (1984).
33. Holben, B. N. & Fraser, R. S. *Int. J. Remote Sens.* **5**, 145-160 (1984).
34. Schutz, C. & Gates, W. L. Preprint (Rand Corporation, R-1029-ARPA, Santa Monica, 1972).
35. Larcher, W. & Bauer, H. in *Physiological Plant Ecology* (eds Lange, O. L., Nobel, P. S., Osmond, C. B. & Ziegler, H.) 403-437 (Springer, Berlin, 1981).
36. Swift, M. J., Heal, O. W. & Anderson, J. M. *Decomposition in Terrestrial Ecosystems*, 308-309 (Blackwell, Oxford, 1979).
37. Linkins, A. E., Melillo, J. M. & Sinsabaugh, R. L. *Current Perspectives in Microbial Ecology* (eds Klug, M. J. & Reddy, C. A.) 572-579 (American Society of Microbiology, Washington, DC, 1984).

Solvation energy in protein folding and binding

David Eisenberg & Andrew D. McLachlan*

Molecular Biology Institute and Department of Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, California 90024, USA
* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

We have developed a method for calculating the stability in water of protein structures, starting from their atomic coordinates. The contribution of each protein atom to the solvation free energy is estimated as the product of the accessibility of the atom to solvent and its atomic solvation parameter. Applications of the method include estimates of the relative stability of different protein conformations, estimates of the free energy of binding of ligands to proteins and atomic-level descriptions of hydrophobicity and amphiphilicity.

OF the forces that guide a polypeptide chain to its folded form in water, solvent interactions, including the hydrophobic interaction, are thought to be among the most important[1-4]. Yet these interactions are currently evaluated by relatively primitive methods. This is in contrast to the other component energies that stabilize proteins, such as hydrogen bonds, for which atom-atom potential functions have been devised[5-7]. Present methods for estimating the contribution of solvation energy to protein stability include the assumption that the hydrophobic character of each amino-acid residue can be summarized by a single number, the amino-acid hydrophobicity[8-12]. This is an oversimplification for amino acids such as Trp, Tyr, Glu, Gln, Lys and Arg which have both polar and apolar parts. Some authors have tried to overcome this limitation by devising more elaborate residue hydrophobicities to improve the description[8,13,14]. Also an oversimplification is the commonly used approximation that the hydrophobic energy of a folded protein molecule is proportional to the total protein surface that is accessible to water, regardless of whether the exposed surface is apolar, polar or charged[15].

To develop an explicitly atomic description of the interaction of water with a protein, we extend the ideas of Langmuir[16], Cohn and Edsall[17], and others[18,19]. The basic assumption is that the free energy of interaction of a solute with water can be considered as a sum of energies of atomic groups. We follow Langmuir in using the exposed surface areas of a group as a measure of its interaction with solvent. In practice we use the solvent-accessible surface area of Lee and Richards[20] and other recent workers[21-24]. This is defined as the area over which the centre of a water molecule of radius 1.4 Å can move while maintaining unobstructed contact with the group. In our method, the sign and strength of the water-solvent interaction are specified by the atomic solvation parameter (ASP) of each atom accessible to water. These values are determined, as described below, from free energies of transfer[8-12]. Thus, our method effectively combines two common approaches for evaluating hydrophobic forces: computation of solvent-accessible surface areas; and estimating energies from free energies of transfer. Our method, however, extends the first approach by weighting the effect of each atom by its polar or apolar character, and extends the second in permitting calculation of the solvation energy from the coordinates of individual atoms in each residue. It advances both methods in permitting estimates of the free energy of transfer of small molecules and of the contribution of solvation to the free energy of binding of small molecules to proteins.

Atomic contributions

Here we express the contribution of protein-solvent interactions to the free energy of protein folding as a sum over all atoms of the structure (except hydrogen atoms, which are not treated explicitly). The term for each atom is the product of its solvent-accessible area[20] and its ASP for transfer from the interior of a protein to aqueous solution. The transfer to solution of a single atom, i, immersed in protein is depicted in Fig. 1*a*. Let the accessible surface area of the atom be given by A_i, and its atomic solvation parameter be given by $\Delta\sigma_i$. Then the free energy of transfer (ΔG) is

$$\Delta G_i = \Delta\sigma_i A_i \quad (1)$$

Now consider the transfer to water of amino-acid residue R immersed in protein (Fig. 1*a*). We assume that the change in free energy for this process can be approximated as a sum of atomic terms, each like that of equation (2)

$$\Delta G_R = \sum_{\text{atoms i}} \Delta\sigma_i A_i \quad (2)$$

Note that the areas A_i now depend on the conformation of the amino acid, and that some atoms will be obscured by their

