6
Weak Interactions in Protein Folding: Hydrophobic Free Energy, van der Waals Interactions, Peptide Hydrogen Bonds, and Peptide Solvation

Robert L. Baldwin

6.1
Introduction

Hydrophobic free energy has been widely accepted as a major force driving protein folding [1, 2], although a dispute over its proper definition earlier made this issue controversial. When a hydrocarbon solute is transferred from water to a nonaqueous solvent, or a nonpolar side chain of a protein is buried in its hydrophobic core through folding, the transfer free energy is referred to as hydrophobic free energy. The earlier dispute concerns whether the transfer free energy can be legitimately separated into two parts and the free energy of hydrophobic hydration treated separately from the overall free energy change [3–5]. If the hydrophobic free energy is defined as the entire transfer free energy [5], then there is general agreement that transfer of the nonpolar solute (or side chain) out of water and into a nonaqueous environment drives folding in a major way. A related concern has come forward, however, and scientists increasingly question whether the energetics of forming the hydrophobic core of a protein should be attributed chiefly to packing interactions (van der Waals interactions, or dispersion forces) rather than to burial of nonpolar surface area. This question is closely related to the issue of whether the hydrophobic free energy in protein folding should be modeled by liquid–liquid transfer experiments or by gas–liquid transfer experiments.

The energetic role of peptide hydrogen bonds (H-bonds) was studied as long ago as 1955 [6] but the subject has made slow progress since then, chiefly because of difficulty in determining how water interacts with the peptide group both in the unfolded and folded forms of a protein. Peptide H-bonds are likely to make a significant contribution to the energetics of folding because there are so many of them: about two-thirds of the residues in folded proteins make peptide H-bonds [7]. Peptide backbone solvation can be predicted from electrostatic algorithms but experimental measurements of peptide solvation are limited to amides as models for the peptide group.

This chapter gives a brief historical introduction to the ''weak interactions in protein folding'' and then discusses current issues. It is not a comprehensive review and only selected references are given. The term ''weak interaction'' is somewhat misleading because these interactions are chiefly responsible for the folded struc-

Protein Folding Handbook. Part I. Edited by J. Buchner and T. Kiefhaber

ISBN: 3-527-30784-2

tures of proteins. The problem of evaluating them quantitatively lies at the heart of the structure prediction problem. Although there are methods such as homology modeling for predicting protein structures that bypass evaluation of the weak interactions, de novo methods of structure prediction generally rely entirely on evaluating them. Thus, the problem of analyzing the weak interactions will continue to be a central focus of protein folding research until it is fully solved.

6.2
Hydrophobic Free Energy, Burial of Nonpolar Surface and van der Waals Interactions

6.2.1
History

The prediction in 1959 by Walter Kauzmann [1] that hydrophobic free energy would prove to be a main factor in protein folding was both a major advance and a remarkable prophecy. No protein structure had been determined in 1959 and the role of hydrophobic free energy in structure formation could not be deduced by examining protein structures. The first protein structure, that of sperm whale myoglobin, was solved at 2 Å resolution only in 1960 [8]. On the other hand, the predicted structure of the α-helix [9] given by Pauling and coworkers in 1951, which was widely accepted, suggested that peptide H-bonds would prove to be the central interaction governing protein folding. Peptide H-bonds satisfied the intuitive belief of protein scientists that the interactions governing protein folding should be bonds with defined bond lengths and angles. This is not a property of hydrophobic free energy.

Kauzmann [1] used the ambitious term ''hydrophobic bonds,'' probably aiming to coax protein scientists into crediting their importance, while Tanford [10] introduced the cautious term ''the hydrophobic effect.'' ''Hydrophobic interaction'' has often been used because a factor that drives the folding process should be an interaction. However, hydrophobic interaction is also used with a different meaning than removal of nonpolar surface from contact with water, namely the direct interaction of nonpolar side chains with each other. The latter topic is discussed under the heading ''van der Waals interactions.'' The term ''hydrophobic free energy'' is used here to signify that nonpolar groups help to drive the folding process. Tanford [10] points out that a hydrophobic molecule has both poor solubility in water and good solubility in nonpolar solvents. Thus, mercury is not hydrophobic because it is insoluble in both solvents. Early work leading to the modern view of hydrophobic free energy is summarized by Tanford [11] and a recent discussion by Southall et al. [12] provides a valuable perspective.

6.2.2
Liquid–Liquid Transfer Model

Kauzmann [1] proposed the liquid–liquid transfer model for quantitating hydrophobic free energy. His proposal was straightforward. Hydrophobic molecules

prefer to be in a nonpolar environment rather than an aqueous one and the free energy difference corresponding to this preference should be measurable by partitioning hydrocarbons between water and a nonaqueous solvent. Nozaki and Tanford [13] undertook a major program of using the liquid–liquid transfer model to measure the contributions of nonpolar and partially nonpolar side chains to the energetics of folding. They measured the solubilities of amino acids with free α-COO^- and α-NH_3^+ groups, while Fauchère and Pliska [14] later studied amino acids with blocked end groups, because ionized end groups interfere with the validity of assuming additive free energies of various groups. They measured partitioning of solutes between water and *n*-octanol (saturated with water), which is less polar than the two semi-polar solvents, ethanol and dioxane, used by Nozaki and Tanford [13]. Wimley et al. [15] used a pentapeptide host to redetermine the partition coefficients of the amino acid side chains between water and water-saturated *n*-octanol and they obtained significantly different results from those of Fauchère and Pliska. They emphasize the effect of neighboring side chains (''occlusion'') in reducing the exposure of a given side chain to water. Radzicka and Wolfenden [16] studied a completely nonpolar solvent, cyclohexane, and observed that the transfer free energies of hydrocarbons are quite different when cyclohexane is the nonaqueous solvent as compared to *n*-octanol.

In Figure 6.1 the transfer free energies of model compounds for nonpolar amino

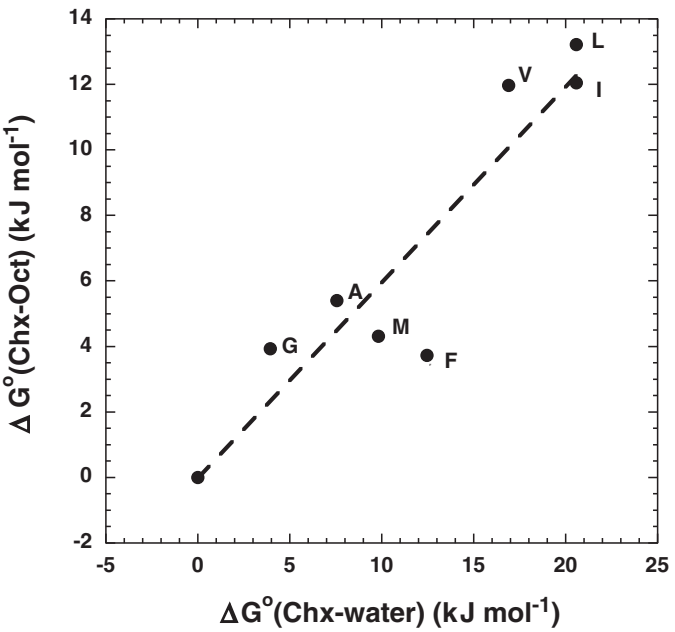

Fig. 6.1. Transfer free energies from cyclohexane to water compared with ones from cyclohexane to water-saturated *n*-octanol (data from Ref. [16]). The model solutes undergoing transfer represent the amino acid side chains shown on the plot. Note that the transfer free energies between cyclohexane and *n*-octanol are more than half as large as those from cyclohexane to water.

acid side chains are compared using either cyclohexane or *n*-octanol as the nonaqueous solvent [16]. If different nonaqueous solvents may be used equally well to model the hydrophobic core of a protein, then the transfer free energies of hydrocarbons from cyclohexane to *n*-octanol should be small compared with the transfer free energies from cyclohexane to water. Figure 6.1 shows this is not the case: the transfer free energies measured between cyclohexane and *n*-octanol are more than half as large as the ones between cyclohexane and water. Thus, these results pose the first serious question about the use of the liquid–liquid transfer model: which nonaqueous solvent should be used to model the hydrophobic core and how valid are the results if no single solvent is a reliable model?

6.2.3
Relation between Hydrophobic Free Energy and Molecular Surface Area

A second important step in quantifying hydrophobic free energy was taken when several authors independently observed that the transfer free energy of a nonpolar solute is nearly proportional to the solute's surface area for a homologous series of solutes [17–19]. This observation agrees with the intuitive notion that the transfer free energy of a solute between two immiscible solvents should be proportional to the number of contacts made between solute and solvent (however, see Section 6.2.6). Lee and Richards [20] in 1971 developed an automated algorithm for measuring the water-accessible surface area (ASA) of a solute by rolling a spherical probe, with a radius equivalent to that of a water molecule (1.4 Å) (10 Å = 1 nm), over its surface. Their work showed how to make use of the surface area of a solute to analyze its hydrophobicity. Proportionality between transfer free energy and ASA does not apply to model compounds containing polar side chains because polar groups interact strongly and specifically with water.

A plot of transfer free energy versus ASA is shown in Figure 6.2 for linear alkanes. The slope of the line for linear (including branched) hydrocarbons is 31 cal mol^{-1} $Å^{-2}$ (1.30 J mol^{-1} nm^{-2}) when partition coefficients on the mole fraction scale [21] are used. Earlier data for the solubilities of liquid hydrocarbons in water are used to provide the transfer free energies. In Figure 6.2 the transfer free energy is nearly proportional to the ASA of the solute. The line does not pass through (0,0), but deviation from strict proportionality is not surprising for small solutes [22].

Hermann [22] points out that linear hydrocarbons exist in a broad range of configurations in solution and each configuration has a different accessible surface area. He also points out [17] that the transfer free energy arises from a modest difference between the unfavorable work of making a cavity in a liquid and the favorable van der Waals interaction between solute and solvent. Consequently, a moderate change in the van der Waals interaction can cause a large change in the transfer free energy. Tanford [10] analyzes the plot of transfer free energy versus the number of carbon atoms for hydrocarbons of various types, and discusses data for the different slopes of these plots.

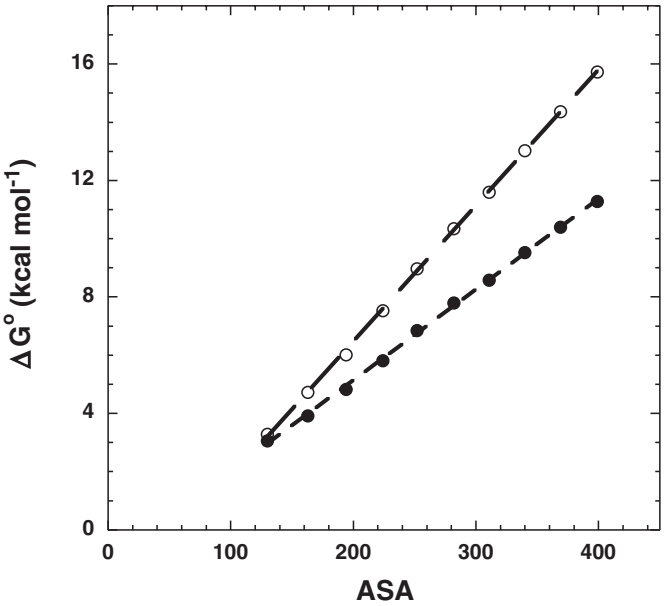

Fig. 6.2. Transfer free energies of linear alkanes (from 1 to 10 carbon atoms) from liquid alkane to water, measured from hydrocarbon solubility in water. They are plotted against water-accessible surface area (ASA in $Å^2$). Data are from Ref. [21]. The uncorrected transfer free energies (filled circles) refer to the mole fraction scale while the corrected values (open circles) refer to the molarity scale and are corrected for the ratio of molecular volumes, solute/solvent, according to Sharp et al. [21]. The data here and in Figure 6.3 are given in kcal mol^{-1} to conform with the literature on this subject. Note that the plots do not pass through 0,0 and note the larger slope (47 cal mol^{-1} $Å^{-2}$) of the corrected plot versus the uncorrected plot (31 cal mol^{-1} $Å^{-2}$).

6.2.4
Quasi-experimental Estimates of the Work of Making a Cavity in Water or in Liquid Alkane

In modern solution chemistry, the solvation free energy of a solute is defined as its transfer free energy from the gas phase into the solvent, when the appropriate standard state concentration (1 M) is used in each phase, as specified by Ben-Naim and Marcus [23]. Gas–liquid transfer free energies are used to analyze the nature of liquid–liquid transfer free energies. The reason for adopting the 1 M standard state concentration in both the gas and liquid phases is to ensure that the density of the solute is the same in both phases at the standard state concentration. Then the transfer free energy gives the free energy change for transferring the solute from a fixed position in the gas phase to a fixed position in the liquid phase [23]. The solute in the gas phase can be treated as an ideal gas [24, 25]

and the nonideal behavior of real gases at 1 M concentration can be omitted from consideration.

