Protein Science (1997), 6:1976–1984. Cambridge University Press. Printed in the USA.

Empirical free energy calculation: Comparison to calorimetric data

ZHIPING WENG, CHARLES DELISI, AND SANDOR VAJDA
Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215

(RECEIVED February 10, 1997; ACCEPTED May 12, 1997)

Abstract

An effective free energy potential, developed originally for binding free energy calculation, is compared to calorimetric data on protein unfolding, described by a linear combination of changes in polar and nonpolar surface areas. The potential consists of a molecular mechanics energy term calculated for a reference medium (vapor or nonpolar liquid), and empirical terms representing solvation and entropic effects. It is shown that, under suitable conditions, the free energy function agrees well with the calorimetric expression. An additional result of the comparison is an independent estimate of the side-chain entropy loss, which is shown to agree with a structure-based entropy scale. These findings confirm that simple functions can be used to estimate the free energy change in complex systems, and that a binding free energy evaluation model can describe the thermodynamics of protein unfolding correctly. Furthermore, it is shown that folding and binding leave the sum of solute–solute and solute–solvent van der Waals interactions nearly invariant and, due to this invariance, it may be advantageous to use a nonpolar liquid rather than vacuum as the reference medium.

Keywords: binding free energy; free energy calculation; heat capacity; protein unfolding

Calculating the free energy change in protein folding and association is a classical problem in biophysical chemistry. In principle, free energy differences can be obtained by molecular dynamics and Monte Carlo simulations that allow for similar molecules to be interconverted, and the relative free energies determined by perturbation or integration techniques (Mezei & Beveridge, 1986; Reynolds et al., 1992). However, simulation methods are far too expensive computationally for free energy calculation in conformational search, docking, and design (Wilson et al., 1991). The simplest remedy is to neglect solvation and entropic contributions in applications, but it is well known that energy-type target functions are frequently unable to distinguish between correct and incorrect proteins folds (Novotny et al., 1988), or correct and incorrect docked conformations (Shoichet & Kuntz, 1991). An alternative approach is to estimate free energy by empirical methods that are computationally viable and yet can better discriminate between correct and incorrect structures than conformational energy alone. A number of effective free energy functions have been proposed during the last few years (Novotny et al., 1989; Wilson et al., 1991; Horton & Lewis, 1992; Wesson & Eisenberg, 1992; Stouten et al., 1993; Bohm, 1994; Smith & Honig, 1994; Vajda et al., 1994; Holloway et al., 1995; Jackson & Sternberg, 1995; Nauchitel et al., 1995; Verkhivker et al., 1995; Wallqvist et al., 1995; Zhang & Koshland, 1996) and some kind of free energy calculation is quickly becoming the standard in computer-aided molecular design (Ajay & Murcko, 1995).

We have developed a relatively complete empirical free energy function, and evaluated it against a range of structural and thermodynamic data (Vajda et al., 1994, 1995; Gulukota et al., 1996; King et al., 1996; Weng et al., 1996). The free energy change, $\Delta G = G_2 - G_1$, between two states is calculated according to the expression:

$$\Delta G = \Delta E + \Delta G_d - T\Delta S_c + \Delta G_{other}, \quad (1)$$

where ΔE, ΔG_d, and ΔS_c represent the energy change, the desolvation free energy, and the change in conformational entropy, respectively. The last term, ΔG_{other}, includes all other free energy changes associated with translational, rotational, vibrational, cratic, and protonation/deprotonation effects (Novotny et al., 1989; Vajda et al., 1994).

