Ångströms and calories

Joël Janin

Address: Laboratoire d'Enzymologie et de Biologie Structurales, UPR 9063 CNRS, Bât. 34, 91198-Gif-sur-Yvette, France.

E-mail: janin@lebs.cnrs-gif.fr

Electronic identifier: 0969-2126-005-00473

Structure 15 April 1997, 5:473–479

Summary

Calorimetry and solution studies yield the enthalpy (ΔH) and free enthalpy (Gibbs energy; ΔG) of protein folding and ligand binding. How do we relate these quantities to the three-dimensional structure? The surface area model offers one possible empirical approach to this question: ΔH and ΔG are estimated with the help of thermodynamic cycles in which folding or binding is assumed to take place in the gas phase or in an organic solvent and is followed by transfer to a water solution. Proportionality coefficients derived from small molecule data relate the transfer ΔH and ΔG to the area of the protein surface made inaccessible to the solvent as a result of folding or binding (ΔASA). As gas phase enthalpies can be evaluated with the help of molecular mechanics, the method can in principle yield quantitative estimates of ΔH. However, this will require the proportionality coefficients to be much more accurate and reliable than they are at present, especially for the hydration of polar groups which is a major contributor to both folding and binding.

Introduction

Protein crystallography and NMR studies have as standard output a molecular structure expressed as a set of atomic coordinates. The output of physical chemical studies of folding or molecular interaction is more often an energy, enthalpy or free enthalpy. Thus, structural biologists tend to think in Ångström units (or possibly nanometres if they care to follow international conventions), whereas other scientists who study the same objects are more familiar with calories or Joules. Folding studies yield the free enthalpy change ΔG between the native and unfolded states of a protein. The enthalpy change ΔH, and the heat capacity change ΔCp, which is the temperature derivative of ΔH, can be measured directly in a calorimeter by differential scanning calorimetry [1]. In recent years, microcalorimeters have been developed that measure heats of reaction at constant temperature. Isothermal titration calorimetry experiments use these devices to titrate binding sites, which gives access to both the enthalpy and the free enthalpy changes upon association [2]. Thus, ligand binding, protein–protein and protein–nucleic acid recognition can now be studied by calorimetry just like folding [3]. Classical relationships between thermodynamic state functions and the equilibrium constant K are recalled in Table 1. We shall use below state functions which pertain to unfolding or ligand dissociation, as these reactions often have positive standard state enthalpies, free enthalpies and heat capacities. State functions for the reverse reactions, folding or association, differ only by the sign.

Structural biologists locate atoms to within a fraction of an Ångström; calorimetrists measure enthalpy changes to within a fraction of a kilocalorie per mole. When the two disciplines work on the same macromolecule, the same protein–ligand or protein–DNA complex, one would like to see their results tied together, but how? Can we translate Ångströms into calories when we discuss folding or recognition? The problem of relating structure to thermodynamics is central to the physical chemistry of biological molecules and has been offered various solutions over the years. It is a very active field and an abundantly reviewed one [4–9]. Here, we concentrate on models that link structural and calorimetric data, and point out their successes and limitations.

Exact and empirical models of solvent–solute interactions

The force fields of molecular mechanics are a straightforward way to convert atomic coordinates into a conformational energy [10–11]. These force fields are familiar to crystallographers and NMR spectroscopists, who use them in structure refinement to complement their own experimental data. Accurate atomic models from high resolution X-ray studies are near a minimum of the conformational energy E, a comforting finding except that it is always a local minimum and one of many. In the absence of experimental constraints, extensive energy minimization on a small protein, such as crambin, yields atomic positions that differ by ~1 Å root mean square from the X-ray structure, and also by 1 Å from one force field to another [12]. However, E is not relevant to thermodynamic experiments: it is the internal energy of an isolated molecule

Table 1

Thermodynamic state functions and the equilibrium constant.

Free enthalpy, enthalpy, entropy	$\Delta G = \Delta H - T\Delta S$
Standard state free enthalpy	$\Delta G^\circ = -RT \ln K$
Temperature derivatives	
Entropy	$\Delta S = -d(\Delta G)/dT$
Heat capacity	$\Delta Cp = d(\Delta H)/dT = T[d(\Delta S)/dT]$

Gas constant $R \sim 2$ cal mol^{-1} K^{-1}. Standard state: pressure $p^\circ = 1$ bar, concentration $c^\circ = 1$ mol l^{-1}. Equilibrium constants: folding/unfolding $K = U/N$, $\Delta G^\circ = \Delta G_{unf}$. Binding/dissociation $K = K_{diss}/c^\circ$, $\Delta G^\circ = \Delta G_{diss}$.

in vacuo at zero K, no solvent is present, nothing moves and the entropy is zero. To be useful, the model must include the solvent and give access to the free energy as well as the internal energy, both of which are ensemble averages, not attached to a particular structure. We may neglect the difference between enthalpy and internal energy, which is related to volume changes and matters only at very high pressure; we certainly may not neglect the entropy.