Modern theories of solvation indicate that the gas–liquid transfer process can be formally divided into two steps of an insertion model of solvation: see discussion by Lee [24, 25] and basic theory by Widom [26]. In step 1 thermal fluctuations create a cavity in the liquid with a size and shape appropriate for containing the solute. The structure of the liquid undergoes reorganization to make the cavity [24]. In step 2 the solute is inserted into the cavity, van der Waals interactions occur between the solute and the solvent, and the solvent structure undergoes further reorganization at the surface of the cavity. Lee [24] determines quasi-experimental values for the entropy and enthalpy changes in the two steps of the insertion model: (1) making a cavity in the liquid, and (2) inserting the solute into the cavity. Experimental transfer data are used for each of five alkanes undergoing transfer from the gas phase to the liquid phase, either to water or to neat liquid alkane. The transfer thermodynamics then are combined with literature estimates for the van der Waals interaction energies, obtained by Jorgensen and coworkers [27] from Monte-Carlo simulations. The results give quasi-experimental estimates of the enthalpy and entropy changes in each step of the insertion model. Table 6.1 gives these values for the free energy cost of making a cavity to contain the alkane solute both in water and in liquid alkane. The 1977 theory of hydrophobic solvation by Pratt and Chandler [28] divides the process of solvation into the two steps of cavity formation and solute insertion, and the authors consider the rules for separating the solvation process into these two steps. The two steps of dissolving an alkane in water have also been simulated by molecular dynamics and the results analyzed by the free energy perturbation method [29].

The following conclusions can be drawn from Lee's data [24, 25]. (1) The work of making a cavity in water is much larger than in liquid alkane and this difference

Tab. 6.1. Quasi-experimental estimates of the free energy of cavity formation and simulation-based results for van der Waals interaction energies between solvent and solute[a].

Hydrocarbon	ΔG_c*(water)*	ΔG_c*(alkane)*	E_a*(water)*	E_a*(alkane)*
Methane	20.4	9.6	−12.1	−14.7
Ethane	27.7	18.1	−20.1	−26.8
Propane	35.8	23.7	−27.6	−34.4
Isobutane	42.4	24.4	−32.7	−35.1
Neopentane	46.5	27.6	−36.0	−39.6

[a] Values in kJ mol^{-1}. ΔG_c is the free energy cost of making a cavity in the solvent to contain the hydrocarbon solute, from a study by Lee [24]. E_a is the van der Waals interaction energy between solute and solvent; values from Lee [24], based on parameters from a Monte-Carlo simulation study by Jorgensen and coworkers [27]. For water as solvent, conditions are 25 °C, 1 atm; for neat hydrocarbon as solvent, either the temperature or pressure is chosen that will liquefy the hydrocarbon.

is the major factor determining the size of the hydrophobic free energy in liquid–liquid transfer. For example, it costs 46.5 kJ mol^{-1} to make a cavity for neopentane in water but only 27.6 kJ mol^{-1} in liquid neopentane. To understand hydrophobic free energy, it is necessary first of all to understand the free energies of cavity formation in water and in nonpolar liquids. The work of making a cavity in water is large because it depends on the ratio of cavity size to solvent size [30, 31] and water is a small molecule (see Section 6.2.7). It is more difficult to make a cavity of given size by thermal fluctuations if the solvent molecule is small. (2) The work of making a cavity in a liquid is chiefly entropic [24], while the van der Waals interactions between solute and solvent are enthalpic. (3) The van der Waals interaction energy between an alkane solute and water is nearly the same as between the alkane solute and liquid alkane (see Table 6.1). For example, the interaction energies between neopentane and water versus neopentane and liquid neopentane are -36.0 kJ mol^{-1} and -39.6 kJ mol^{-1} [24]. Earlier, Tanford [32] used interfacial tensions to show that the attractive force between water and hydrocarbon is approximately equal, per unit area, to that between hydrocarbon and hydrocarbon. Scaled particle theory predicts well the work of making a cavity either in water or in liquid alkane [24], but it predicts only semi-quantitatively the enthalpy of solvent reorganization for these cavities.

6.2.5
Molecular Dynamics Simulations of the Work of Making Cavities in Water

In 1977 a physico-chemical theory of hydrophobic free energy by Pratt and Chandler [28] already gave good agreement between predicted and observed transfer free energies of linear alkanes, both for gas–liquid and liquid–liquid transfer. Molecular dynamics simulations can be used to obtain the free energy cost of cavity formation in liquids and the results are of much interest because they basically depend only on the specific water model used for the simulations. It should be kept in mind that the physical properties of water used as constraints when constructing water models do not normally include surface tension, and consequently good agreement between the predicted and known surface tension of water is not necessarily to be expected. (For macroscopic cavities, the work of making a cavity equals surface tension times the surface area of the cavity.) In 1982 Berendsen and coworkers [33] determined the free energy of cavity formation in water for cavities of varying size and compared the results to values predicted by scaled particle theory, with reasonable agreement. Remarkably, the comparison with scaled particle theory also gave a value for the surface tension of water close to the known value. Because of the importance of the problem, simulations of cavity formation in water by molecular dynamics continued in other laboratories (see references in [34]). An important result is the development by Hummer and coworkers [34] of an easily used information-theory model to represent the results for water in the cavity size range of interest. Some applications of the information theory model are mentioned below.

6.2.6
Dependence of Transfer Free Energy on the Volume of the Solute

Evidence is discussed in Section 6.2.3 that the transfer free energy is correlated with the surface area (ASA) of the solute. Because it is straightforward to compute ASA [20] from the structure of a peptide or protein, this correlation provides a very useful means of computing the change in hydrophobic free energy that accompanies a particular change in conformation. In recent years, evidence has grown, however, that the transfer free energy of a nonpolar solute depends on its size and shape for reasons that are independent of hydrophobic free energy. In 1990 DeYoung and Dill [35] brought the problem forcibly to the attention of protein chemists by demonstrating that the transfer free energy of benzene from liquid hydrocarbon to water depends on the size and shape of the liquid hydrocarbon molecules. Section 6.2.2 reviews evidence that liquid–liquid transfer free energies depend on the polarity (and perhaps on water content) of the nonaqueous solvent. But in the study by DeYoung and Dill [35] the size and shape of the nonaqueous solvent molecules affect the apparent hydrophobic free energy. A large literature has developed on this subject and recently Chan and Dill [36] have provided a comprehensive review.

Chandler [37] briefly discusses the reason why solvation free energy in water depends on the volume of a sufficiently small nonpolar solute. This dependence can be found in both the information-theory model [34] and the Lum-Chandler-Weeks theory [38] of hydrophobic solvation. Effects of the size and shape of the solute are taken into account in the Pratt and Chandler theory [28].

Stimulated by the results of DeYoung and Dill [35], Sharp and coworkers [21] used a thermodynamic cycle and an ideal gas model to relate the ratio of sizes, solute to solvent, to the transfer free energy for gas–liquid transfer. They conclude that the transfer free energy depends on the ratio of solute/solvent molecular volumes. Their paper has generated much discussion and controversy. In 1994 Lee [39] gave a more general derivation for the transfer free energy, based on statistical mechanics, and considered possible assumptions that will yield the result of Sharp et al. [21].

The Lum-Chandler-Weeks theory of hydrophobic solvation [38] predicts a crossover occurring between the solvation properties of macroscopic and microscopic cavities when the cavity radius is 10 Å. Huang and Chandler [40] point out that the ratio of the work of making a cavity in water to its surface area reaches a plateau value for radii above 10 Å, and this value agrees with the known surface tension of water at various temperatures. On the other hand, the hydrophobic free energy found from hydrocarbon transfer experiments increases slightly with temperature (see Section 6.2.9), implying that the work of making a sufficiently small cavity in water increases with temperature. Chandler [37] explains that these two different outcomes, which depend on solute size, arise naturally from the H-bonding properties of water, because the sheath of water molecules that surrounds a nonpolar solute remains fully H-bonded when the solute is sufficiently small but

not when the solute radius exceeds a critical value. Southall and Dill [41] find that a highly simplified model of the water molecule (the ''Mercedes-Benz'' model), which reproduces several remarkable properties of water, also predicts such a transition from microscopic to macroscopic solvation behavior.

The question of interest to protein chemists is: should a transfer free energy be corrected for the ratio of solute/solvent volumes or not? Figures 6.2 and 6.3 compare the uncorrected with the volume-corrected plots of transfer free energy versus ASA, for both liquid–liquid transfer (Figure 6.2) and gas–liquid transfer (Figure 6.3). Both correlations show good linearity. However, the hydrophobic free energy corresponding to a given ASA value is substantially larger if the volume-corrected transfer free energy is used (see Ref. [21]). Whether the volume correction should be made remains controversial. Scaled particle theory emphasizes the role of surface area in determining the free energy of cavity formation while the information-theory model [34] and the Lum-Chandler-Weeks theory [38] both emphasize the

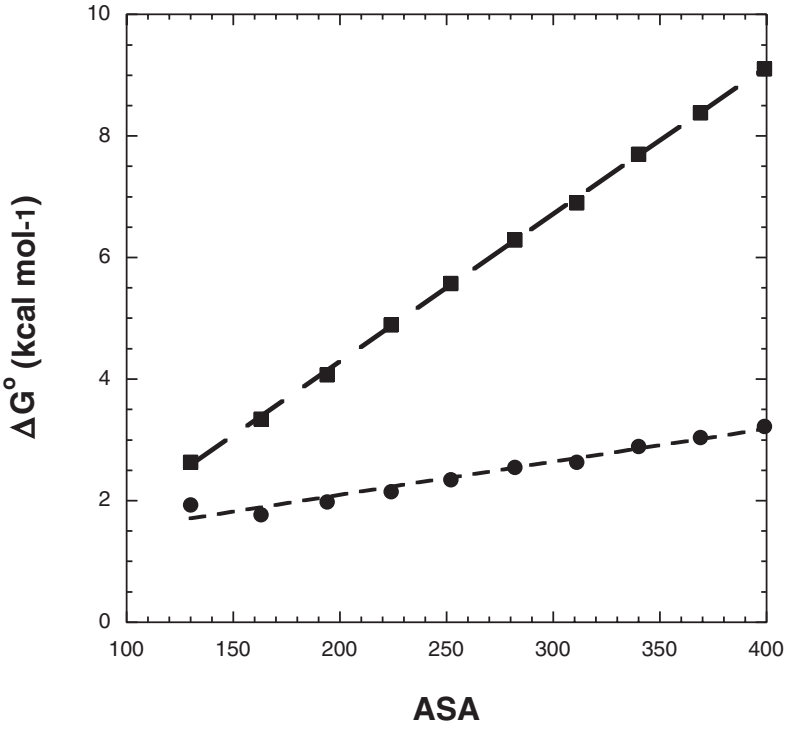

Fig. 6.3. Transfer free energy from the gas phase to liquid water for linear alkanes (from 1 to 10 carbon atoms) plotted against water-accessible surface area (in $Å^2$). Data are from Ref. [21]. The uncorrected transfer free energies (filled circles) refer to the mole fraction scale while the corrected values (filled squares) refer to the molarity scale and are volume-corrected according to Ref. [21]. The slopes of the lines are 5.5 cal mol^{-1} $Å^{-2}$ (uncorrected) and 24 cal mol^{-1} $Å^{-2}$ (corrected).

role of molecular volume. The molecular dynamics simulations of Berendsen and coworkers [33] were interpreted by scaled-particle theory while those of Hummer and coworkers were interpreted by the information-theory model [34], which has much in common with the Lum-Chandler-Weeks theory. Pohorille and Pratt [42] give a detailed discussion of how the scaled particle interpretation [33] may be reconciled with their own analysis.

6.2.7 Molecular Nature of Hydrophobic Free Energy

The molecular nature of hydrophobic free energy has been controversial for a long time [3, 4, 11, 12]. A long-standing proposal, supported by liquid–liquid transfer data at 25 °C [1] and by simulation results [27], is that the arrangement of water molecules in the solvation shell around a dissolved hydrocarbon is entropically unfavorable. Consequently the unfavorable entropy change for dissolving a hydrocarbon in water should provide the driving force for expelling the solute from water. This proposal required modification when it was learned that the hydrophobic free energy found from liquid–liquid transfer is purely entropic only at an isolated temperature near room temperature [43]. Hydrophobic free energy becomes increasingly enthalpy-driven as the temperature increases and it becomes entirely enthalpic upon reaching its maximum value at a temperature above 100 °C (see Section 6.2.9). The characteristic property of hydrophobic free energy that dominates its temperature-dependent behavior is the large positive value of ΔC_p, the difference between the heat capacities of the hydrocarbon in water and in the nonaqueous solvent.