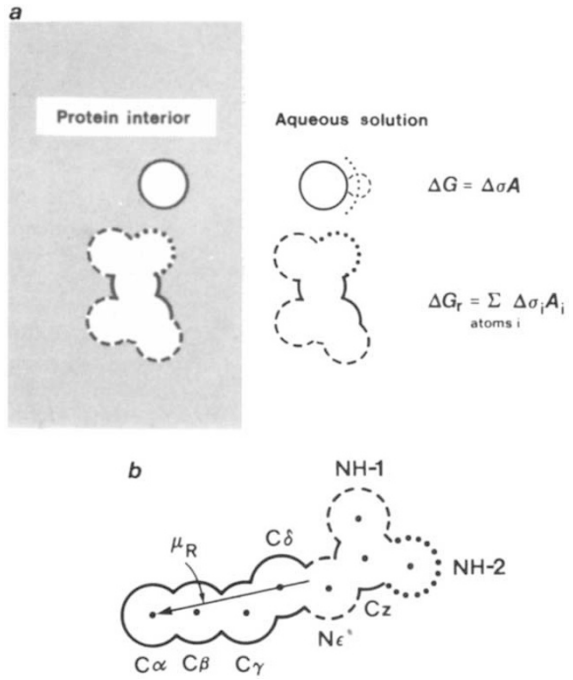

Fig. 1 *a*, Transfer of an atom and a molecule immersed in protein to water. The molecule contains three different types of atoms, represented by solid, dashed and dotted outlines. The definition of accessible surface area is illustrated for the atom at the right, where the centre of a sphere of radius 1.4 Å, tangent to the atom, sweeps out the accessible area. *b*, The atomic hydrophobic moment for the Arg side chain, shown as a vector running from the side-chain centre to the α-carbon atom. The chain takes the conformation specified in the COREL S[40] computer program dictionary. The dotted NH-2 atom is the most exposed nitrogen and is accordingly treated as the charged nitrogen.

neighbours. We assume that conformation is not significantly changed during transfer.

The validity of equation (2) can be tested by a fit to experimental free energies of transfer. For proteins the fit is carried out by considering atoms in five classes: carbon, neutral oxygen and nitrogen (N/O), charged oxygen (O^-), charged nitrogen (N^+) and sulphur. Then the free energies of the common amino acids can be represented by equations of the type

$$\begin{aligned}\Delta G_R = {} & \Delta\sigma(C) \sum_{C\ \text{atoms i}} A(C_i, R) \\ & + \Delta\sigma(N/O) \sum_{N,\,O\ \text{atoms i}} A(N/O_i, R) \\ & + \Delta\sigma(O^-) \sum_{O^-\ \text{atoms i}} A(O_i^-, R) \\ & + \Delta\sigma(N^+) \sum_{N^+\ \text{atoms i}} A(N_i^+, R) \\ & + \Delta\sigma(S) \sum_{S\ \text{atoms i}} A(S_i, R) \qquad (3)\end{aligned}$$

in which $\Delta\sigma(C)$ is the ASP for carbon and $A(C_i, R)$ is the solvent-accessible surface area of carbon atom i in a standard conformation[21] of residue R, and so forth for the other terms. The O^- term describes the charged oxygen atoms in Glu and Asp; the N^+ term describes the charged nitrogen atoms in Lys, Arg and His residues; the N/O term describes all uncharged N and O atoms. In Asp, Glu and Arg the most exposed N or O atom is taken to be the atom on which the charge is concentrated. Only certain terms of equation (3) are needed to describe any given residue. For example, Asn requires only C and N/O terms. The *p*H dependence of charges is ignored, an assumption that might be thought to be poor for His, but the solvent-accessible areas of the N atoms are not large.