The contribution of the solvent is just as important. Solvent–solute and solvent–solvent interactions make a large contribution to the energy and entropy of the system. The relaxation of molecular dipoles in water profoundly affects all electrostatic interactions in a way that cannot be modelled by just a macroscopic dielectric constant [13]. In folding or binding, interactions with water are made or unmade, order increases or decreases in the liquid phase. Crystallography and NMR locate those water molecules that are immobilized, but they say little about the many others that undergo energy and entropy changes in the vicinity of the macromolecule. Solvent can be included in molecular dynamic simulations, which gives kinetic energy to atoms and follow their trajectory over time. The force fields are the same as for molecular mechanics, but now time averaging yields the internal energy and fluctuations yield the entropy, provided all populated states of the system have been sampled. With macromolecules, trajectories can be followed only for timescales in the order of a nanosecond and cover a very limited region of phase space [10]. Local changes in a protein may be accessible to the simulation, but certainly not folding or major conformation changes, at least not with the intention of matching biochemical and thermodynamic data.

As an alternative, an entirely empirical approach may be taken. This approach assumes that solvent-related thermodynamic state functions scale linearly with the amount of water in contact with the macromolecule. This is conveniently estimated as a solvent accessible surface area (ASA) [14]. Given atomic coordinates, the ASA is calculated by rolling a sphere representing the solvent molecule on the van der Waals surface of each atom of the solute. It is a geometric quantity, objective except for the choice of the van der Waals radii of chemical groups. With a water molecule of radius 1.4 Å, there is about one solvent molecule in contact per 10 Å^2 ASA. Relevant thermodynamic quantities are the standard state enthalpy, free enthalpy and heat capacity for transferring a molecule from either the gas phase, the crystalline state or an organic solvent to a water solution. Transfer from the gas phase is simply called hydration. With small molecules, the free enthalpy of hydration may be obtained by measuring the solubility in water, whereas calorimetry yields the corresponding enthalpy and heat capacity changes after appropriate corrections have been made. Similar measurements can be made for transfer from the crystalline solid or from an organic solution. Thanks to the work of physical chemists, thermodynamic parameters for hydration and transfer are known experimentally for many model compounds. With alkanes and most aliphatic compounds, including amino acid sidechains [15], the linear dependence on either the ASA or the number of carbon atoms, is generally excellent. We may write the enthalpy (ΔH^{aliph}) and free enthalpy (ΔG^{aliph}) of hydration of aliphatic groups as:

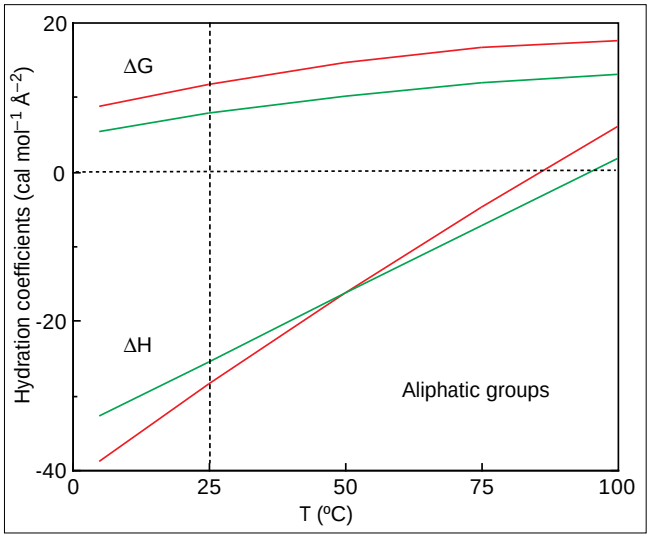
$$\Delta H^{aliph} \sim \eta \, ASA \qquad \Delta G^{aliph} \sim \gamma \, ASA$$

where η and γ are hydration coefficients. If we assume that similar relationships apply to other types of chemical groups and that hydration is additive, we have:

$$\Delta H^{hyd} \sim \Sigma_i \, \eta_i \, ASA_i \qquad \Delta G^{hyd} \sim \Sigma_i \, \gamma_i \, ASA_i$$

The coefficients η_i and γ_i for individual atom types are energies per unit surface, like a surface tension. At 25°C, η is negative and large for aliphatic groups (Fig. 1): the hydration enthalpy favours the solute state because alkanes make van der Waals interactions with water, whilst in the gas phase no interaction is made. However, under the same conditions γ is positive: the hydration free enthalpy is unfavourable, mostly due to the entropic cost of making a cavity in water [16–17]. Alkanes cannot make hydrogen bonds and van der Waals interactions are too weak to offset the entropy change. In addition, η increases quickly with temperature, so that the enthalpy is zero near 90°C and