In contradiction to the thesis developed by Kauzmann [1], Privalov and Gill [3, 4] proposed that the hydration shell surrounding a dissolved hydrocarbon tends to stabilize the hydrocarbon in water while the van der Waals interactions between the hydrocarbon solute and the nonaqueous solvent account for the hydrophobicity of the solute. They assume that the van der Waals interaction energy between solute and solvent is large in the nonaqueous solvent compared to water. However, Lee's data (see Table 6.1) show that the large work of making a cavity in water is responsible for the hydrophobic free energy while a hydrocarbon solute makes nearly equal van der Waals interactions with water and with a liquid hydrocarbon [24]. Privalov and Gill coupled two proposals: (1) the van der Waals interactions between nonpolar side chains drive the formation of the hydrophobic core of a protein, and (2) the hydration shell surrounding a dissolved hydrocarbon tends to stabilize it in water. The latter proposal is now widely believed to be incorrect but there is increasing interest in the first proposal.

There are restrictions both on the possible orientations of water molecules in the solvation shell around a hydrocarbon solute and on the hydrogen bonds they form [12]. Models suggesting how these restrictions can explain the large positive values of ΔC_p found for nonpolar molecules in water have been discussed as far back as 1985, in Gill's pioneering study of the problem [44]. Water-containing clathrates of nonpolar molecules surrounded by a single shell of water molecules have been

crystallized and their X-ray structures determined [45]. The water molecules form interconnected 5- and 6-membered rings. Water molecules in ice I are oriented tetrahedrally in a lattice and the oxygen atoms form six-membered rings [45].

In 1985 Lee [30] used scaled particle theory to argue that the low solubilities of nonpolar solutes in water, and the magnitude of hydrophobic free energy, depend strongly on the solute/solvent size ratio, and he reviews prior literature on this subject. Rank and Baker [31] confirmed his conclusion with Monte-Carlo simulations of the potential of mean force between two methane molecules in water. The remarkable temperature-dependent properties of hydrophobic free energy (see Section 6.2.9) are determined, however, chiefly by ΔC_p, which depends on the hydrogen bonding properties of water according to most authors (see Ref. [46]).

6.2.8
Simulation of Hydrophobic Clusters

Formation of the hydrophobic core of a protein during folding must proceed by direct interaction between nonpolar side chains. Yet direct interaction between two hydrocarbon molecules in water is known to be extremely weak (compare Ref. [28]). Benzene dimers or complexes between benzene and phenol in water are barely detectable [47]. Raschke, Tsai and Levitt used molecular dynamics to simulate the formation of hydrophobic clusters, starting from a collection of isolated hydrocarbon molecules in water [48]. Their results give an interesting picture of the thermodynamics of the process. The gain in negative free energy from adding a hydrocarbon molecule to a hydrocarbon cluster in water increases with the size of the cluster until limiting behavior is reached for large clusters. The simulations of cluster formation yield a proportionality between transfer free energy and burial of nonpolar surface area that is similar to the one found from liquid–liquid transfer experiments [48]. The simulation results make the important point that a standard molecular force field is able to simulate the thermodynamics of hydrophobic free energy when a hydrophobic cluster is formed in water [48]. Rank and Baker [49] found that solvent-separated hydrocarbon clusters precede the desolvated clusters found in the interior of large hydrocarbon clusters. Thus, a hydrocarbon desolvation barrier may be important in the kinetics of protein folding [49]. In both simulation studies [48, 49], the authors find that the molecular surface area (defined by Richards [50]) is more useful than water-accessible surface area in analyzing cluster formation.

6.2.9
ΔC_p and the Temperature-dependent Thermodynamics of Hydrophobic Free Energy

Although the gas–liquid transfer model is now often used instead of the liquid–liquid transfer model to analyze hydrophobic free energy, nevertheless thermodynamic data for the transfer of liquid hydrocarbons to water are remarkably successful in capturing basic thermodynamic properties of hydrophobic free energy in protein folding. Figure 6.4A shows ΔH versus temperature for the process

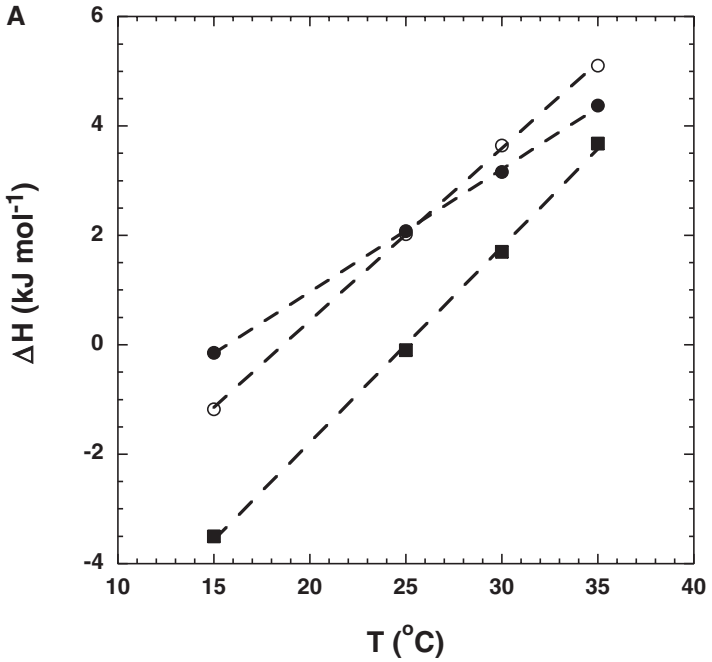

Fig. 6.4. (A) Enthalpy of transfer plotted against temperature for three liquid hydrocarbons (benzene (filled circles), ethylbenzene (open circles) and cyclohexane (filled squares)) undergoing transfer from liquid hydrocarbon to water. Data are from Ref. [51]. Note that the plots are straight lines in this temperature range, where ΔC_p = constant, and ΔH passes through 0 near room temperature.

of transferring three hydrocarbon solutes from neat liquid hydrocarbon to water, taken from the calorimetric study by Gill and coworkers [51]. These results illustrate the large positive values of ΔC_p found when nonpolar molecules are dissolved in water, a property discovered by Edsall [52] in 1935. Although ΔH depends strongly on temperature, ΔC_p (the slope of ΔH versus T) is nearly constant in the temperature range 15–50 °C (see figure 12 of Ref. [3]). ΔC_p decreases perceptibly at temperatures above 50 °C [3, 4].

The following thermodynamic expressions take a simple form when ΔC_p is constant. A strong dependence of ΔH on temperature must be accompanied by a strong dependence of ΔS° on temperature. When ΔC_p is constant, then:

$$\Delta H(T_2) = \Delta H(T_1) + \Delta C_p(T_2 - T_1) \tag{1}$$

$$\Delta S^\circ(T_2) = \Delta S^\circ(T_1) + \Delta C_p T_2 \ln(T_2/T_1) \tag{2}$$

Figure 6.4B compares $\Delta H, T\Delta S^\circ$ and ΔG° as functions of temperature for the transfer of benzene from liquid benzene to water. (The behavior of a liquid alkane,

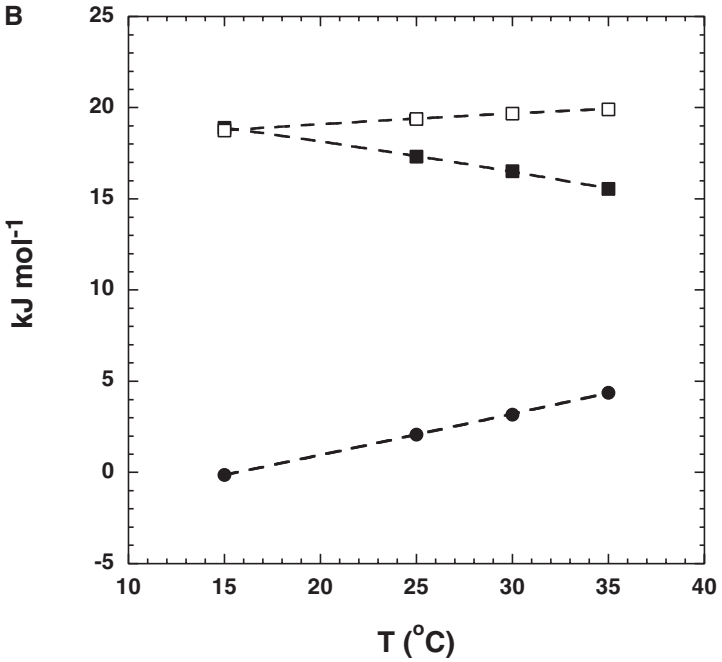

Fig. 6.4. (B) Transfer free energy and the contributions to free energy from ΔH (filled circles) and $T\Delta S$ (filled squares), plotted against temperature, for the transfer of benzene from liquid benzene to water. Data are from Ref. [43]. Note that the relative contribution to ΔG (open squares) from ΔH increases with temperature while the contribution of $T\Delta S$ decreases, and ΔG increases only slightly with temperature. The plots illustrate how entropy–enthalpy compensation affects hydrophobic free energy.

pentane, is fairly similar to that of the aromatic hydrocarbon benzene, see figure 12 of Ref. [3].) Substituting Eqs (1) and (2) into the standard relation

$$\Delta G^\circ = \Delta H - T\Delta S^\circ \tag{3}$$

gives

$$\Delta G^\circ(T_2) = \Delta G^\circ(T_1) + \Delta C_p(T_2 - T_1) - T_2\Delta C_p \ln(T_2/T_1) \tag{4}$$

Equation (4) shows that the ΔC_p-induced changes in ΔH and ΔS° with temperature tend to compensate each other to produce only a small net increase in $-\Delta G^\circ$ as the temperature increases. This property of enthalpy–entropy compensation is one of the most characteristic features of hydrophobic free energy. (ΔG and ΔS depend strongly on the solute concentration and the superscript $^\circ$ emphasizes that ΔG° and ΔS° refer to the standard state concentration.)

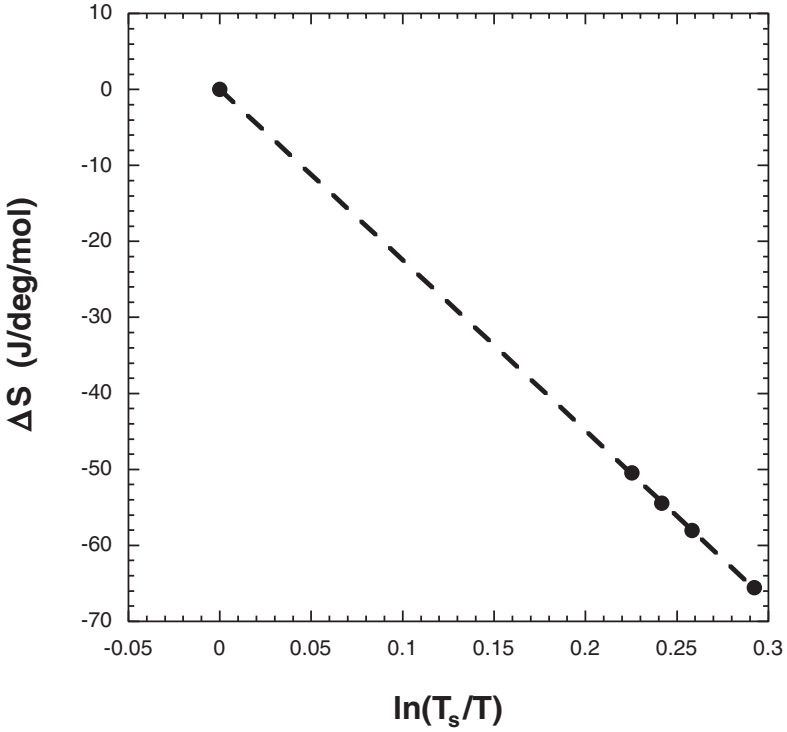

Fig. 6.5. Decrease of $-\Delta S$ towards 0 as temperature (K) approaches T_s (386 K), the temperature at which $\Delta S = 0$. The transfer of benzene from liquid benzene to water is shown. Data are from Ref. [43]. Note the linear decrease in $-\Delta S$ with temperature [43], which depends on using data from the temperature range (15–50 °C) where ΔC_p = constant (see figure 12 of Ref. [3]). $T_s = 386$ K is an average value for several hydrocarbon solutes [43].

Figures 6.4A and 6.4B illustrate some basic thermodynamic properties of hydrophobic free energy as modeled by liquid–liquid transfer. The enthalpy change ΔH is zero near room temperature (the exact temperature depends on the hydrocarbon) and the transfer process is entropy-driven around room temperature. However, as the temperature increases above 25 °C, the transfer process gradually becomes enthalpy-driven. A surprising property is that of entropy convergence: if ΔC_p = constant, then different hydrocarbons share a convergence temperature at which $\Delta S^\circ = 0$ (386 K or 113 °C) [43]. Data taken between 15 and 35 °C are extrapolated linearly versus ln T in Figure 6.5, according to Eqn (2). When the gradual decrease in ΔC_p at temperatures above 50 °C [3] is taken into account by using a curved extrapolation, hydrocarbon data for ΔS° still approach a common value near T_s, the temperature at which the entropy change is 0, which is approximately 140 °C [3, 4]. The property of entropy convergence is predicted both by the information-theory model [53] and scaled particle theory [54].