Table 1 Amino-acid sovation free energies and group moments

Amino-acid residue	ΔG_{OBS}*	ΔG_R†	ΔG_{TAN}‡	μ_a§	cos θ
Gly	(0)	(0)	(0)	(0)	—
Ala	0.42	0.67	0.50	0	—
Val	1.66	1.5	1.50	0.48	0.84
Leu	2.32	1.9	1.80	1.0	0.89
Ile	2.46	1.9	—	1.2	0.99
Pro	0.98	1.2	—	0.18	0.22
Cys¶	1.34	0.38	—	0.17	0.76
Met	1.68	2.4	1.30	1.9	0.94
Thr	0.35	0.52	0.40	1.5	0.09
Ser	−0.05	0.01	−0.30	0.73	−0.67
Phe	2.44	2.3	2.50	1.1	0.92
Trp	3.07	2.6	3.40	1.6	0.67
Tyr	1.31	1.6	2.30	1.8	−0.93
Asn	−0.82	−0.60	—	1.3	−0.86
Gln	−0.30	−0.22	—	1.9	−1.0
Asp	−1.05	−1.2	—	1.9	−0.98
Glu	−0.87	−0.76	—	3.0	−0.89
His	0.18	0.64	0.50	0.99	−0.75
Lys	−1.35	−0.57	—	5.7	−0.99
Arg	−1.37	−2.1	—	10.0	−0.96

Group	ΔG_{OBS}‖ (ref. 17)	ΔG_R# (equation (3))
$-CH_2-$	0.67	0.56
$-C_6H_5$	1.90	1.6
-OH	−0.76	−0.66
$-CH_2-CONH-$	−1.14	0.66
Gly ionization	−3.7	−2.6

Observed and calculated free energies of transfer are given in units of kcal mol^{-1}, calculated residue hydrophobic moments μ_a are given in kcal mol^{-1} Å. ΔG_{OBS} and ΔG_{TAN} are experimental free energies of transfer relative to glycine. ΔG_R is the calculated value; μ_a is the calculated residue hydrophobic moment; and θ is the angle between the direction of the residue hydrophobic moment and a line from the α-carbon nucleus to the geometric centre of each side chain.

* ΔG_{OBS} is the observed value of π reported by Fauchere and Pliska[25] multiplied by 2.30 RT = 1.36 kcal mol^{-1}; π(side chain) is defined as log D(acetyl amino-acid amide) − log D(acetyl glycine amide), where D is the distribution coefficient for octanol/water.

† Calculated from equation (3).

‡ Observed values for transfer from ethanol to water from ref. 9.

§ Atom-based hydrophobic moment, defined in equation (5), below, using solvent-accessible surface areas from ref. 21 and atomic coordinates from ref. 39. These moments depend on the detailed conformation of the side chain, but not sensitively, because the corresponding values based on the coordinate dictionary supplied with the computer program COREL S[40] differ by an average of only 3%.

‖ Experimental values from ref. 17 (from p. 212 except for the value for the -OH group which is from p. 206). Gly ionization is for the process: $NH_2-CH_2-COOH \rightarrow {}^+NH_3-CH_2-COO^-$.

Calculated from equation (3) as follows: $-CH_2-$, using the average accessible are of 8 $-CH_2$ groups in 5 residues; Gly ionization from tabulated areas[21]; and the -OH group as ΔG_R(Ser) − ΔG_R(Ala) and $-C_6H_5-$ as ΔG_R(Phe)-G_R(Ala), corresponding to the experimental values[17].

¶ Both ΔG_{OBS} and ΔG_R refer to half-cystine. We were unable to derive a reliable value for -SH.

The assumption in equation (3) is that 20 residue hydrophobicities (ΔG_R) can be represented by five ASPs ($\Delta\sigma$s), provided that suitable values of $\Delta\sigma$ can be found. We now show that a consistent set of $\Delta\sigma$s can be derived from experimental data on free energies of transfer. The accessible areas A_i of residues in standard conformations can be calculated from known structures; we have adopted the areas calculated by Shrake and Rupley[21], which are close to those found by C. Chothia (personal communication); they are for the amino acid X in a Gly-X-Gly sequence in a typical extended conformation. The values of the $\Delta\sigma$s were determined by a linear least-squares fit of equations (3) to observed values for the free energy of

transfer, ΔG_R, of the amino-acid side chains. The measured ΔGs were those determined by Fauchere and Pliska[25] from transfer free energies of amino-acid residue analogues from *n*-octanol to water. Our estimated ASP values (with their standard deviations) are:

$$\Delta\sigma(\mathrm{C}) = \ 16 \pm 2 \text{ cal Å}^{-2} \text{ mol}^{-1}$$
$$\Delta\sigma(\mathrm{N/O}) = -6 \pm 4$$
$$\Delta\sigma(\mathrm{O}^-) = -24 \pm 10$$
$$\Delta\sigma(\mathrm{N}^+) = -50 \pm 9$$
$$\Delta\sigma(\mathrm{S}) = \ 21 \pm 10$$

The quality of the fit is illustrated in Fig. 2, and the observed and calculated values are given in Table 1. The fit is generally good except for Lys and Cys. Note that the non-polar atoms C and S increase the free energy of the system as they are transferred from the interior of the protein to water. Polar atoms decrease the free energy in the same process; charged atoms cause a much larger decrease.