Figure 1



Hydration of aliphatic groups. Proportionality coefficients η and γ relate the hydration enthalpy ΔH and free enthalpy ΔG to the ASA of aliphatic groups. The entropy change can be derived as $\Delta S = (\Delta H - \Delta G)/T$. Values taken from the Makhatadze-Privalov set of coefficients [19,20,8] (red) and the Oobatake-Ooi set [18,7] (green) are plotted as a function of temperature. The slope of the ΔH curve is ΔCp, the slope of the ΔG curve is $-\Delta S$; both values are positive and ΔCp is large.

unfavourable above. The slope of the η line is the hydration heat capacity which, for nonpolar groups, is positive and large. Aromatic groups behave very much the same, but as they are more polar and interact better with water, η is larger in absolute value for aromatic than aliphatic groups and γ negative throughout the temperature range.

In Figure 1, we show two estimates of the aliphatic η and γ hydration coefficients as a function of temperature, those of Oobatake and Ooi (the OO set) [7,18] and a more recent set of coefficients determined by Makhatadze and Privalov (MP) [8,19,20]. The coefficients differ by 10–30%, largely because the data from different model compounds are fitted differently, and to a lesser extent, because of a different choice of van der Waals radii in estimating ASAs. The discrepancy is about the same for aromatic groups. We may take these discrepancies as a measure of how good the surface area model is and assume that hydration parameters of nonpolar groups are known to within ~20%. This is quite acceptable given the simplicity of the model and we shall discuss below the more serious problem of polar groups.

When the transfer to water is from an organic solvent instead of the gas phase, cavities and van der Waals interactions occur in both states. Thus, both the enthalpy and the entropy changes are less than for hydration. For alkanes, the free enthalpy change is still very much in favour of the less polar solvent. Estimates of the γ transfer coefficient range from 16 cal mol^{-1} Å^{-2} for octanol to 31 cal mol^{-1} Å^{-2} for cyclohexane [21,22] at 25°C. These large values are attributed to the hydrophobic effect which makes water a different solvent from octanol or cyclohexane [23]. The γ transfer coefficient has been correlated with the surface tension at a water–solvent interface, a macroscopic property than can be viewed as an extrapolation to very large areas of what we may observe at the surface of a macromolecule.

The transfer and the hydration methods

The relation of γ to the surface tension is still unclear and beyond our concern in this review [22,24]. On the other hand, γ is often used in estimating $\Delta\Delta$Gs and this needs be discussed. $\Delta\Delta$G is a notation describing, for example, the change in unfolding free enthalpy of a protein in which a point mutation has deleted a buried methylene group, or in the dissociation free enthalpy of a ligand missing the same group. In either case, we expect the deletion to lower the free enthalpy of the state where the methylene group is water accessible and, therefore, to facilitate unfolding or dissociation.

As transferring a methylene group of ASA ~25 Å^2 from octanol to water costs ~0.4 kcal mol^{-1} free enthalpy at 25°C, we may expect that deleting this group from a protein is energetically worth the same amount. Experimental data on the stability of mutant proteins, however, suggest that the octanol value is too low. Many observed $\Delta\Delta$Gs are ~1 kcal mol^{-1} per methylene [25,26], lending weight to the suggestion that the γ transfer coefficient should be raised [22,27]. But then, is the model where $\Delta\Delta$G is a transfer free energy a realistic one? The model requires that nothing but the deletion happens in the folded or associated state, and assumes that octanol is a good model for the environment of the methylene group within the protein. Whereas the first point can be checked by structural studies, the second contradicts what we know of the protein interior. The protein interior is definitely not a liquid, it is ordered and much more dense and compact [28] than an organic solvent, with an average density 60% larger than for octanol. Organic solids are probably a better model than liquids [6]. A methylene group makes more van der Waals interactions inside a protein than in octanol, while the entropic cost of making a cavity must be less. Even if the solvent-to-water transfer model correctly predicted $\Delta\Delta$G, it would fail in its predictions for the enthalpic and entropic components, which the calorimetry of mutant protein unfolding shows to be largely uncorrelated [29].

To further evaluate the transfer model, we should firstly explain the thermodynamic cycle to which it refers. In Figure 2, we show the cycles for the solvent-to-water (transfer) and for the gas-to-water (hydration) methods. In Figure 2a, changes in a state function, such as the unfolding free enthalpy, can be first estimated in the reference organic solvent before the protein (folded and unfolded) is transferred to water. The cycle requires that:

$$\Delta G_{unf} = \Delta G^{trans} + \Delta G^{orga} \quad (1)$$

where ΔG_{unf} is free enthalpy of unfolding in water, ΔG^{orga} is the unfolding free enthalpy in the organic solvent and ΔG^{trans} the balance of transfer values pertaining to the native and unfolded states. Within the surface area model, ΔG^{trans} is simply given by:

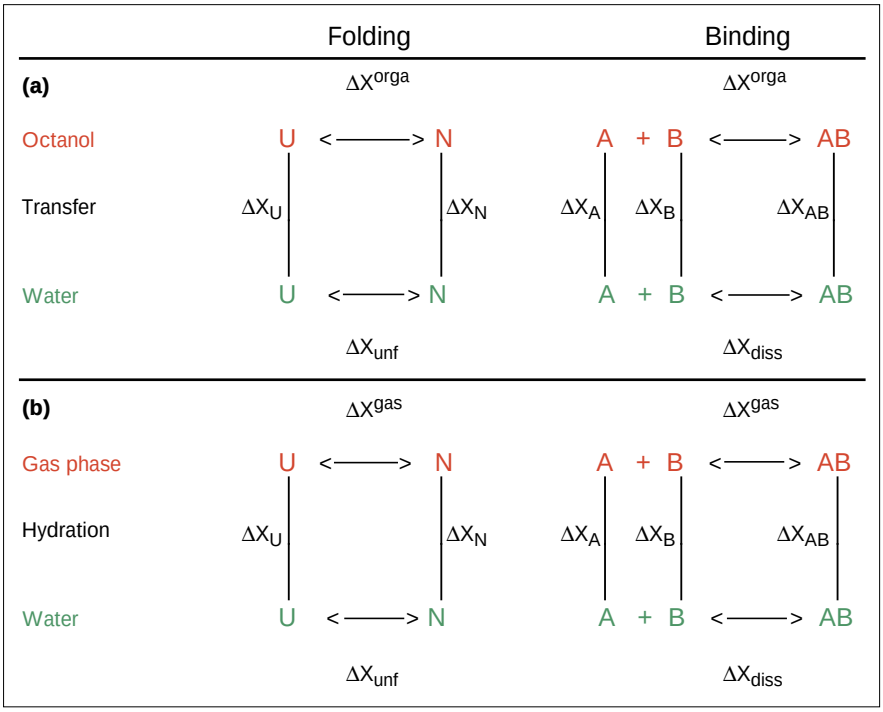
$$\Delta G^{trans} = \Delta G_U - \Delta G_N \sim \Sigma_i \gamma_i \Delta ASA_i \quad (2)$$

Here, ΔASA_i is a buried surface area, the solvent accessible area lost upon folding by chemical groups of type i. Any state function, other than ΔG, may be obtained after replacing γ_i in Equation 2 by appropriate coefficients. Moreover, Figure 2a shows that binding can be analyzed in the same way: components A and B associate in octanol, and are then transferred to water together with the complex AB. Equation 1 yields the dissociation free enthalpy and, in Equation 2, where ΔASA_i now represents the areas buried at the AB interface:

$$\Delta G^{trans} = \Delta G_A + \Delta G_B - \Delta G_{AB} \sim \Sigma_i \gamma_i \Delta ASA_i \quad (3)$$

How do we estimate ΔG^{orga} in Equation 1 and the buried surface areas in Equations 2 and 3? One possible answer to

Figure 2



Thermodynamic cycles for folding and binding. (a) Solvent-to-water transfer. (b) Gas phase-to-water transfer (hydration). X represents a state function of the system: the free enthalpy, enthalpy, entropy or heat capacity. In water, X changes by the quantity ΔX_{unf} upon going from the native state N to the unfolded state U, and by ΔX_{diss} upon dissociation of the complex AB to its components A and B. The corresponding change is ΔX_{orga} for unfolding or dissociation in an organic solvent, or ΔX_{gas} in the gas phase. ΔX_U, ΔX_N, etcetera, are changes that accompany the transfer of U, N, etcetera, to water solution. Only ΔX_{unf} and ΔX_{diss} are directly accessible to experiment.

the first question is to omit ΔG^{orga} altogether. We did just that in interpreting the methylene group deletion: we assumed that the contribution of interactions inside the protein exactly cancelled those with the solvent. This is a gross approximation and one that depends on the solvent. Yet, what else could we do as evaluating free enthalpies is hardly simpler in octanol than in water? Secondly, we need to address the problem of how to estimate buried surface areas. Estimates of the unfolding ΔASA require a model of the unfolded state, usually the extended polypeptide chain. This is a poor representation of the average thermally denatured protein molecule which probably has a smaller ASA, especially if disulphide bonds are present. Moreover, the ASA of a polypeptide depends on mainchain dihedral angles and on sidechain conformations. The literature shows discrepancies between unfolding ΔASAs in excess of 20%. When analyzing binding, ΔASA can be derived from the atomic coordinates of the AB complex alone, provided components A and B associate as rigid bodies. Rigid-body binding is a fair approximation for a number of protein–protein complexes [30], but not for protein–DNA complexes [31]. With peptides and flexible ligands, we also need models of the dissociated components to estimate ΔASA.