Privalov [55] discovered in 1979 that values for the specific entropy change on protein unfolding converge near 113 °C when results for some different proteins

are extrapolated linearly versus temperature. His observation, taken together with the entropy convergence shown by hydrocarbon solutes, suggests that the hydrophobic entropy change of a protein unfolding reaction might be removed by extrapolating the total entropy change to T_s for hydrocarbon transfer [43]. However, when Robertson and Murphy [56] analyzed unfolding data from several laboratories, they found more scatter in the data and they did not find a common intercept for the specific entropy change on protein unfolding. The extrapolation offers a possible route to determining the change in conformational entropy on unfolding, one of the basic unsolved problems in protein folding. The hydrophobic contribution to the entropy of unfolding is opposite in sign and, at room temperature, comparable in size to the change in conformational entropy [43].

Schellman [57] points out that there are advantages to using $\Delta G/T$ as an index of stability instead of ΔG itself when considering the temperature dependence of hydrophobic free energy. The low solubilities of hydrocarbons in water have minimum values near room temperature, corresponding to the maximum values of $\Delta G/T$ and to $\Delta H = 0$, whereas ΔG itself increases steadily but slowly with temperature until it reaches a maximum value above 100 °C, at a temperature where $\Delta S = 0$.

A standard test of whether the hydrophobic free energy found from liquid–liquid transfer experiments applies to protein folding experiments is to compare the unfolding free energy change caused by a ''large to small'' mutation with the free energy change predicted from the change in ASA on unfolding. Pace [58] surveyed the literature on this subject (see also Ref. [21]). He concludes there is good agreement provided the proportionality coefficient between $\Delta\Delta G$ and ΔASA is the volume-corrected value 47 cal mol^{-1} $Å^{-2}$ [21]. However, this type of comparison between liquid–liquid transfer and protein folding thermodynamics is complicated by the presence of cavities produced when large-to-small mutations are made [59].

A direct test of the correspondence between liquid–liquid transfer thermodynamics and protein folding is to compare the value of ΔC_p for a protein unfolding experiment with the value predicted from the change in ASA on unfolding. No other factor besides the burial of nonpolar surface area is known to make a significant positive contribution to ΔC_p. This comparison was studied in 1991 by Record and coworkers [60], who found that values of ΔC_p measured in unfolding experiments do agree with the ones predicted from the change in nonpolar ASA on unfolding. The issue became somewhat complicated, however, by contemporary work showing that polar groups (and especially the peptide group) contribute negatively to ΔC_p for unfolding. Makhatadze and Privalov [61] gave extensive model compound data and used them to estimate contributions to the enthalpy and heat capacity changes on unfolding of all groups present in proteins. Freire and coworkers [62] used empirical calibration of protein unfolding data to give the contributions to ΔC_p expected from the changes in polar and nonpolar ASA on unfolding. The issue is seriously complicated, however, by evidence from the study of model compounds that the enthalpies of interaction with water are not related in any simple way to polar ASA: see Table 6.2 and Section 6.3.2. To make progress in analyzing this problem, it is important to get data for the ΔC_p values accompanying unfold-

ing of peptide helices and progress has been reported recently by Richardson and Makhatadze [63].

A long-standing puzzle in comparing the thermodynamics of protein unfolding with those of liquid–liquid transfer has been the unfolding results produced by high pressures [64]. The information-theory model [65] provides new insight into the problem by predicting that water penetrates the hydrophobic cores of proteins during pressure-induced unfolding, as observed experimentally [66], but does not penetrate the hydrophobic cores during thermal unfolding.

6.2.10 Modeling Formation of the Hydrophobic Core from Solvation Free Energy and van der Waals Interactions between Nonpolar Residues

As explained above, solvation free energies of solutes are based on gas–liquid transfer data. They provide an alternative model for the energetics of forming the hydrophobic core of a protein. Compared with the liquid–liquid transfer model, this approach has two major advantages. (1) It considers the van der Waals interactions explicitly, and (2) it avoids the question of which nonaqueous solvent to use for modeling the hydrophobic core. This approach was pioneered by Ooi and Oobatake [67, 68] and by Makhatadze and Privalov [61]. Simonson and Brünger [69] analyze model compound data for gas–liquid transfer. They report that transfer free energies for gas to liquid transfer of cyclic hydrocarbons fall well below the line for linear alkanes when plotted against ASA.

The use of gas–liquid transfer to probe the energetics of folding is illustrated in the cycle shown in Scheme 6.1, in which the native protein (N) is formed by folding the unfolded protein (U) either in the gas phase (g) or in aqueous solution (aq). The cycle is completed by the process of transferring U from aqueous solution to the gas phase and by transferring N from the gas phase back to aqueous solution. Only the nonpolar side chains and nonpolar moieties of polar side chains are considered here. Transfer of the polar peptide NH and CO groups between the gas phase and aqueous solution is considered in Section 6.3.8, together with formation of the peptide H-bonds. The polar moieties of polar side chains are supposed to be fully exposed to water in both U and N, and so their transfer between aqueous solution and the gas phase cancels energetically in steps 1 and 3. The change in conformational entropy when the unfolded polypeptide U folds to the native protein N is supposed to be the same in steps 2 and 4.

In Scheme 6.1, two processes contribute to ΔG°_{UN}, the standard free energy

$$\begin{array}{ccc} \mathrm{U\ (g)} & \Leftarrow (2) \Rightarrow & \mathrm{N\ (g)} \\ \Uparrow\Downarrow (1) & & \Uparrow\Downarrow (3) \\ \mathrm{U\ (aq)} & \Leftarrow (4) \Rightarrow & \mathrm{N\ (aq)} \end{array}$$

Scheme 6.1

change for folding in aqueous solution. Process A is burial (removal from water) of nonpolar side chains (or moieties) in steps 1 and 3, and process B is formation of van der Waals interactions between nonpolar side chains (or moieties) in step 2. Let ΔG_{bur} and ΔG_{vdw} be the contributions to ΔG°_{UN} from processes A and B in Eq. (5).

$$\Delta G^{\circ}_{UN} = \Delta G_{bur} + \Delta G_{vdw} \tag{5}$$

Let ΔASA be the net change in nonpolar solvent-accessible surface area in steps $1 + 3$. Evidence is presented above that ΔG_{bur} can be related to ΔASA even for microscopic cavities. Thus ΔG_{bur} can be represented by

$$\Delta G_{bur} = \beta \Delta ASA \tag{6}$$

in which β is the proportionality coefficient between ΔG°_{solv} and ASA for alkanes in transfer experiments from the vapor phase to aqueous solution. Figure 6.3 shows good linearity between ΔG°_{solv} and ASA for several alkanes, and β is reasonably constant. However, the value of β changes substantially (from 5.5 to 24 cal mol^{-1} $Å^{-2}$ [21]) when volume-corrected transfer free energy is used. Evaluation of ΔG_{vdw} depends sensitively on parameters that are difficult to determine experimentally, and there is little discussion in the literature of how mutations cause ΔG_{vdw} to vary. See, however, the discussion of ΔG_{vdw} values for selected proteins by Makhatadze and Privalov [61] and note that the *number* of van der Waals contacts is being discussed in mutational studies of packing [70].

For the process of forming the hydrophobic core, ΔG_{vdw} can be estimated approximately in the following manner. First, consider the empirical relation between ΔG°_{UN} and ΔASA given by the liquid–liquid transfer model.

$$\Delta G^{\circ}_{UN} \sim B(\Delta ASA) \tag{7}$$

In Eq. (7), ΔASA is the net value of ASA buried upon folding and B is the proportionality factor between liquid–liquid transfer free energy and ASA. Combining Eqns (5)–(7) gives:

$$\Delta G_{vdw} \sim (B - \beta)\Delta ASA \tag{8}$$

The value of $(B - \beta)$ depends on whether or not the volume correction is made to the transfer free energies. If the volume correction is made, then $(B - \beta) = 47 - 25 = 22$ cal mol^{-1} $Å^{-2}$ [21] and ΔG_{vdw} should account for approximately $22/47 = 47\%$ of the free energy change on forming the hydrophobic core. If the volume correction is not made and the mole fraction scale is used for computing transfer free energies, then $(31 - 5)/31$ [21] $= 84\%$, the percentage of ΔG°_{UN} that is assigned to ΔG_{vdw}. This argument follows the one given by Havranek and Harbury [71], who used different numbers for B and β from the ones given by Sharp et al. [21].

Data for the melting of solid crystalline alkanes show that ΔH is substantial and favors the solid alkane, but ΔG for melting is small because there is a substantial ΔS that favors the liquid alkane [72]. These results have been used to argue that a liquid hydrocarbon should model satisfactorily the hydrophobic core of a protein [72], and therefore the van der Waals interactions need not be considered explicitly.

6.2.11
Evidence Supporting a Role for van der Waals Interactions in Forming the Hydrophobic Core

An important role for van der Waals interactions in protein folding is suggested by the unusual close-packed arrangement of the side chains in folded proteins. This argument was pointed out in 1971 by Klapper, who compares the protein core to a wax ball rather than an oil drop [73], and in 1977 by Richards, who compares the protein core to an organic crystal [50]. The fraction of void space in folded proteins is quite small and comparable to that in close-packed spheres, whereas there is a sharp contrast between the small void space in folded proteins and the large void space in nonpolar liquids, as pointed out by these authors. In 1994 Chothia and coworkers [74] made a detailed study of the volumes of aliphatic side chains inside proteins and found they are substantially smaller than in aqueous solution. Water molecules are nevertheless rather well packed around hydrocarbons in aqueous solution and hydrocarbons occupy larger volumes in nonpolar solvents than in water, as pointed out by Klapper [73], who found good agreement between experimental data and predictions based on scaled particle theory.

Stites and coworkers have made extensive mutational experiments on the hydrophobic core of staphylococcal nuclease (SNase) [70, 75] including many single, double, triple, and quadruple mutants. The authors measure interaction energies by taking the difference between the ΔG° values of two single mutants and the double mutant, and so forth. Their overall results suggest that the packing arrangement of nonpolar side chains in the core is an important determinant of SNase stability because favorable interaction energies are associated with certain pairs of residues [75]. They tested this interpretation by constructing seven multiple mutants that were predicted to be hyperstable [75]. The X-ray structures of five mutants were determined. The mutants uniformly show high temperatures for thermal unfolding (T_m). The largest increase in T_m is 22.9 °C and the increase in T_m correlates with the number of van der Waals contacts. The overall results indicate that increased protein stability is correlated with an increased number of van der Waals contacts, and certain pairs of buried side chains make characteristic contributions to the number of van der Waals contacts.

A novel approach to analyzing the thermodynamics of burial of nonpolar side chains in protein folding was undertaken by Varadarajan and coworkers [76, 77]. They measure $\Delta\Delta H$ of folding for mutants with varying extents of burial of nonpolar surface area. Note that $\Delta\Delta H$ should be small near 25 °C for burial of nonpolar surface area, according to liquid–liquid transfer experiments with hydrocarbons (see Figures 6.4A and 6.4B). On the other hand, van der Waals packing interactions

are enthalpy-driven and larger values of $\Delta\Delta H$ are expected if the folding thermodynamics are dominated by packing interactions. The authors are able to measure ΔH directly at fixed temperatures by titration calorimetry for their ribonuclease S system (which is composed of two dissociable species, S-peptide + S-protein) by titrating unfolded S-peptide versus folded S-protein. They determine the X-ray structures of the mutant RNase S species and measure ΔH for folding S-peptide at varying temperatures. They find substantial $\Delta\Delta H$ values for the mutants and conclude that van der Waals packing interactions are responsible; see also the study of cavity mutants of T4 lysozyme by Matthews and coworkers [59].

Zhou and Zhou [78, 79] present an interesting analysis of the energetics of burying different side-chain types. They construct a database of mutational experiments on protein folding from the literature and extract the ''transfer free energy'' contributed by each amino acid side chain to the unfolding free energy: the transfer process is from the folded protein to the unfolded protein in aqueous solution. Then the authors simulate the mutational results by a theoretical analysis based on a potential of mean force, and plot the results as $\Delta\Delta G$ versus ΔASA for each side-chain type. The resulting plots are similar for the various nonpolar amino acids, and the slopes are in the range 25–30 cal mol^{-1} $Å^{-2}$. Interestingly, the plots for polar amino acids are similar in character but have substantially smaller slopes. The authors conclude that it is energetically favorable to fold polar as well as nonpolar side chains, but folding polar side chains contributes less to stability. There should be an energetic penalty if the polar moiety of a side chain is buried without making an H-bond, because the polar group interacts favorably with water (see below).

6.3
Peptide Solvation and the Peptide Hydrogen Bond

6.3.1
History

Interest in the role of the peptide hydrogen bond in stabilizing protein folding took off when Pauling and Corey proposed structures for the α-helix and the parallel and antiparallel β-sheets, and when Schellman analyzed [6] the factors that should determine the stability of the helix backbone in water. He estimated the net enthalpy change for forming a peptide H-bond in water as -1.5 kcal mol^{-1}, based on heat of dilution data for urea solutions [80], and he concluded that an isolated α-helix backbone should be marginally stable in water and might or might not be detectable [6]. Klotz and Franzen [81] investigated by infrared spectroscopy whether H-bonded dimers of *N*-methylacetamide are detectable in water and concluded that they have at best marginal stability: see also a later study of the dimerization of δ-valerolactam [82].