Solvation contribution

With $\Delta\sigma$ values, it is possible to estimate the solvation contribution to the free energy of protein folding, ΔG_s. This energy with respect to a reference state *r* is given by

$$\Delta G_s = \Delta\sigma(\mathrm{C}) \sum_{\text{C atoms i}} (A_i - A_i^r) + \Delta\sigma(\mathrm{N/O}) \sum_{\text{N, O atoms i}} (A_i - A_i^r) + \Delta\sigma(\mathrm{O}^-) \sum_{\text{O}^- \text{ atoms i}} (A_i - A_i^r) + \Delta\sigma(\mathrm{N}^+) \sum_{\text{N}^+ \text{ atoms i}} (A_i - A_i^r) + \Delta\sigma(\mathrm{S}) \sum_{\text{S atoms i}} (A_i - A_i^r) \quad (4)$$

in which A is the solvent-accessible surface area of an atom in the folded state and A^r is that in the reference state[21]. In the usual way in which equation (4) is used, the reference state does not matter because the energies of the same polypeptide chain are compared in two conformations having a common reference state. This is true for the application described below.

Improperly folded structures

One potential use of equation (4) is to assess the stabilities of modified or redesigned proteins. To illustrate such a calculation, the models of Novotný *et al.*[1] for incorrectly folded proteins are useful. These models were generated by exchanging the 113 side chains on the correctly folded α-carbon backbone of haemerythrin[26] from *T. dyscritum* with those from the correctly folded, β-sheet backbone of an immunoglobulin variable light-chain (V_L) domain[27] having the same number of residues. Novotný *et al.* used an energy-minimizing computer program to make small annealing adjustments in coordinates, and to reduce the potential energies of the two incorrectly folded structures to local minima. These annealed structures then had satisfactory van der Waals contacts, with normal valence angles and bond lengths, but no explicit account had been taken of solvation effects. Novotný *et al.*[1] noted that there is greater exposure of apolar residues in the misfolded structures. To these model misfolded structures, as well as to their energy-annealed, correctly folded analogues, we have applied equation (4) to estimate the solvation contribution to the free energy of folding. Our procedure consists simply of: first, computing[22] solvent-accessible atomic areas of the four structures from atomic coordinates; and second, weighting each atomic area by its ASP, and summing.

The solvation free energy is lower for the correct structures (Table 2). Both real structures have solvation free energies lower