The major difficulty of the transfer method, estimating state functions in an organic solvent, is circumvented in the hydration method. Following the practice of Privalov and collaborators, folding is assumed to take place in the gas phase, where interactions between molecules may be ignored, and is followed by a hydration step as shown in Figure 2b. The hydration method can also be applied to binding studies [32]. Unfolding (ΔG_{unf}) or dissociation (ΔG_{diss}) free enthalpies are evaluated by adding the free enthalpy of hydration and the enthalpy in the gas phase:

$$\Delta G_{unf} \text{ or } \Delta G_{diss} = \Delta G^{hyd} + \Delta G^{gas}$$

where ΔG^{hyd} is from Equation 2 or 3.

In the hydration method, state functions are calculated in the gas phase and hydration coefficients applied to the buried areas. One quality of this method is that sets of coefficients exist for thermodynamic parameters other than just the free enthalpy at 25°C. Thus, enthalpy, entropy and heat capacity changes may be calculated and their temperature dependence predicted. However, we must comment on their accuracy. We mentioned above the OO and MP coefficients and noted that they differ by 10–30% for nonpolar groups. With polar groups, they are wholly inconsistent, and MP predicts much larger values of the enthalpy and free enthalpy of polar group hydration in favour of folding or binding. In Figure 3, ΔH_{hyd} and ΔG_{hyd} were calculated at 25°C for four different proteins unfolding from the native to the extended chain. As the values scale linearly with the protein size, to a good approximation, only the average per gram of protein is shown. Molar values may be obtained by multiplying each

number by the protein molecular weight, or by multiplying by ~110 if values per amino acid residue are required.

Energetics of folding and binding

Calculations using the MP and OO coefficient sets agree on the contribution of nonpolar hydration: at 25°C, the enthalpy of nonpolar hydration opposes folding by 12–14 cal g^{-1}, yet its free enthalpy weakly favours folding by 2–3 cal g^{-1} due to a compensating entropy change. Polar hydration opposes folding. The corresponding enthalpy and free enthalpy changes are the largest terms in both calculations, and are a factor of five larger in the MP derivation ($\Delta H^{hyd} \sim -100$ cal g^{-1}). On this scale, observed values of the unfolding enthalpy and free enthalpies are ~0. This is an important point in the analysis of folding energetics, and one that needs comment, as ΔH_{unf} and ΔG_{unf} are the only numbers in the analysis that result from actual experiments on proteins. Like most proteins, those in our test sample are only marginally stable at 25°C. Their unfolding free enthalpies are small, in the usual 5–15 kcal mol^{-1} range which amounts to less than 1 cal g^{-1}. Less well publicised is the fact that at 25°C, unfolding enthalpies are also small: 0–4 cal g^{-1} in our test sample. Due to the large positive unfolding heat capacity, ΔH_{unf} is much smaller and thermal transitions are much less cooperative at 25°C than at 50 or 75°C. ΔH_{unf} may even be negative, which leads to cold denaturation [4].

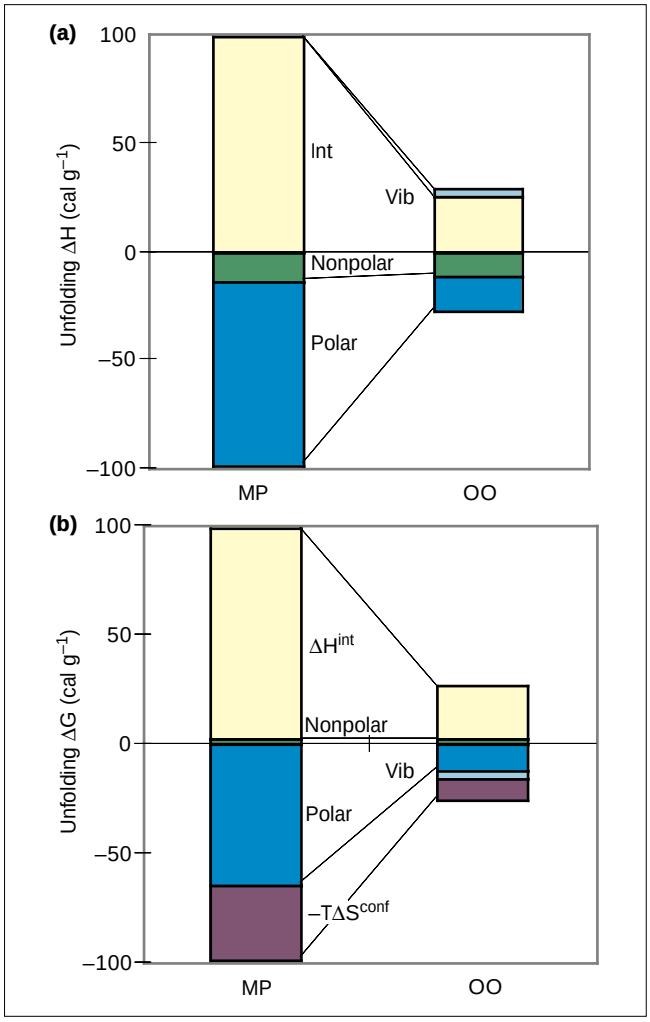
To fit the experiment, the balance of both the enthalpy and the free enthalpy terms must therefore be ~0. In other words, the gas phase terms must cancel the hydration terms. As, in the gas phase, most of the folding enthalpy is from intramolecular van der Waals and electrostatic interactions and most of the folding entropy is conformational, we note these terms ΔH^{int} and ΔS^{conf}, respectively. We now have:

$$\Delta H_{unf} = \Delta H^{int} + \Delta H^{hyd} \sim 0$$

$$\Delta G_{unf} = \Delta H^{int} - T\Delta S^{conf} + \Delta G^{hyd} \sim 0$$

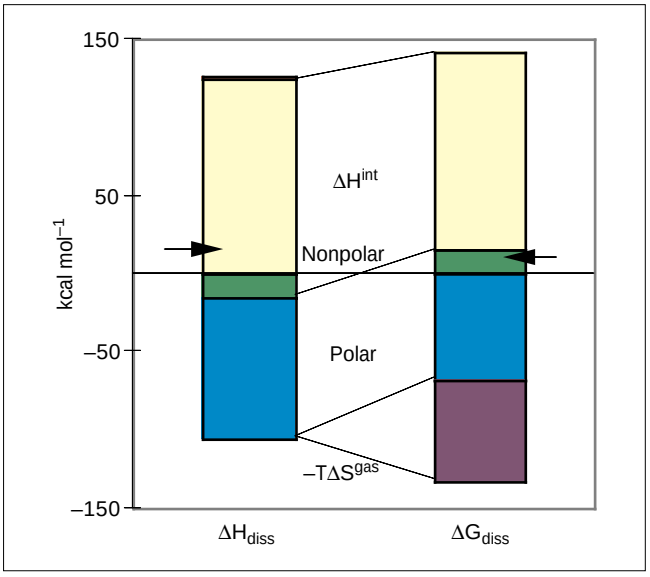
As a consequence, the hydration method gives access to the two very important quantities ΔH^{int} and ΔS^{conf}, albeit it accumulates errors and approximations in their estimate. MP predicts $\Delta H^{int} \sim 100$ cal g^{-1} and $T\Delta S^{conf} \sim 35$ cal g^{-1}, equivalent to 11 kcal mol^{-1} and 4 kcal mol^{-1} per residue, respectively. The alternative analysis that uses the OO coefficients predicts values three to four times smaller. It also has vibrational terms, yet these are comparatively small and the discrepancy originates mostly from the polar hydration term. Obviously, such large differences in estimates of ΔH^{int} and ΔS^{conf} affect the interpretation of folding energetics. Is the average energy of van der Waals and electrostatic interactions 11 kcal per mole of residue as Makhatadze and Privalov conclude, or is it 3 kcal? Does folding reduce the entropy of the polypeptide chain by the equivalent of 4 kcal per mole of residue, or only 1 kcal? At this stage, we might as well abandon the empirical approach and go back to molecular mechanics. Even though the solvent and entropy calculations are beyond the reach of molecular mechanics, it can evaluate conformational energies. This was done with the CHARMM

Figure 3



The energetics of protein folding. The hydration method, represented by the thermodynamic cycle of Figure 2b, has been applied to four proteins: cytochrome c (PDB file, 5CYT); ribonuclease A (PDB file, 7RSA); hen lysozyme (PDB file, 1LZL); and myoglobin (PDB file, 1MBO). Their unfolding enthalpy ΔH **(a)** and free enthalpy ΔG **(b)** have been evaluated. Enthalpy changes due to nonpolar hydration are represented in green, polar hydration (dark blue) and vibrational (light blue; vib). The enthalpy ΔH^{int} of intramolecular van der Waals and electrostatic interactions (light yellow; int) is derived by assuming $\Delta H^{hydr} + \Delta H^{vib} + \Delta H^{int} \sim 0$. The conformational entropy term $-T\Delta S^{conf}$ (purple), is derived by assuming that $\Delta G^{hydr} + \Delta G^{vib} + \Delta H^{int} - T\Delta S^{con} \sim 0$. Data in columns labelled MP are adapted from [8] at 25°C; the data do not include vibrational terms and differ from earlier data calculated with a tripeptide model that grossly overestimates the ΔASA of the unfolded protein [39,40]. For the same reason, data in columns labelled OO are taken from [41] and not from [7].