Determination of the first X-ray structures of proteins sparked interest in the problem of whether the peptide sequence for a protein helix can form the helix in water. Initial results based on circular dichroism (CD) (or optical rotatory disper-

sion) spectroscopy were disappointing. In 1968 Epand and Scheraga [83] examined two peptides obtained by cyanogen bromide cleavage from the highly helical protein sperm whale myoglobin and found very low values for the possible helix contents. In 1969 Taniuchi and Anfinsen [84] cleaved staphylococcal nuclease, which has 149 residues, between residues 126 and 127 and found that both fragments 1–126 and 127–149 (each of which has a helix of modest size) were structureless by CD as well as by other physical criteria. In 1969–1971 Klee obtained tantalizing results, suggestive of some helix formation at low temperatures, with the ''S-peptide'' (residues 1–20) [85] and ''C-peptide'' (residues 1–13) [86] of ribonuclease A. (The RNase A helix contains residues 3–12.) The helix problem languished for 11 years, perhaps because of the disappointing results found in the laboratories of Anfinsen and Scheraga. In 1982, Bierzynski et al. [87] used NMR as well as CD spectroscopy to reinvestigate the claim of Brown and Klee [86] and confirmed that partial helix formation occurs. Bierzynski et al. found that helix formation is strongly pH dependent and the pK values indicate that both a His residue and a Glu residue are required for helix formation, thus implicating specific side-chain interactions [87]. Several kinds of helix-stabilizing side-chain interactions were later studied, such as salt bridges, amino–aromatic interactions, and charge–helix dipole interactions [88]. The helix formed by C-peptide or S-peptide has basic properties in common with the protein helix of RNase A: both the peptide and protein helices have the same two helix-stabilizing side-chain interactions and the same helix termination signals. Today hundreds of peptides with sequences derived from proteins form partly stable helices in water [89]. The helix termination signals are well-studied [90] and consistent results are found from protein and peptide helices. The AGADIR algorithm of Muñoz and Serrano [89] fits the helix contents of peptides by statistically evaluating the many parameters describing side-chain interactions and helix propensities.

In 1989 Marqusee et al. [91] found that alanine-rich peptides form stable helices in water without the aid of specific side-chain interactions. Of the 20 natural amino acids, only alanine has this ability [92]. Because alanine has only a $-CH_3$ group for a side chain, this result strongly suggests that the helix backbone itself has measurable stability in water and that side chains longer than alanine somehow destabilize the helix. In 1985 Dyson and coworkers used NMR to detect a reverse turn conformation in a nine-residue peptide [93], and later studied the specific sequence requirements for forming this class of turn in peptides [94]. In 1994 Serrano and coworkers [95] found the β-hairpin conformation in a protein-derived peptide, and other β-hairpin sequences were soon found. Thus, H-bonded secondary structures derived from proteins can in general be studied in peptides in aqueous solution.

Determination of the energetic nature of the peptide H-bond was a chief motivation for the pioneering development in 1979 of a molecular force field by Lifson and coworkers [96]. They concluded that the peptide H-bond is represented to a good first approximation by an electrostatic model based on partial charges residing on the peptide CO and NH groups. Their analysis indicates that the peptide

CO and NH groups interact with water as well as with each other. Solvation of the peptide polar groups has been analyzed by using amides such as *N*-methylacetamide (NMA) as models for the peptide group.

6.3.2 Solvation Free Energies of Amides

In a classic experiment, Wolfenden demonstrated in 1978 that amides interact very strongly with water [97]. He measured the equilibrium distribution of a radioactively labeled amide between aqueous solution and the vapor phase, and calculated its transfer free energy from the vapor phase to aqueous solution. The (hypothetical) standard state concentration of the amide solute is 1 M in the vapor phase as well as in the liquid phase (see Section 6.2.4). Wolfenden and coworkers next measured the transfer free energies of model compounds for amino acid side chains [98]. The ionized forms of the basic and acidic residues are excluded from their study because species containing a full formal charge interact too strongly with water to be measured. Their results show that the interaction of water with amides is stronger than with any of the model compounds representing side chains. Saturated hydrocarbons, which serve as models for the side chains of leucine, isoleucine and valine, are preferentially excluded from aqueous solution and have unfavorable solvation free energies. The basic cause is the large unfavorable work of making a cavity in water, which exceeds the favorable van der Waals interaction energy (see Refs [24, 25] and Section 6.2.4). Figure 6.3 shows a direct proportionality between the transfer free energy ΔG° and ASA for transfer of alkanes from aqueous solution to the vapor phase.

For polar solutes as well as nonpolar solutes, there is an unfavorable work of making a cavity in water for the solute and a favorable van der Waals interaction energy between the solute and water. These effects must be removed from the observed solvation free energy to obtain the interaction free energy between water and the polar groups. Currently the sum of the cavity and van der Waals terms is approximated by experimental data for the transfer free energies of nonpolar solutes, as suggested by Makhatadze and Privalov [61] and Sitkoff et al. [99]. The observed solvation free energy in Eq. (9) is split into two terms: one for the polar groups ($\Delta G(\text{pol})$) and one for the sum of the cavity and van der Waals terms ($\Delta G(\text{np})$). The value of $\Delta G(\text{np})$ in Eq. (10) is taken from the plot of transfer free energy versus ASA for alkanes and the proportionality coefficient β is taken from this plot (Figure 6.3).

$$\Delta G^\circ(\text{obs}) = \Delta G(\text{pol}) + \Delta G(\text{np}) \tag{9}$$

$$\Delta G(\text{np}) = \beta \text{ASA} \tag{10}$$

Values of $\Delta G(\text{pol})$ are given for four amides in Table 6.2, taken from Avbelj et al. [100]. Although the correction term $\Delta G(\text{np})$ is obtained by the approximate procedure of using experimental data for nonpolar solutes, the correction term is fairly

Tab. 6.2. Solvation free energy data for amides as models for the free peptide group[a].

Quantity[b] *(kJ mol^{-1})*	*Acetamide*	*N-methyl-acetamide*	*N,N-dimethyl-acetamide*	*Propionamide*
ΔG°(obs)	−40.58	42.13	−35.73	−39.25
ΔG(np)	8.16	−9.04	9.71	8.87
ΔG(pol)	−48.74	−51.17	−45.44	−48.12
ΔH(obs)	−68.28	−71.42	−69.29	−73.01
ΔH(np)	−19.54	−23.35	−26.15	−22.64
ΔH(pol)	−48.74	−48.07	−43.14	−50.38
ASA(pol) (nm^2)	0.95	0.51	0.32	0.88

[a] Data are from Ref. [100].
[b] ΔG°(obs) is the observed transfer free energy from gas to water. Data are taken from Ref. [100], based on a standard state concentration of 1 M in both the gas and liquid phases and without volume correction. ΔG(np) is the correction for the cavity and van der Waals terms according to Eq. (10) and using the value of β given in Ref. [100] (6.4 cal mol^{-1} Å^{-2}), which is not volume-corrected. ΔG(pol) is the solvation free energy of the polar groups in the amide, found from Eq. (9). The quantities ΔH(obs), ΔH(np) and ΔH(pol) have corresponding meanings and are explained in the text. Values for ΔH(obs) are taken from various calorimetric papers, and references are given in Ref. [100]. ASA(pol) is the value of ASA assigned to the polar groups.

small (9.0 kJ mol^{-1} for *N*-methylacetamide) compared with the observed solvation free energy (−42.1 kJ mol^{-1}) or the value of ΔG(pol) (−51.2 kJ mol^{-1}).

The value of β (Eq. (6)) used by Sitkoff et al. [99] (to obtain values for ΔG(np) in Eq. (10)) is based on transfer free energies that are not volume-corrected [21, 99]. The resulting values of ΔG(pol) (obtained from Eq. (9) for a database of model compounds) have been used to calibrate the PARSE parameter set of the DelPhi algorithm [99] (see Section 6.3.5). For this reason, the values of ΔG(np) in Table 6.1 are based on transfer free energies that are not volume-corrected.

The enthalpy of interaction between the polar groups and water, ΔH(pol), has been obtained [100] from calorimetric data in the literature after approximating ΔH(np) by experimental data for alkanes, as before.

$$\Delta H(\text{obs}) = \Delta H(\text{pol}) + \Delta H(\text{np}) \tag{11}$$

The value for ΔH(obs) in Eq. (11) is found by combining results from two calorimetric experiments. The first experiment yields the heat of vaporization of the liquid amide and the second experiment yields the heat of solution of the liquid amide in water. The value of ΔH(np) is found from the heats of solution, plotted against ASA, of gaseous hydrocarbons in water, for which the heats of solution depend chiefly [24] on the favorable van der Waals interaction energy of the hydrocarbon solute with water.

Some interesting conclusions may be drawn from Table 6.2: (1) The four different amides have similar values for ΔG(pol) and also for ΔH(pol). Although there

is a threefold difference between the values of polar ASA for acetamide and *N*,*N*-dimethylacetamide, their values of ΔG(pol) differ by less than 10%. Thus, the computationally convenient approximation of assuming proportionality between ΔG(pol) and polar ASA [101] is not a satisfactory assumption for amides. The polar groups of an amide interact strongly with water even when they are only partly exposed. (2) The value of ΔG(pol) is nearly equal to the value of ΔH(pol) for each amide, and so the $T\Delta S$(pol) term must be small by comparison. This conclusion is surprising if one expects a large entropy change when water interacts with polar groups. It is not surprising, however, after considering the entropies of hydration of crystalline hydrate salts [102] or the prediction [103] made from the Born equation that ΔG(pol) is almost entirely enthalpic. The basic reason why the entropy term looks small is that the enthalpy term is large. (3) The values of ΔG(pol) are impressively large for amides. NMA is sometimes taken as a model for studying the interaction between water and the peptide group, and its value of ΔG(pol) is -51.2 kJ mol^{-1}. This is a huge value for only one peptide group: $-\Delta G^\circ$ for folding a 100-residue protein is typically 40 kJ mol^{-1} or less. Evidently solvation of the peptide group is a major factor in the energetics of folding: see Wolfenden [97] and Makhatadze and Privalov [61].

6.3.3 Test of the Hydrogen-Bond Inventory

Protein chemists often use the H-bond inventory to interpret results of mutational experiments when the experiment involves a change in the number of hydrogen bonds [104, 105]. Experimental data for amides provide a direct test of the H-bond inventory [106]. A key feature of the H-bond inventory is that water (W) is treated as a reactant and the change in the total number of W•••W H-bonds is included in the inventory. When discussing the formation of a peptide H-bond (CO•••HN) in water, the H-bond inventory takes the form of Eq. (12).

$$\text{CO}\bullet\bullet\bullet\text{W} + \text{NH}\bullet\bullet\bullet\text{W} = \text{CO}\bullet\bullet\bullet\text{HN} + \text{W}\bullet\bullet\bullet\text{W} \qquad (12)$$

The peptide NH and CO groups are predicted to make one H-bond each to water in the unfolded peptide, and the CO and NH groups are predicted not to interact further with water after the peptide H-bond is made. Gas phase calculations of H-bond energies indicate that all four H-bond types in Eq. (12) have nearly the same energy [107], -25 ± 4 kJ mol^{-1}, and consequently the net enthalpy change predicted for forming a peptide H-bond in water is 0 ± 4 kJ mol^{-1}. The enthalpy of sublimation of ice has been used to give an experimental estimate of -21 kJ mol^{-1} for the enthalpy of the W•••W H-bond [45].

When a dry molecule of amide (Am) in the gas phase becomes hydrated upon solution in water, the H-bond inventory model postulates that the CO and NH groups each make one H-bond to water and one W•••W H-bond is broken:

$$\text{HN–Am–CO} + \text{W}\bullet\bullet\bullet\text{W} = \text{W}\bullet\bullet\bullet\text{HN–Am–CO}\bullet\bullet\bullet\text{W} \qquad (13)$$

Thus, the H-bond inventory predicts a net increase of one H-bond when a dry amide becomes solvated, with a net enthalpy change of -25 ± 4 kJ mol^{-1}. But the experimental enthalpy change for solvating the amide polar groups is ~ -50 kJ mol^{-1} (Table 6.2) and the H-bond inventory fails badly.