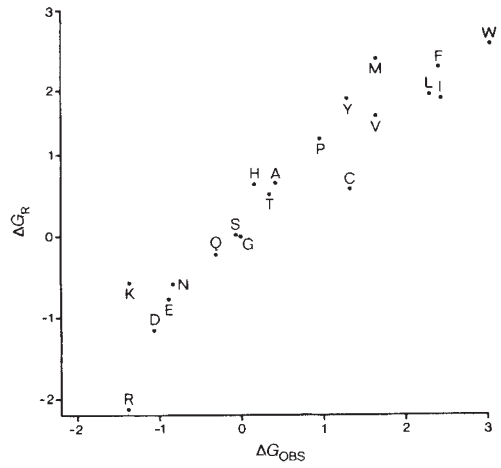

Fig. 2 Comparison of experimental values ΔG_{OBS} for free energies of transfer for amino-acid residues with the values of ΔG_R calculated from equation (3), using ASPs determined by linear least squares. The details of the least-squares fit are as follows: Fauchere and Pliska[25] deduced the hydrophobic parameter π for 19 amino-acid side chains as the difference between the logarithm of the measured distribution coefficients of the acetyl amino-acid amides and acetyl-glycine amide. Equation (3) was fitted to the free energies of transfer, $2.303\ RT\pi = \Delta G_{OBS}$, with the aid of NAG library routines G02BFF and G02CHF (Numerical Algorithms Group Ltd, Mayfield House, 256 Banbury Road, Oxford OX2 7DE, UK). In this fitting procedure, the difference between ΔG_R and ΔG_{OBS} is minimized by least-squares determination of the ASP values. The values of A for the appropriate atoms in each side chain are taken from ref. 21. For all residues, only the differences in area with the Gly residue are considered. Thus backbone atoms are ignored, except for the difference in exposed area for the α-carbon atom of glycine and of the other side chains. The deviation of points of the figure from a straight line can be explained by any of: (1) errors in measurements of ΔG_{OBS}; (2) inappropriate areas in equation (3); or (3) faulty assumptions in the formulation of equation (3) (see below). The multiple correlation coefficient R is 0.95, and the value of its square corrected for the number of parameters (ASPs), R^2_{COR}, is 0.88. The R^2_{COR} suggests that the separate ASPs for O^- and N^+ are justified, as is the sulphur ASP: with only a single type of charged atom, R^2_{COR} falls to 0.85 and R to 0.93; for only C/S and N/O ASPs, R^2_{COR} falls to 0.11 and R to 0.45. Fits to other hydrophobicity scales give different values for the ASPs. For example, the consensus scale of Eisenberg *et al.*[12], an average of five scales, yields $R = 0.97$ for a five-term fit. The ASP values for the consensus fit are similar to those reported here, but the O^- and N^+ ASPs have larger values. Despite the better fit to the consensus scale, the ASPs based on the data of Fauchere and Pliska[25] have been used in our calculations, in part because the absolute level of the consensus scale is unknown and in part because the new ASPs for the charged atoms seem more reasonable.

than the hypothetical, unfolded reference state by some 100 kcal mol^{-1}, ~1 kcal per residue. Moreover, the solvation free energies for the real structures are markedly lower than for the incorrectly folded ones. Not only does each correctly folded structure have lower solvation free energy than the same polypeptide sequence folded improperly, but also each correct structure has lower free energy than the other sequence folded with the same backbone pattern. These results can be contrasted with the behaviour of traditional energy terms[1], which are not significantly different in the folded and misfolded structures.

When contributions to the solvation energy of folding for the natural structures are examined, interesting details emerge. For example, in the haemerythrin structure, 95 of the 113 side chains stabilize the solvation energy. Only one side chain, Asp 11, is destabilizing by so much as 1 kcal mol^{-1}. For this side chain, both oxygen atoms are somewhat shielded from solvent relative to the extended reference structure. But for the protein as a whole, the process of folding from the extended reference state

Table 2 Solvation free energies of folding ΔG_s for an immunoglobulin V_L domain and haemerythrin

	V_L immunoglobulin domain	Haemerythrin
Correct fold	−106	−113
Incorrect fold (with the same sequence)	−72	−96
Net stabilization of correct structure	−34	−17

Values are calculated for the natural structures and the incorrectly folded structures (with exchanged amino-acid sequences) constructed by Novotný *et al.* Energies are given in kcal mol^{-1}. Equation (4) is used, with areas determined by the program ACCESS[22]. The uncertainty in the net stabilization of the correct structure was estimated from the standard deviations in the ASPs and by assuming that the fractional uncertainty in the sums of the atomic areas of each type is 0.05; this results in an overall uncertainty of ~10 kcal mol^{-1}.

yields −159 kcal for sequestering the 606 C and 3 S atoms away from the solvent (the sum of the first and last terms of equation (4)), while the partial covering of the 297 N/O, 17 O^-, and 21 N^+ atoms consumed only 25, 12 and 11 kcal mol^{-1}, respectively. Significant contributions to the stability of the folded state come from the apolar carbon atoms of the 29 Lys, Arg, Asp and Glu residues (a total of 41 kcal mol^{-1}). These contributions would be obscured or neglected in energetic models that are cast in terms of whole residue hydrophobicities; clearly, whole-residue hydrophobicities cannot account adequately for the folding energies of proteins.

Agreement with experimental data

Although it is reassuring that the estimated free energy of folding is lower for the real protein structures than for the incorrectly folded chains, this result falls well short of experimental verification of the method proposed here. Here we examine its assumptions and compare its predictions with experiment.