Figure 4



The energetics of protein–protein association. The hydration method, represented by the thermodynamic cycle of Figure 2b, has been used to analyze hen lysozyme binding to the HyHEL5 Fab fragment (PDB file, 2HFL). The dissociation enthalpy ΔH_{diss} and free enthalpy ΔG_{diss} in the 1 M standard state at 25°C have been evaluated [34]. Coefficients from the OO set were applied to areas buried at the interface to yield the hydration enthalpy and free enthalpy of nonpolar (green) and polar groups (blue). ΔH^{int} (light yellow) is the contribution of van der Waals and electrostatic interactions calculated by molecular mechanics. The balance of the enthalpic terms (left arrow) is 19 kcal mol^{-1}, close to the calorimetric ΔH_{diss} = 22.6 kcal mol^{-1} [36]. The gas phase entropic term $-T\Delta S^{gas}$ (purple) includes estimated contributions of external (rotational/translational) and internal (sidechain and vibrational) degrees of freedom. The balance of the free enthalpy terms is 8 kcal mol^{-1} (right arrow), to be compared with the experimental ΔG_{diss} = 14.5 kcal mol^{-1} derived from the equilibrium constant at 25°C.

force field by Karplus and collaborators [33] in parallel to the MP calculation [8]. For the four proteins in Figure 3, gas phase energy minimization gave unfolding energies that were 75–80% of the ΔH^{int} values predicted by MP. The electrostatic component was only 30% of that predicted by MP, supporting the idea that the latter overestimates the enthalpy of polar group hydration.

The OO set of hydration coefficients were used to perform a similar calculation on three protein–protein complexes where the rigid-body approximation applies [34]. For the lysozyme–HyHEL5 antibody complex [35], the gas phase interaction energy estimated by molecular mechanics was $\Delta H^{int} \sim 125$ kcal mol^{-1}, its hydration counterpart $\Delta H^{hyd} \sim -106$ kcal mol^{-1}. The balance of 19 kcal mol^{-1} is reasonably close to the value of ΔH_{diss} given by isothermal titration calorimetry [36]. As with folding, most of the ΔH^{hyd} of a protein–protein interaction is polar hydration. It is obvious that in this case as well, the MP set would predict much larger values, difficult to reconcile with the calorimetric data. In addition to the enthalpy change, the calculation using the OO coefficients fitted the experimental ΔG_{diss}, which suggests that major contributions to the dissociation entropy were correctly taken into account (Fig. 4). In this sense, this calculation performed better than an earlier calculation based on the solvent-to-water transfer method, which did not evaluate ΔH_{diss} and overestimated ΔG_{diss} [37]. However, the fit to the experimental data on the lysozyme–HyHEL5 complex was rather artificial, as the error bar in each term is large, several kcal mol^{-1}.

At the present stage, the conclusion has to be the same for binding and folding. Whereas the solvent-to-water transfer method is at best qualitative, the hydration method has a sound thermodynamic basis. The hydration method can in principle estimate enthalpies as well as free enthalpies, and fit calorimetric data as well as equilibrium constants. However, the excessive uncertainty concerning polar hydration coefficients makes numerical results unreliable. When a protein folds, many peptide groups and polar sidechains are removed from contact with the water. When a protein binds a ligand, only some of the mainchain is buried, but polar sidechains and polar groups on the ligand (small molecule, peptide or DNA) are extensively buried . Estimates from the literature of the hydration enthalpy and free enthalpy, however, vary grossly for groups containing nitrogen or oxygen atoms [38]; some parameters are lacking altogether and additivity remains to be tested [24]. There is still a major step to be made before converting (square) Ångströms into calories and once again we need the help of physical chemists and more small molecule data to make it safely.

Acknowledgements

This review was largely inspired by discussions with NN Khechinashvili (Puschino, Russia). I am grateful to G Bains (Johns Hopkins University) for critical reading, and for financial support to Université Paris-Sud, Agence Nationale de Recherche contre le SIDA, and Association pour la Recherche contre le Cancer.

References

1. Privalov, P.L. & Potekhin, S.A. (1986). Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.* 131, 4–51.
2. Fisher, H.F. & Singh, N. (1995). Calorimetric methods for interpreting protein–ligand interactions. *Methods Enzymol.* 259, 194–221.
3. Ladbury, J.E. & Chowdhry, B.Z. (1996). Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. *Chem. Biol.* 3, 791–801.
4. Privalov, P.L. & Gill, S.J. (1988). Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39, 191–234.
5. Privalov, P.L. (1989). Thermodynamic problems of protein structure. *Annu. Rev. Biophys. Biophys. Chem.* 18, 47–69.
6. Murphy, K.P. & Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem.* 43, 313–361.
7. Oobatake, M. & Ooi, T. (1993). Hydration and heat stability effects on protein unfolding. *Prog. Biophys. Mol. Biol.* 59, 237–284.
8. Makhatadze, G.I. & Privalov, P.L. (1995). Energetics of protein structure. *Adv. Protein Chem.* 47, 307–425.
9. Honig, B. & Yang, A.S. (1995). Free energy balance in protein folding. *Adv. Protein Chem.* 46, 27–60.