6.3.4 The Born Equation

In 1920 Max Born wanted to know why atom-smashing experiments can be visualized in Wilson's cloud chamber [108]: an ion leaves a track of tiny water droplets as it passes through supersaturated water vapor. He computed the work of charging a spherical ion in vacuum and in water, and he treated water as a continuum solvent with dielectric constant D. He gave the favorable change in free energy for transferring the charged ion from vacuum to water as:

$$-\Delta G = (q^2/2r)[1 - (1/D)] \tag{14}$$

in which q is the charge and r is the radius of the ion [109]. This simple calculation answers Born's question: the free energy change is enormous, of the order of -420 kJ mol^{-1} for a monovalent ion transferred to water. Consequently the ion easily nucleates the formation of water droplets. Inorganic chemists – notably W. M. Latimer – realized that Born's equation provides a useful guide to the behavior of solvation free energies. Note that the solvation free energy predicted by Born's equation is inversely proportional to the ionic radius, but which radius should be used? Ions are strongly solvated and the solvated radius is subject to argument. Rashin and Honig [103] argued that the covalent radius is a logical choice and gives consistent results for different monovalent anions and cations. They also point out that one can obtain the enthalpy of solvation from Born's equation and it predicts that, if the solvent is water with a high dielectric constant (near 80), the solvation free energy is almost entirely enthalpy. Quite recently the solvation free energies of monovalent ions have been analyzed by a force field employing polarizable water molecules [110], with excellent agreement between theory and experiment.

6.3.5 Prediction of Solvation Free Energies of Polar Molecules by an Electrostatic Algorithm

The success of the Born equation in rationalizing solvation free energies of monovalent anions and cations suggests that it should be possible to predict the solvation free energies of polar molecules by using an appropriate electrostatic algorithm plus knowledge of the structure and partial charges of the polar molecule. The problem has a long history. Kirkwood and Westheimer [111] developed a theory in 1938, based on a simple geometrical model, that gives the effect of a low dielectric environment within the molecule on the separation between the two pK values of a dicarboxylic acid. The problem was treated later by an electrostatic algorithm that is free from geometrical assumptions about the shape of the molecule

[112]. The latter treatment includes the effect of electrostatic solvation (the ''Born term'') as well as the electrostatic interaction between the two charged carboxylate groups.

Various electrostatic algorithms are in current use [99, 113] today, including one based on using Langevin dipoles [113] to treat the polarization of water molecules in the vicinity of charged groups. The focus here is on the DelPhi algorithm of Honig and coworkers [99], because the PARSE parameter set of DelPhi has been calibrated against a database of experimental solvation free energies for small molecules that includes amides [99]. Thus, DelPhi may plausibly be used to predict the solvation free energies of the polar CO and NH groups of peptides in various conformations. There are no adjustable parameters and the predicted values of electrostatic solvation free energy (ESF) may be compared directly with experiments if suitable data are available. The DelPhi algorithm uses a low dielectric ($D = 2$) cavity to represent the shape of the solute while the partial charges of the solute are placed on a finely spaced grid running through the cavity, and the solvent is represented by a uniform dielectric constant of 80 [99]. The results are calculated by Poisson's equation if the solvent does not contain mobile ions, or by the Poisson-Boltzmann equation if it does.

6.3.6
Prediction of the Solvation Free Energies of Peptide Groups in Different Backbone Conformations

A basic prediction about the electrostatic interactions among polar CO and NH groups in the peptide backbone is that the interactions depend strongly on the peptide backbone conformation [114–116]. The peptide group is normally fixed in the trans conformation while the CO and NH dipoles of adjacent peptide units are aligned antiparallel in the extended β-strand conformation and parallel in the α-helix conformation. When adjacent peptide dipoles are parallel they make unfavorable interactions unless the dipoles are placed end to end, when formation of peptide H-bonds occurs. These simple observations have important consequences, as pointed out especially by Brant and Flory [114, 115] and by Avbelj and Moult [116]. Thus, intrachain electrostatic free energy favors the extended β conformation over the compact α conformation and this factor tends to make β the default conformation in unfolded peptides [114]. There is a very large difference, $\sim$20 kJ mol^{-1}, between the local electrostatic free energy of the α- and β-strand conformations when calculated with $D = 1$ [116]. Nucleation of an α-helix is difficult for this electrostatic reason [115] as well as for the commonly cited loss in backbone conformational entropy: the peptide dipoles in the helix nucleus are parallel and make unfavorable electrostatic interactions. When helix propagation begins and additional H-bonded residues are added onto the helical nucleus, the favorable H-bond energy drives helix formation.

Calculations of polar group solvation (ESF) in peptides [100, 121] using DelPhi show that ESF depends strongly on two factors: the access of solvent to the peptide group and the local electrostatic potential in the peptide chain. Nearby side chains

Tab. 6.3. Calculated solvation free energies of peptide groups in different backbone conformations[a].

Structure[b]	*ESF (kJ mol^{-1})*	*Reference*
Helix, H-bonded, solvent-exposed	−10.5	100
Helix, not H-bonded, solvent-exposed	−39.8	100
β-hairpin, H-bonded, exposed	−10.5	121
β-hairpin, H-bonded, buried	0	121
β-strand[c], not H-bonded, exposed	−35.6	
Polyproline II[b], not H-bonded, exposed	−38.1	

[a] The solvation free energies of the peptide polar groups (CO, NH) are calculated by DelPhi, as explained in the text (see Section 6.3.5). No adjustable parameters are used except those that describe the structure. The calculations listed here are made for all-alanine peptides and they refer to interior peptide groups, not to N- or C-terminal groups.
[b] The H-bonded, solvent-exposed helix refers to the central residue of a 15-residue helical peptide. The solvent-exposed helix, not H-bonded, refers to the central residue of a five-residue peptide in the α-helical conformation ($\phi, \psi = -65^\circ, -40^\circ$). The β-hairpin (H-bonded, solvent-exposed) refers to typical H-bonded residues in a 15-residue peptide with a β-hairpin conformation taken from a segment of the GB1 structure. The solvent-exposed, extended β-strand, not H-bonded ($\phi, \psi = -120^\circ, 120^\circ$) refers to a nine-residue peptide (F. Avbelj and R. L. Baldwin, to be published). The ESF value given here is more negative than the value previously given [100]. The earlier value by accident had a conformation deviating from ($\phi, \psi = -120^\circ, 120^\circ$) because it was used for an Ala to Val substitution and had a conformation suitable for receiving a Val residue. The polyproline II structure (F. Avbelj and R. L. Baldwin, to be published) refers to the central residue of a nine-residue peptide and has ($\phi, \psi = -70^\circ, 150^\circ$).

hinder access of solvent to the peptide group and reduce the negative ESF value. Consequently, helix and β-structure propensities of the different amino acids should depend on the ESF values of the peptide groups [100, 116–118]. The different helix propensities of the amino acids are often attributed instead to the loss in side-chain entropy when an unfolded peptide forms a helix [119]. Because ESF is almost entirely enthalpic (see Table 6.2), these alternative explanations can be tested by determining if the helix propensity differences are enthalpic or entropic. Temperature-dependence results [120] measured for the nonpolar amino acids (see also Ref. [63]) indicate that the helix propensity differences are largely enthalpic.

Table 6.3 gives some ESF values for alanine peptide groups in different backbone conformations. There are several points of interest: (1) A peptide group in a non-H-bonded alanine peptide has substantially different ESF values in the three major backbone conformations β, α_R, and P_{II} (polyproline II). Consequently any analysis based on group additivity that predicts the overall enthalpy change by assigning constant energetic contributions (independent of backbone conformation) to the polar peptide CO and NH groups cannot be valid. The overall enthalpy change con-

tains contributions from both ESF and the local intrachain electrostatic potential (E_{local}) [116–118], both of which depend on backbone conformation. (2) Because ESF depends on the accessibility of water to a peptide group, the N-terminal and C-terminal peptide groups of an all-alanine peptide have more negative ESF values than interior peptide groups. (3) Replacement of Ala by a larger or more bulky residue such as Val [100] changes the ESF not only at the substitution site but also at neighboring sites (Avbelj and Baldwin, to be published). (4) *N*-Methylacetamide, whose ΔG_{pol} (or ESF) is -51.2 kJ mol^{-1} (Table 6.2) is a poor model for the interaction with water of a free peptide group. There is a 16 kJ mol^{-1} difference between the ESF of NMA and that of an alanine peptide group in the β-conformation (Tables 6.2 and 6.3).

6.3.7
Predicted Desolvation Penalty for Burial of a Peptide H-bond

ESF calculations for completely buried peptide groups, with no accessibility to water, show that their ESF values fall to zero [121]. On the other hand, DelPhi calculations for H-bonded and solvent-exposed peptide groups in either the α-helical [100] or β-hairpin [121] conformation show that H-bonded peptide groups interact with solvent and have highly significant ESF values, about -10.5 kJ mol^{-1} for alanine peptides. Of this, -8.5 kJ mol^{-1} is assigned to the peptide CO group and -2 kJ mol^{-1} to the peptide NH group [100]. Consequently, there should be a large desolvation penalty (equal to the ESF of the solvent-exposed, H-bonded peptide group) if a solvent-exposed peptide H-bond becomes completely buried [100, 121]. This deduction agrees with the prediction by Honig and coworkers [122] of an even larger desolvation penalty (16 kJ mol^{-1}) for transferring an amide H-bond in a NMA dimer from water to liquid alkane. Two points of interest for the mechanism of protein folding arise from these ESF calculations. First, the ESF values of H-bonded, solvent-exposed peptide groups are predicted to be a major factor stabilizing molten globule folding intermediates [100, 121]. Second, complete burial of H-bonded secondary structure should involve a substantial desolvation penalty [121]. The size of the penalty depends on how solvated each peptide group is when the secondary structure is solvent-exposed. Neighboring side chains larger or more bulky than Ala reduce the negative ESF of the peptide H-bond.

When Ben-Tal et al. [122] studied the transfer of a H-bonded NMA dimer from water to liquid alkane, they found the formation of the H-bonded dimer in water to be stable by -5 kJ mol^{-1}. Their results give the penalty for forming the H-bond in water and then transferring it to liquid alkane as $-5 + 16 = 11$ kJ mol^{-1}. When their predictions are compared with values for an alanine peptide helix, the results are numerically different (not surprising, in view of the structural difference) but qualitatively the same. Ben-Tal et al. omit the loss in translational and rotational entropy for forming the NMA dimer and their ΔG value of -5 kJ mol^{-1} [122] may be compared with the measured ΔH (-4.0 kJ mol^{-1} [123, 124]) for forming the alanine peptide helix.

Myers and Pace [125] discuss whether or not peptide H-bonds stabilize protein

folding by considering mutational data for H-bonds made between specific pairs of side chains. The ESF studies considered here emphasize the role of context: solvent-exposed peptide H-bonds should stabilize folding but buried H-bonds should detract from stability.

The ESF perspective provides a simple explanation for why alanine, alone of the 20 amino acids in proteins, forms a stable helix in water. The H-bonded, solvent-exposed peptide group in an alanine helix has the most stabilizing ESF value of any nonpolar amino acid except glycine, and glycine fails to form a helix because of the exceptional flexibility of the glycine peptide linkage.

The interaction energy between water and the peptide group is a critical quantity in predicting the structures of membrane proteins from amino acid sequences [126]. Accurate values are needed for the energy of desolvation both of the free and H-bonded forms of the peptide group. The solvent *n*-octanol, which has been used extensively to model the interiors of water-soluble proteins (see Sections 6.2.2 and 6.2.3), has also been used as an experimental model for the lipid bilayer environment of membrane proteins, when measuring transfer free energies of peptides between water and a membrane-like solvent. The small value of the transfer free energy found for the free glycine peptide group transferred from water to water-saturated *n*-octanol (4.8 kJ mol^{-1} [15]) emphasizes the role of the water contained in this solvent, which increases its attraction for the peptide group. Cyclohexane has been used as an alternative solvent that provides a nearly water-free environment when partitioning model compounds [16]. The transfer free energies of amides are much larger in the cyclohexane/water pair than the 4.8 kJ mol^{-1} determined for the peptide group in octanol/water: 25 and 21 kJ mol^{-1} for acetamide and propionamide, respectively [16].

6.3.8
Gas–Liquid Transfer Model

The gas–liquid transfer model has been used in Section 6.2.8 to discuss the respective roles of van der Waals interactions and burial of nonpolar surface area in protein folding (Scheme 6.1). The gas–liquid transfer model can be adapted to discuss the roles of peptide H-bonds and peptide solvation in folding. Scheme 6.2 is written for a simple folding reaction involving only a single peptide H-bond.

Here the conformation of the unfolded form U must be considered because changes in both local electrostatic free energy (E_{local}) and ESF contribute to ΔH

$$\begin{array}{ccccc} U(\gamma)(g) & \Leftarrow (2) \Rightarrow & U(\delta)(g) & \Leftarrow (3) \Rightarrow & N(g) \\ \Uparrow\Downarrow (1) & & \Uparrow\Downarrow (7) & & \Uparrow\Downarrow (4) \\ U(\gamma)(aq) & \Leftarrow (5) \Rightarrow & U(\delta)(aq) & \Leftarrow (6) \Rightarrow & N(aq) \end{array}$$

Scheme 6.2

and they depend on backbone conformation. In order to form the peptide H-bond, a change in peptide backbone conformation must usually take place that involves substantial changes in E_{local} and ESF [116]. The unfolded form U(γ) normally exists as a complex equilibrium mixture of conformations and U(δ) represents the new conformation needed to make the peptide H-bond.