The principal assumption in the use of equation (3) to derive the ASPs is that the free energy of transfer is a linear function of the accessible areas of the constitutent atoms of the transferred molecule. This assumption is supported indirectly by the observed linearity of free energy with surface area for the transfer of many compounds from various non-aqueous liquids into water[15,28-31]. Of more direct bearing on the assumption is the accumulated evidence that molecular free energies of transfer can often be reckoned as a sum of energies for atomic groups[19]. This is not a new idea. Cohn and Edsall[17] made a critical analysis of evidence which suggested that the free energy of transfer of a chemical group such as $-CH_2-$, from ethanol and other apolar solvents to water is often independent of the molecule in which it is found. Some values for the free energy of transfer for various groups as compiled by Cohn and Edsall are given in Table 1, where they are compared with our new values calculated from equation (3). Generally, the computed values agree well in sign and magnitude with the experimental values, especially considering the diversity of the groups considered. The calculated values for side-chain transfers are also in reasonable agreement with the well-known measured values of Nozaki and Tanford[9] (Table 1). An exception to agreement in Table 1 is the free energy of transfer of the glycyl residue, for which equation (3) gives the wrong sign. This is probably because our model takes virtually no account of the partial charges on the amide N and O atoms. The reason is that the backbone atoms are omitted in the fitting procedure used to determine the ASPs, because the data being fitted are all differences between acetyl glycyl amide and the corresponding derivative of other amino acids. From the value in Table 2 for the free energy of transfer of the glycyl residue from ethanol to water, it is possible to estimate a crude value for the ASP for the amide N and O atoms at ~ −50 cal $Å^{-2}$ mol^{-1}. If this value is qualitatively correct, then our solvation energies of folding are somewhat overestimated. To resolve this uncertainty, it will be necessary to fit a larger database of comparable free energies of transfer, which is not yet available.

An additional assumption in the use of equation (4) is that a non-polar liquid is an adequate model for the interior of a protein, as the ASPs derived from transfers between liquids are applied by equation (4) to transfer from liquid to protein interior. This assumption has been examined by Richards[24], who also considered amino-acid crystals as models for the protein interior. Richards concluded that an ASP somewhere in the range of 13-26 cal $Å^{-2}$ mol^{-1} is appropriate for the transfer of apolar atoms. Our $\Delta\sigma$(C) and $\Delta\sigma$(S) values of 16 and 21 cal $Å^{-2}$ mol^{-1} lie in this range, but may deserve revision as further free-energy data become available. The value for $\Delta\sigma$(S), resting on only two measurements, is particularly subject to revision.

Atomic hydrophobic moments

In earlier work we described the amphiphilicity (asymmetry of hydrophobicity) of segments of regular protein secondary structure in terms of hydrophobic moments[32]. It was found that the hydrophobic dipole moments of neighbouring segments of secondary structure tend to oppose each other in correctly folded proteins[12], but not in incorrectly folded ones[33], and that hydrophobic moments can be used to classify helices on the basis of amino-acid sequences[34,35] and to detect periodicities[36,37]. These applications of hydrophobic moments were all formulated in terms of residue moments and can perhaps be made more precise with atomic-level moments.

The hydrophobic moment can be defined as an atomic property using the ASP values determined here. The hydrophobic dipole moment of a group of atoms is given by a sum over atoms i

$$\boldsymbol{\mu}_a = \sum_{\text{atoms i}} \Delta\sigma_i A_i \boldsymbol{r}_i - \langle \Delta\sigma A \rangle \sum_{\text{atoms i}} \boldsymbol{r}_i \qquad (5)$$

in which $\boldsymbol{r}_i$ is a vector from any origin to the position of atom i, and where the brackets indicate the mean value for all atoms of the group. The second term ensures that the value of $\boldsymbol{\mu}_a$ is independent of the choice of origin. When the sum in equation (5) is restricted to the atoms in a single side chain, a residue hydrophobic dipole moment is defined. This quantity is a measure of the amphiphilicity of the side chain, as distinct from its hydrophobicity. The values of these moments are tabulated in Table 1.

From these values it is apparent that the residues of greatest intrinsic amphiphilicity are Arg, Lys and Glu. In contrast, the most hydrophobic residues, Trp, Phe, Leu and Ile, all have small amphiphilicities. The direction of the hydrophobic moment is also defined by equation (5). This is expressed in Table 1 as the cosine of the angle between the direction of the moment and the direction of the vector from the α carbon to the side-chain centre. For the highly amphiphilic residues, the direction of the moment is nearly antiparallel to the α-carbon side-chain vector ($\cos\theta$ near to −1). This is illustrated in Fig. 1*b* for Arg. Some of the most hydrophobic side chains (for example, Ile, Met and Phe) have moments nearly parallel to the α-carbon-centre vector ($\cos\theta$ near to +1). Note that when the phenolic oxygen (with negative $\Delta\sigma$) is added to Phe to form Tyr, it reverses the direction of the moment (from +0.92 to −0.93).