10. van Gunsteren, W.F., Weiner, P.K. & Wilkinson, A.J. (1993). Computer simulations of biomolecular systems. In *Theoretical and Experimental Applications, Vol. 2*. ESCOM, Leiden, Netherlands.
11. Halgren, T.A. (1995). Potential energy functions. *Curr. Opin. Struct. Biol.* **5**, 205–210.
12. Teeter, M. & Case, D. (1990). Harmonic and quasi-harmonic descriptions of crambin. *J. Phys. Chem.* **94**, 8091–8097.
13. Honig, B. & Nicholls, A. (1995). Classical electrostatics in biology and chemistry. *Science* **268**, 1144–1149.
14. Lee, B. & Richards, F.M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.
15. Chothia, C. (1974). Hydrophobic bonding and accessible surface area in proteins. *Nature* **248**, 338–339.
16. Lee, B. (1995). Analyzing solvent reorganization and hydrophobicity. *Methods Enzymol.* **259**, 555–576.
17. Kocher, J.P., Prévost, M., Wodak, S.J. & Lee, B. (1996). Properties of the protein matrix revealed by the free energy of cavity formation. *Structure* **4**, 1517–1529.
18. Ooi, T., Oobatake, M., Némethy, G. & Scheraga, H.A. (1987). Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc. Natl. Acad. Sci. USA* **84**, 3086–3090.
19. Makhatadze, G.I. & Privalov, P.L. (1993). Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration. *J. Mol. Biol.* **232**, 639–659.
20. Privalov, P.L. & Makhatadze, G.I. (1993). Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J. Mol. Biol.* **232**, 660–679.
21. Eisenberg, D. & McLachlan, A.D. (1986). Solvation energy in protein folding and binding. *Nature* **319**, 199–203.
22. Sharp, K.A., Nicholls, A., Fine, R.F. & Honig, B. (1991). Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. *Science* **252**, 106–109.
23. Tanford, C. (1979). Interfacial free energy and the hydrophobic effect. *Proc. Natl. Acad. Sci. USA* **76**, 4175–4176.
24. Ben-Naim, A. (1994). Solvation: from small to macro molecules. *Curr. Opin. Struct. Biol.* **4**, 264–268.
25. Pace, C.N. (1995). Evaluating contribution of hydrogen bonding and hydrophobic bonding to protein folding. *Methods Enzymol.* **259**, 538–554.
26. Matthews, B.W. (1995). Studies on protein stability with T4 lysozyme. *Adv. Protein Chem.* **46**, 249–281.
27. Vajda, S., Weng, Z. & DeLisi, C. (1995). Extracting hydrophobicity parameters from solute partition and protein mutation/unfolding experiments. *Protein Eng.* **8**, 1081–1092.
28. Harpaz, Y., Gerstein, M. & Chothia, C. (1994). Volume changes on protein folding. *Structure* **2**, 641–649.
29. Sturtevant, J.M. (1994). The thermodynamic effects of protein mutations. *Curr. Opin. Struct. Biol.* **4**, 69–78.
30. Janin, J. & Chothia, C.(1990). The structure of protein–protein recognition sites. *J. Biol. Chem.* **265**, 16027–16030.
31. Spolar, R.S., Ha, J.-H. & Record, M.T., Jr. (1994). Coupling of local folding to site-specific binding of proteins and DNA. *Science* **263**, 777–783.
32. Janin, J. (1996). Protein–protein recognition. *Prog. Biophys. Mol. Biol.* **64**, 145–165.
33. Lazaridis, T., Archontis, G. & Karplus, M. (1995). Enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics. *Adv. Protein Chem.* **47**, 231–306.
34. Janin, J. (1995). Elusive affinities. *Proteins* **21**, 30–39.
35. Sheriff, S., *et al.*, & Davies, D.R. (1987). Three-dimensional structure of an antibody–antigen complex. *Proc. Natl. Acad. Sci. USA* **84**, 8075–8079.
36. Hibbits, K.A., Gill, D.S. & Willson, R.C. (1994). Isothermal titration calorimetric study of the association of hen egg lysozyme and the anti-lysozyme antibody Hy-HEL5. *Biochemistry* **33**, 3584–3590.
37. Novotny, J., Bruccoleri, R.E. & Saul, F.A. (1989). On the attribution of the binding energy in the antigen–antibody complexes McPC603, D1.3 and HyHEL-5. *Biochemistry* **28**, 4735–4749.
38. Juffer, A.H., Eisenhaber, F., Hubbard, S.J., Walther, D. & Argos, P. (1995). Comparison of atomic solvation parametric sets: applicability and limitations in protein folding and binding. *Protein Eng.* **4**, 2499–2509.
39. Makhatadze, G.I. & Privalov, P.L. (1994). Hydration effects in protein unfolding. *Biophys. Chem.* **51**, 291–309.
40. Spolar, R.S., Livingstone, J.R. & Record, M.T., Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry* **31**, 3947–3955.
41. Khechinashvili, K.N., Janin, J. & Rodier, F. (1995). Thermodynamics of the temperature-induced unfolding of globular proteins. *Protein Sci.* **4**, 1315–1324.