In principle, Scheme 6.2 may be adapted to predict the enthalpy change for forming an alanine peptide helix, whose experimental value is known [123, 124], by assigning provisional values to the different steps. Step 1, desolvating the free peptide group, has $-\text{ESF} = 36$ kJ mol^{-1} if the free peptide conformation is an extended β-strand (Table 6.2), and this $-$ESF value gives the contribution to ΔH because ESF and ΔH(pol) are the same within error for amides (Table 6.2). Likewise, solvating the alanine helix in step 4 can be assigned a contribution to ΔH equal to the ESF of the solvent-exposed, H-bonded peptide group, -10.5 kJ mol^{-1} [100]. In step 3, making the peptide H-bond in the gas phase, the H-bond energy has been calculated by quantum mechanics to be -28 kJ mol^{-1} for the H-bonded NMA dimer [122]. Step 3 also contains, however, an additional unknown quantity, the van der Waals interaction energy made in forming the helix backbone. In step 1 there is also an additional problem. The ESF depends not only on the appropriate mixture of backbone conformations (values for α_R, β, and P_{II} are given in Table 6.3), but also on the tendency of a peptide to bend back on itself, which reduces the exposure to solvent of the peptide groups [100]. Likewise, there are unknown – and possibly quite large – changes in E_{local} in both steps 2 and 3. Thus, the puzzle of predicting ΔH even for a simple folding reaction, like that of forming an alanine peptide helix is far from being completely understood.

Nevertheless, the gas–liquid transfer model provides the background for a simple interpretation of the enthalpy of forming the alanine helix. The observed enthalpy of helix formation, ΔH(H-C), can be written as a sum of three terms, and the enthalpy of interaction with water of the helix or of the coil can be approximated by the ESF value, as explained above.

$$\Delta H(\text{H-C}) = \Delta H(\text{hb}) + \Delta H(\text{H-W}) - \Delta H(\text{C-W}) \tag{15}$$

In Eq. (15), H refers to helix, C to coil (the unfolded peptide), W to water and hb to the peptide H-bond. ΔH(H-C) is the observed enthalpy of helix formation in water (-4.0 kJ mol^{-1}, see above). It equals ΔH(hb), the enthalpy of forming the peptide H-bond (-27.6 kJ mol^{-1} [122]), plus ΔH(H-W), the enthalpy of interaction between water and the helix (-10.5 kJ mol^{-1} (see Table 6.3), minus ΔH(C-W), the enthalpy of interaction between water and the coil (unknown, for reasons explained above). The unknown term ΔH(C-W) is found by solving the equation to be $-27.6 - 10.5 + 4.0 = -34.1$ kJ mol^{-1}. This is a reasonable value for the enthalpy of interaction between water and the coil, in view of the ESF values given in Table 6.3 for different unfolded conformations and given also that bending back of the unfolded peptide on itself is likely to reduce the negative ESF value.

Acknowledgments

I thank Franc Avbelj, David Baker, David Chandler, Ken Dill, Pehr Harbury, B.-K. Lee and John Schellman for discussion, and David Chandler, Alan Grossfield, George Makhatadze and Yaoqi Zhou for sending me their papers before publication.

Footnote added in proof

An important paper [127] was overlooked in writing section 6.2.11 on the role of van der Waals interactions in forming the hydrophobic cores of proteins. Loladze, Ermolenko and Makhatadze [127] made a set of large-to-small mutations (e.g., Val → Ala) in ubiquitin and used scanning calorimetry to measure ΔH, $T\Delta S^\circ$ and ΔG° for mutant unfolding. In agreement with a related study by Varadarajan, Richards and coworkers [76, 77], they find that large-to small mutations are characterized chiefly by unfavorable enthalpy changes, and not by the expected unfavorable entropy changes. Surprisingly, they find favorable entropy changes for these mutants, indicating that some kind of loosening of the structure occurs, perhaps at residues lining the cavity surface. Thus, they conclude that burial of nonpolar surface area stabilizes folding primarily by a favorable enthalpy change and not by the favorable change in $T\Delta S^\circ$ predicted by the liquid–liquid transfer model. The measurements are reported at 50 °C [127], where the favorable free energy change for burial of nonpolar surface area given by the liquid–liquid transfer model (Section [6.2.9]) is approximately 30% enthalpy and 70% $T\Delta S^\circ$. The changes in conformational entropy that occur in large-to-small mutants [127] invalidate testing the relation between buried ASA and ΔG° for folding unless the enthalpy changes are also measured and analyzed. There is the further problem of a dependence of $\Delta\Delta G^\circ$ on the size of the cavity, discussed by Matthews and coworkers [59]. Makhatadze and coworkers find values of ΔC_p in these experiments that are too small to be determined accurately [128], in agreement with predictions by the liquid–liquid transfer and other models.

References

1 Kauzmann, W. (1959). Factors in interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1–63.
2 Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155.
3 Privalov, P. L. & Gill, S. J. (1988). Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* **39**, 191–234.
4 Privalov, P. L. & Gill, S. J. (1989). The hydrophobic effect: a reappraisal. *Pure Appl. Chem.* **61**, 1097–1104.
5 Dill, K. A. (1990). The meaning of hydrophobicity. *Science* **250**, 297–298.
6 Schellman, J. A. (1955). The stability of hydrogen-bonded peptide structures in aqueous solution. *C.R. Trav. Lab. Carlsberg Sér Chim.* **29**, 230–259.
7 Stickle, D. F., Presta, L. G., Dill,

K. A., & Rose, G. D. (1992). Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**, 1143–1159.

8 Kendrew, J. C., Dickerson, R. E., Strandberg, B. E. et al. (1960). Structure of myoglobin. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422–427.

9 Pauling, L., Corey, R. B. & Branson, H. R. (1951). The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl Acad. Sci. USA* **37**, 205–211.

10 Tanford, C. (1980). *The Hydrophobic Effect*, 2nd edn. John Wiley & Sons, New York.

11 Tanford, C. (1997). How protein chemists learned about the hydrophobic factor. *Protein Sci.* **6**, 1358–1366.

12 Southall, N. T., Dill, K. A., & Haymet, A. D. J. (2002). A view of the hydrophobic effect. *J. Phys Chem. B* **106**, 521–533.

13 Nozaki, Y. & Tanford, C. (1971). The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. *J. Biol. Chem.* **246**, 2211–2217.

14 Fauchère, J.-L. & Pliska, V. (1983). Hydrophobic parameters π of amino-acid side chains from partitioning of N-acetyl-amino-acid amides. *Eur. J. Med. Chem.* **18**, 369–375.

15 Wimley, W. C., Creamer, T. P., & White, S. H. (1996). Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* **35**, 5109–5124.

16 Radzicka, A. & Wolfenden, R. (1988). Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol and neutral aqueous solution. *Biochemistry* **27**, 1644–1670.

17 Hermann, R. B. (1972). Theory of hydrophobic bonding. II. The correlation of hydrocarbon solubility in water with solvent cavity surface area. *J. Phys. Chem.* **76**, 2754–2759.

18 Chothia, C. (1974). Hydrophobic bonding and accessible surface area in proteins. *Nature* **248**, 338–339.

19 Reynolds, J. A., Gilbert, D. B., & Tanford, C. (1974). Empirical correlation between hydrophobic free energy and aqueous cavity surface area. *Proc. Natl Acad. Sci. USA* **71**, 2925–2927.

20 Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.

21 Sharp, K. A., Nicholls, A., Friedman, R., & Honig, B. (1991). Extracting hydrophobic free energies from experimental data: relationship to protein folding and theoretical models. *Biochemistry* **30**, 9686–9697.

22 Hermann, R. B. (1977). Use of solvent cavity area and number of packed molecules around a solute in regard to hydrocarbon solubilities and hydrophobic interactions. *Proc. Natl Acad. Sci. USA* **74**, 4144–4145.

23 Ben-Naim, A. & Marcus, Y. (1984). Solvation thermodynamics of nonionic solutes. *J. Chem. Phys.* **81**, 2016–2027.

24 Lee, B. (1991). Solvent reorganization contribution to the transfer thermodynamics of small nonpolar molecules. *Biopolymers* **31**, 993–1008.

25 Lee, B. (1995). Analyzing solvent reorganization and hydrophobicity. *Methods Enzymol.* **259**, 555–576.

26 Widom, B. (1982). Potential-distribution theory and the statistical mechanics of fluids. *J. Phys. Chem.* **86**, 869–872.

27 Jorgensen, W. L., Gao, J., & Ravimohan, C. (1985). Monte Carlo simulations of alkanes in water. Hydration numbers and the hydrophobic effect. *J. Phys. Chem.* **89**, 3470–3473.

28 Pratt, L. R. & Chandler, D. (1977). Theory of the hydrophobic effect. *J. Chem. Phys.* **67**, 3683–3704.

29 Gallichio, E., Kubo, M. M., & Levy, R. M. (2000). Enthalpy-entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation. *J. Phys. Chem. B* **104**, 6271–6285.

30 Lee, B. (1985). The physical origin of the low solubility of nonpolar solutes in water. *Biopolymers* **24**, 813–823.

31 Rank, J. A. & Baker, D. (1998). Contributions of solvent-solvent hydrogen bonding and van der Waals interaction to the attraction between methane molecules in water. *Biophys. Chem.* **71**, 199–204.

32 Tanford, C. (1979). Interfacial free energy and the hydrophobic effect. *Proc. Natl Acad. Sci. USA* **76**, 4175–4176.

33 Postma, J. P. M., Berendsen, H. J. C., & Haak, J. R. (1982). Thermodynamics of cavity formation in water: a molecular dynamics study. *Faraday Symp. Chem. Soc.* **17**, 55–67.

34 Hummer, G., Garde, S., García, A., Pohorille, A., & Pratt, L. R. (1996). An information theory model of hydrophobic interactions. *Proc. Natl Acad. Sci. USA* **93**, 8951–8955.

35 DeYoung, L. R. & Dill, K. A. (1990). Partitioning of nonpolar solutes into bilayers and amorphous n-alkanes. *J. Phys. Chem.* **94**, 801–809.

36 Chan, H. S. & Dill, K. A. (1997). Solvation: how to obtain microscopic energies from partitioning and solvation experiments. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 425–459.

37 Chandler, D. (2004). Hydrophobicity: two faces of water. *Nature (London)* **417**, 491–493.

38 Lum, K., Chandler, D., & Weeks, J. D. (1999). Hydrophobicity at small and large length scales. *J. Phys. Chem. B* **103**, 4570–4577.

39 Lee, B. (1994). Relation between volume correction and the standard state. *Biophys. Chem.* **51**, 263–269.

40 Huang, D. M. & Chandler, D. (2000). Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *Proc. Natl Acad. Sci. USA* **97**, 8324–8327.

41 Southall, N. T. & Dill, K. A. (2000). The mechanism of hydrophobic solvation depends on solute radius. *J. Phys. Chem. B* **104**, 1326–1331.

42 Pohorille, A. & Pratt, L. R. (1990). Cavities in molecular liquids and the theory of hydrophobic solubility. *J. Am. Chem. Soc.* **112**, 5066–5074.

43 Baldwin, R. L. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl Acad. Sci. USA* **83**, 8069–8072.

44 Gill, S. J., Dec, S. F., Olofsson, G., & Wadsö, I. (1985). Anomalous heat capacity of hydrophobic solvation. *J. Phys. Chem.* **89**, 3758–3761.

45 Pauling, L. (1960). *The Nature of the Chemical Bond*, 3rd edn, pp. 468–472. Cornell University Press, Ithaca, NY.

46 Lazaridis, T. (2001). Solvent size versus cohesive energy as the origin of hydrophobicity. *Acc. Chem. Res.* **34**, 931–937.

47 Stellner, K. L., Tucker, E. E., & Christian, S. D. (1983). Thermodynamic properties of the benzene-phenol dimer in dilute aqueous solution. *J. Sol. Chem.* **12**, 307–313.

48 Raschke, T. M., Tsai, J., & Levitt, M. (2001). Quantification of the hydrophobic interaction by simulations of the aggregation of small hydrophobic solutes in water. *Proc. Natl Acad. Sci. USA* **98**, 5965–5969.

49 Rank, J. A. & Baker, D. (1997). A desolvation barrier to cluster formation may contribute to the rate-limiting step in protein folding. *Protein Sci.* **6**, 347–354.

50 Richards, F. M. (1977). Areas, volumes. packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.

51 Gill, S. J., Nichols, N. F., & Wadsö, I. (1976). Calorimetric determination of enthalpies of solution of slightly soluble liquids. II. Enthalpy of solution of some hydrocarbons in water and their use in establishing the temperature dependence of their solubilities. *J. Chem. Thermodynam.* **8**, 445–452.

52 Edsall, J. T. (1935). Apparent molal heat capacities of amino acids and other organic compounds. *J. Am. Chem. Soc.* **57**, 1506–1507.