Implications and extensions

How does the solvation free energy defined in equation (4) correspond to the component of the free energy of protein folding first described by Kauzmann as hydrophobic[3]? Kauzmann suggested that exposed apolar side chains increase the free energy of the system because they decrease the entropy of water. This effect on the free energy is described in equation (4) by the terms $\Delta\sigma$(C) and $\Delta\sigma$(S). Kauzmann's approach is extended in equation (4) by including the terms $\Delta\sigma$(N/O),

$\Delta\sigma(O^-)$ and $\Delta\sigma(N^+)$. This has been done to introduce information from empirical free-energy measurements in a self-consistent way, and to impose a penalty in free energy for burial of charged and polar groups with a magnitude determined from experiments. These terms probably include contributions from electrostatic[38] and hydrogen-bonding effects as well. A full understanding of these contributions would require a knowledge of the structure of the molecules used to derive the ASPs, including states of protonation and hydration in both solvents. In short, ΔG_s includes contributions to the free energy of folding that go beyond the entropy change of the solvent. For this reason we call this energy the solvation free energy, rather than the hydrophobic free energy.

One use of equation (4) is in assessing the stability of proposed new proteins. These can either be genetically redesigned proteins or proteins designed for peptide synthesis. The energy of any model described in atomic coordinates can be tested against equation (4), although it may be necessary to supplement this equation with terms for other components of the folding energy. Another use is in energy-coupled X-ray phase refinement of protein structures[39-42]. In existing procedures, a model is determined that is the best compromise between the potential energy of the structure and the fit to the X-ray data. These methods do not go far enough in that it is the free energy, including solvation effects, that determines the structure. These effects could be incorporated by using ASPs as suggested here, along with the analytical method of Richmond[23], for computing accessible surface areas and their spatial derivatives. Similarly, the addition of a solvation energy term of the form of equation (4) could be made to molecular dynamics calculations. Finally, we note that the general method of computing solvation energy expressed in equation (4) could be developed further to compute the solvation contribution to the free energy for association of proteins with membranes or nucleic acids.

We thank Drs Novotný, Brucolleri and Karplus for making available the coordinates of the incorrectly and correctly folded protein structures described in ref. 1; Dr C. Chothia for his values of accessible surface areas of amino acids in extended polypeptide chains; Dr Richmond for use of his computer program ACCESS; and Drs D. Abrahams, F. R. Salemme and C. Schutt for discussions. D.E. acknowledges a Guggenheim fellowship and support from NIH and NSF. A computer program to calculate solvation energy according to equation (4) is available from the authors.

Received 19 July; accepted 30 October 1985.