53 Garde, S., Hummer, G., García, A., Paulaitis, M. E., & Pratt, L. R. (1996). Origin of entropy convergence

in hydrophobic hydration and protein folding. *Phys. Rev. Lett.* **77**, 4966–4968.

54 Graziano, G. & Lee, B.-K. (2003). Entropy convergence in hydrophobic hydration: a scaled particle theory analysis. *Biophys. Chem.* **105**, 241–250.

55 Privalov, P. L. (1979). Stability of proteins: small globular proteins. *Adv. Protein Chem.* **33**, 167–241.

56 Robertson, A. D. & Murphy, K. P. (1997). Protein structure and the energetics of protein stability. *Chem. Rev.* **97**, 1251–1267.

57 Schellman, J. A. (1997). Temperature, stability and the hydrophobic interaction. *Biophys. J.* **73**, 2960–2964.

58 Pace, C. N. (1992). Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.* **226**, 29–35.

59 Eriksson, A. E., Baase, W. A., Zhang, X.-J. et al. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **255**, 178–183.

60 Livingstone, J. R., Spolar, R. S., & Record, M. T. (1991). Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry* **30**, 4237–4244.

61 Makhatadze, G. I. & Privalov, P. L. (1993). Contribution of hydration to protein folding. I. The enthalpy of hydration. *J. Mol. Biol.* **232**, 639–659.

62 Gómez, J., Hilser, V. J., Xie, Dong, & Freire, E. (1995). The heat capacity of proteins. *Proteins Struct. Funct. Genet.* **22**, 404–412.

63 Richardson, J. M. & Makhatadze, G. I. (2003). Temperature dependence of the thermodynamics of the helix-coil transition. *J. Mol. Biol.*, **335**, 1029–1037.

64 Kauzmann, W. (1987). Thermodynamics of unfolding. *Nature (London)* **325**, 763–764.

65 Hummer, G., Garde, S., García, A., Paulaitis, M. E., & Pratt, L. R. (1998). The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins. *Proc. Natl Acad. Sci. USA* **95**, 1552–1555.

66 Akasaka, K. (2003). Highly fluctuating protein structures revealed by variable-pressure nuclear magnetic resonance. *Biochemistry* **42**, 10875–10885.

67 Ooi, T. & Oobatake, M. (1988). Effects of hydrated water on protein unfolding. *J. Biochem. (Tokyo)* **103**, 114–120.

68 Oobatake, M. & Ooi, T. (1992). Hydration and heat stability effects on protein unfolding. *Prog. Biophys. Mol. Biol.* **59**, 237–284.

69 Simonson, T. & Brünger, A. T. (1994). Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment. *J. Phys. Chem.* **98**, 4683–4694.

70 Chen, J., Lu, Z., Sakon, J., & Stites, W. E. (2000). Increasing the thermostability of staphylococcal nuclease: implications for the origin of protein thermostability. *J. Mol. Biol.* **303**, 125–130.

71 Havranek, J. J. & Harbury, P. B. (2003). Automated design of specificity in molecular recognition. *Nature Struct. Biol.* **10**, 45–52.

72 Nicholls, A., Sharp, K. A., & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296.

73 Klapper, M. H. (1971). On the nature of the protein interior. *Biochim. Biophys. Acta* **229**, 557–566.

74 Harpaz, Y., Gerstein, M., & Chothia, C. (1994). Volume changes on protein folding. *Structure* **2**, 641–659.

75 Chen, J. & Stites, W. E. (2001). Packing is a key selection factor in the evolution of protein hydrophobic cores. *Biochemistry* **40**, 15280–15289.

76 Varadarajan, R., Connelly, P. R., Sturtevant, J. M., & Richards, F. M. (1992). Heat capacity changes for protein-peptide interactions in the ribonuclease S system. *Biochemistry* **31**, 1421–1426.

77 Ratnaparkhi, G. S. & Varadarajan, R. (2000). Thermodynamic and structural studies of cavity formation in proteins suggest that loss of

packing interactions rather than the hydrophobic effect dominates the observed energetics. *Biochemistry* **39**, 12365–12374.

78 Zhou, H. & Zhou, Y. (2002). Stability scale and atomic solvation parameters extracted from 1023 mutation experiments. *Proteins Struct. Funct. Genet.* **49**, 483–492.

79 Zhou, H. & Zhou, Y. (2003). Quantifying the effect of burial of amino acid residues on protein stability. *Proteins Struct. Funct. Genet.* **54**, 315–322.

80 Schellman, J. A. (1955). The thermodynamics of urea solutions and the heat of formation of the peptide hydrogen bond. *C.R. Trav. Lab. Carlsberg, Sér. Chim.* **29**, 223–229.

81 Klotz, I. M. & Franzen, J. S. (1962). Hydrogen bonds between model peptide groups in solution. *J. Am. Chem. Soc.* **84**, 3461–3466.

82 Susi, H., Timasheff, S. N., & Ard, J. S. (1964). Near infrared investigation of interamide hydrogen bonding in aqueous solution. *J. Biol. Chem.* **239**, 3051–3054.

83 Epand, R. M. & Scheraga, H. A. (1968). The influence of long-range interactions on the structure of myoglobin. *Biochemistry* **7**, 2864–2872.

84 Taniuchi, H. & Anfinsen, C. B. (1969). An experimental approach to the study of the folding of staphylococcal nuclease. *J. Biol. Chem.* **244**, 3864–3875.

85 Klee, W. A. (1968). Studies on the conformation of ribonuclease S-peptide. *Biochemistry* **7**, 2731–2736.

86 Brown, J. E. & Klee, W. A. (1971). Helix-coil transition of the isolated amino-terminus of ribonuclease. *Biochemistry* **10**, 470–476.

87 Bierzynski, A., Kim, P. S., & Baldwin, R. L. (1982). A salt bridge stabilizes the helix formed by the isolated C-peptide of RNase A. *Proc. Natl Acad. Sci. USA* **79**, 2470–2474.

88 Baldwin, R. L. (1995). α-Helix formation by peptides of defined sequence. *Biophys. Chem.* **55**, 127–135.

89 Muñoz, V. & Serrano, L. (1994). Elucidating the folding problem of helical peptides using emprical parameters. *Nature Struct. Biol.* **1**, 399–409.

90 Aurora, R. & Rose, G. D. (1998). Helix capping. *Protein Sci.* **7**, 21–38.

91 Marqusee, M., Robbins, V. H., & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine-based peptides. *Proc. Natl Acad. Sci. USA* **86**, 5286–5290.

92 Rohl, C. A., Chakrabartty, A., & Baldwin, R. L. (1996). Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.* **5**, 2623–2637.

93 Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985). The immunodominant site of a synthetic immunogen has a conformational preference in water for a type-II reverse turn. *Nature (London)* **318**, 480–483.

94 Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988). Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn. *J. Mol. Biol.* **201**, 161–200.

95 Bianco, F. J., Rivas, G., & Serrano, L. (1995). A short linear peptide that folds into a stable native β-hairpin in aqueous solution. *Nature Struct. Biol.* **1**, 584–590.

96 Lifson, S., Hagler, A. T., & Dauber, P. (1979). Consistent force field studies of hydrogen-bonded crystals. 1. Carboxylic acids, amides, and the C=O•••H– hydrogen bonds. *J. Am. Chem. Soc.* **101**, 5111–5121.

97 Wolfenden, R. (1978). Interaction of the peptide bond with solvent water: a vapor phase analysis. *Biochemistry* **17**, 201–204.

98 Wolfenden, R., Andersson, L., Cullis, P. M., & Southgate, C. C. B. (1981). Affinities of amino acid side chains for solvent water. *Biochemistry* **20**, 849–855.

99 Sitkoff, D., Sharp, K. A., & Honig, B. (1994). Accurate calculation of

hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **98**, 1978–1988.

100 Avbelj, F., Luo, P., & Baldwin, R. L. (2000). Energetics of the interaction between water and the helical peptide group and its role in determining helix propensities. *Proc. Natl Acad. Sci. USA* **97**, 10786–10791.

101 Eisenberg, D. & McLachlan, A. D. (1986). Solvation energy in protein folding and binding. *Nature (London)* **319**, 199–203.

102 Dunitz, J. D. (1994). The entropic cost of bound water in crystals and biomolecules. *Science* **264**, 670.

103 Rashin, A. A. & Honig, B. (1985). Reevaluation of the Born model of ion hydration. *J. Phys. Chem.* **89**, 5588–5593.

104 Fersht, A. R., Shi, J.-P., Knill-Jones, J. et al. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. *Nature (London)* **314**, 235–238.

105 Fersht, A. R. (1987). The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* **12**, 301–304.

106 Baldwin, R. L. (2003). In search of the energetic role of peptide H-bonds. *J. Biol. Chem.* **278**, 17581–17588.

107 Mitchell, J. B. O. & Price, S. L. (1991). On the relative strengths of amide•••amide and amide•••water hydrogen bonds. *Chem. Phys. Lett.* **180**, 517–523.

108 Galison, P. (1997). *Image and Logic*, p. 68, University of Chicago Press, Chicago.

109 Born, M. (1920). Volumen und Hydrationswärme der Ionen. *Z. Physik* **1**, 45–48.

110 Grossfield, A., Ren, P., & Ponder, J. W. (2003). Ion solvation thermodynamics from simulation with a polarizable force field. *J. Am. Chem. Soc.* **125**, 15671–15682.

111 Kirkwood, J. G. & Westheimer, F. H. (1938). The electrostatic influence of substituents on the dissociation constants of organic acids. *J. Chem. Phys.* **6**, 506–512.

112 Rajasekaran, E., Jayaram, B., & Honig, B. (1994). Electrostatic interactions in aliphatic dicarboxylic acids: a computational route to the determination of pK shifts. *J. Am. Chem. Soc.* **116**, 8238–8240.

113 Florian, J. & Warshel, A. (1997). Langevin dipoles model for ab initio calculations of chemical processes in solution: parameterization and application to hydration free energies of neutral and ionic solutes and conformational analysis in aqueous solution. *J. Phys. Chem. B* **101**, 5583–5595.

114 Brant, D. A. & Flory, P. J. (1965). The configuration of random polypeptide chains. II. Theory. *J. Am. Chem. Soc.* **87**, 2791–2800.

115 Brant, D. A. & Flory, P. J. (1965). The role of dipole interactions in determining polypeptide configurations. *J. Am. Chem. Soc.* **87**, 663–664.

116 Avbelj, F. & Moult, J. (1995). Role of electrostatic screening in determining protein main chain conformational preferences. *Biochemistry* **34**, 755–764.

117 Avbelj, F. & Fele, L. (1998). Role of main-chain electrostatics, hydrophobic effect and side-chain conformational entropy in determining the secondary structure of proteins. *J. Mol. Biol.* **279**, 665–684.

118 Avbelj, F. (2000). Amino acid conformational preferences and solvation of polar backbone atoms in peptides and proteins. *J. Mol. Biol.* **300**, 1335–1359.

119 Creamer, T. P. & Rose, G. D. (1994). α-Helix-forming propensities in peptides and proteins. *Proteins Struct. Funct. Genet.* **19**, 85–97.

120 Luo, P. & Baldwin, R. L. (1999). Interaction between water and polar groups of the helix backbone: an important determinant of helix propensities. *Proc. Natl Acad. Sci. USA* **96**, 4930–4935.

121 Avbelj, F. & Baldwin, R. L. (2002). Role of backbone solvation in determining thermodynamic β propensities of the amino acids. *Proc. Natl Acad. Sci. USA* **99**, 1309–1313.

122 Ben-Tal, N., Sitkoff, D., Topol, I. A., Yang, A.-S., Burt, S. K., & Honig, B.

(1997). Free energy of amide hydrogen bond formation in vacuum, in water, and in liquid alkane solution. *J. Phys. Chem. B* **101**, 450–457.

123 Lopez, M. M., Chin, D.-H., Baldwin, R. L., & Makhatadze, G. I. (2002). The enthalpy of the alanine peptide helix measured by isothermal titration calorimetry using metal-binding to induce helix formation. *Proc. Natl Acad. Sci. USA* **99**, 1298–1302.

124 Goch, G., Maciejczyk, Oleszczuk, M., Stachowiak, D., Malicka, J., & Bierzynski, A. (2003). Experimental investigation of initial steps of helix propagation in model peptides. *Biochemistry* **42**, 6840–6847.

125 Myers, J. K. & Pace, C. N. (1996). Hydrogen bonding stabilizes globular proteins. *Biophys. J.* **71**, 2033–2039.

126 Jayasinghe, S., Hristova, K., & White, S. H. (2001). Energetics, stability and prediction of transmembrane helices. *J. Mol. Biol.* **312**, 927–934.

127 Loladze, V. V., Ermolenko, D. N., & Makhatadze, G. I. (2002). Thermodynamic consequences of burial of polar and nonpolar amino acid residues in the protein interior. *J. Mol. Biol.* **320**, 343–357.

128 Loladze, V. V., Ermolenko, D. N., & Makhatadze, G. I. (2001). Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci.* **10**, 1343–1352.