1. Novotný, J., Bruccoleri, R. & Karplus, M. *J. molec. Biol.* **177**, 787-818 (1984).
2. Kauzmann, W. *Adv. Protein Chem.* **14**, 1-63 (1959).
3. Creighton, T. E. *Proteins* (Freeman, San Francisco, 1984).
4. Tanford, C. *The Hydrophobic Effect* 2nd edn (Wiley, New York, 1980).
5. Karplus, M. & McCammon, A. *A. Rev. Biochem.* **52**, 263-300 (1983).
6. Nemethy, G. & Scheraga, H. A. *Q. Rev. Biophys.* **10**, 239-352 (1977).
7. Levitt, M. & Warshel, A. *Nature* **253**, 694-698 (1975).
8. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. H. *Science* **229**, 834-838 (1985).
9. Nozaki, Y. & Tanford, C. *J. biol. Chem.* **246**, 2211-2217 (1971).
10. Janin, J. *Nature* **277**, 491-492 (1979).
11. Wolfenden, R., Anderson, L., Cullis, P. M. & Southgate, C. C. B. *Biochemistry* **20**, 849-855 (1981).
12. Eisenberg, D., Weiss, R. M., Terwilliger, T. C. & Wilcox, W. *Faraday Symp. chem. Soc.* **17**, 109-120 (1982).
13. Guy, H. R. *Biophys. J.* **47**, 61-70 (1985).
14. Frommel, C. *J. theor. Biol.* **111**, 247-260 (1984).
15. Chothia, C. *Nature* **248**, 338-339 (1974).
16. Langmuir, I. *Colloid Symp. Monogr.* **3**, 48-75 (1925).
17. Cohn, E. J. & Edsall, J. T. in *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Ch. 9 (ACS Monograph Series, Reinhold, New York, 1943).
18. Tanford, C. *J. Am. chem. Soc.* **84**, 4240-4247 (1962).
19. Hansch, C. & Leo, A. in *Substitution Constants for Correlation Analysis in Chemistry and Biology* (Wiley, New York, 1979).
20. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379-400 (1971).
21. Shrake, A. & Rupley, J. A. *J. molec. Biol.* **79**, 351-371 (1973).
22. Richmond, T. J. & Richards, F. M. *J. molec. Biol.* **119**, 537-555 (1978).
23. Richmond, T. J. *J. molec. Biol.* **178**, 63-89 (1984).
24. Richards, F. M. *A. Rev. Biophys. Bioengng* **6**, 151-176 (1977).
25. Fauchere, J.-L. & Pliska, V. *Eur. J. med. Chem.-Chim. ther.* **18**, 369-375 (1983).
26. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. *Acta crystallogr.* **B38**, 784-792 (1978).
27. Segal, D. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 4298-4302 (1974).
28. Hermann, R. B. *J. phys. Chem.* **76**, 2754-2759 (1972).
29. Nandi, P. K. *Int. J. Peptide Protein Res.* **8**, 253-264 (1976).
30. Gelles, J. & Klapper, M. H. *Biochim. biophys. Acta* **533**, 465-477 (1978).
31. Creighton, T. E. *J. molec. Biol.* **129**, 235-264 (1979).
32. Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. *Nature* **299**, 371-374 (1982).
33. Eisenberg, D., Wilcox, W. & McLachlan, A. D. *Proc. UCLA Symp. Protein Structure, Folding and Design* (in the press).
34. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. *J. molec. Biol.* **179**, 125-142 (1984).
35. Pownall, H. J., Knapp, R. D., Gotto, A. M. Jr & Massey, J. B. *FEBS Lett.* **159**, 17-23 (1983).
36. Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. *Proc. natn. Acad. Sci. U.S.A.* **81**, 140-144 (1984).
37. Finer-Moore, J. & Stroud, R. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 155-159 (1984).
38. Warshel, A. & Russell, S. T. *Q. Rev. Biophys.* **17**, 283-422 (1984).
39. Diamond, R. *J. molec. Biol.* **82**, 371-391 (1974).
40. Susman, J. L., Holbrook, S. R., Church, G. M. & Kim, S.-H. *Acta crystallogr.* **A33**, 800-804 (1977).
41. Levitt, M. *J. molec. Biol.* **82**, 393-420 (1974).
42. Konnert, J. H. *Acta crystallogr.* **A32**, 614-617 (1976).

LETTERS TO NATURE

Inadequacy of Coulomb's friction law for particle assemblies

K. Kendall

ICI New Science Group, PO Box 11, The Heath, Runcorn, Cheshire WA7 4QE, UK

Wherever powders are manipulated by mechanical means, adhesive and frictional forces act between the solid grains. For two centuries, these forces have been described by Coulomb's law[1], a law which originally proved useful in designing civil engineering structures, and which has since been profitably extended to the understanding of the powder flows in chemical plants[2]. However, we have now found that Coulomb's law is not adequate to explain the compaction of fine powders in the manufacture of ceramic articles, where it has been observed that friction appears to increase for smaller particles. This apparent increase in friction is explained by considering in detail the contact of loaded, adhesive, elastic spheres.

Consider the ideal experiment (Fig. 1*a*) in which a bed of particles is loaded onto and sheared along a polished plate of the same material[3]. The nominal shear stress F/A (where F is force and A is area) required to slide the powder across the plate under the normal stress W/A (where W is normal load) has usually been interpreted by Coulomb's law:

$$F/A = K + \mu W/A \qquad (1)$$

where K represents the cohesive shear strength of the particle/plate assembly (that is, the shear stress experienced at zero normal load) and μ is the coefficient of friction between the surfaces[4]. This law, in combination with the geometrical constraints on powder motion first mentioned by Reynolds in 1885, has been much used to describe the deformation of powders and soils in complex stress states[5-7]. Here we demonstrate that equation (1) does not follow the experimental results satisfactorily, and we propose a new argument which provides an improved fit to the observed behaviour.

The problem of applying Coulomb's law to assemblies of fine particles is that the friction coefficient calculated from equation (1) does not generally match that measured in a standard friction test between macroscopic blocks of material. First, the calculated μ is often higher than expected: for example, sand particles of diameter 0.5 mm in water give $\mu = 0.45$ whereas two large quartz blocks give $\mu = 0.3$ (Fig. 1*b*). A second difficulty is that the friction coefficient increases as the grains are made smaller,