The function has been developed originally for calculating receptor–ligand binding free energies, and we used a number of simplifying assumptions to calculate the free energy terms. It was assumed that binding does not affect the conformational energy of either molecule substantially (Novotny et al., 1989; Vajda et al., 1994). Due to this assumption, the energy change ΔE is reduced to the receptor–ligand interaction energy $E^{r\text{-}l}$, calculated in a reference medium (vacuum or nonpolar liquid). The desolvation free energy, ΔG_d, is obtained by the expression $\Delta G_d = \Delta G^{rl}_{tr} - \Delta G^{r}_{tr} - \Delta G^{l}_{tr}$, where ΔG^{rl}_{tr}, ΔG^{r}_{tr}, and ΔG^{l}_{tr}, denote the free energies of transferring the complex, the receptor, and the ligand, respectively, from water into the reference medium. We assumed that the protein–

Reprint requests to: Sandor Vajda, Department of Biomedical Engineering, Boston University, 44 Cummington St., Boston, Massachusetts 02215; e-mail: vajda@enga.bu.edu.

protein and the protein–solvent interfaces are equally well packed, to the extent that the intermolecular van der Waals (vdW) interactions in the bound state are balanced by interactions with the solvent in the free state (Adamson, 1982; Novotny et al., 1989; Nicholls et al., 1991; Horton & Lewis, 1992; Krystek et al., 1993; Jackson & Sternberg, 1995; Nauchitel et al., 1995). Due to this vdW cancellation, the vdW contributions were removed both from the desolvation free energy ΔG_d, and from the interaction energy $E^{r\text{-}l}$, thereby reducing the latter to its electrostatic component $E_{el}^{r\text{-}l}$. The term ΔG_{other} was considered to be constant; i.e., independent of the detailed structure of the potential complexes (Novotny et al., 1989; Horton & Lewis, 1992; Vajda et al., 1994; Jackson & Sternberg, 1995; Nauchitel et al., 1995).

The free energy function was shown to be useful in a number of applications. The direct evaluation of the method consisted of comparisons against measured binding free energies for molecules that, to a good first approximation, did not change backbone geometry on complexation. Proteases interacting with their inhibitors fall into this category, and we found that the average difference between calculated and measured binding free energies was approximately 1.3 kcal/mol, representing an error of about 10% (Vajda et al., 1994). The free energy has also been used as a target function for docking (King et al., 1996; Weng et al., 1996), and was shown to discriminate between correct and incorrect docked conformations better than traditional selection criteria (Shoichet & Kuntz, 1991).

The main goal of this paper is to further establish the validity of the effective free energy potential (Equation 1) by comparing the calculated free energies to thermodynamic data from calorimetric observations. Most of the available calorimetric data describe the temperature-induced *unfolding* of proteins (Murphy & Freire, 1992; Makhatadze & Privalov, 1993; Privalov & Makhatadze, 1993). However, we want to avoid the calculation of unfolding free energy, i.e., the free energy difference between unfolded and folded states because, for the unfolded state, there exist neither high-resolution structures nor well established structural models (Lazaridis et al., 1995; Makhatadze & Privalov, 1995; Fried & Bromberg, 1996). The calculation of the binding free energy, i.e., the free energy change in a binding reaction, is more tractable, because one can access high-resolution structures of both the reactants (free receptor and ligand) and the product (receptor–ligand complex) experimentally (Connelly, 1994).

In order to avoid the modeling of the unfolded state, required for folding free energy calculation, we consider the effective potential given by Equation 1, and calculate the binding free energies for a number of protein–protein complexes. Because protein folding and association are governed by the same physical forces, under appropriate conditions, Equation 1 should also apply to folding (Khechinashvili et al., 1995). As we will describe, the folding free energy has been expressed as a linear combination of the changes in apolar and polar solvent-accessible surface areas, with coefficients determined from calorimetric data. To compare this expression to the binding free energy, the latter will also be expressed in terms of surface area changes. In spite of being written in the same form, the two models are based on very different decomposition of the free energy, and are parameterized on virtually nonoverlapping data sets. Nevertheless, we will be able to show a good agreement between them, suggesting that there exists a relatively general empirical framework for free energy calculation, and that the validity of an empirical model may extend well beyond the particular class of problems for which it has been developed.

Calorimetry also provides information on some of the individual terms in Equation 1. Because the free energies determined from the calorimetric data by Freire and his group (Murphy & Freire, 1992) do not include changes in conformational entropy, subtracting Equation 1 from the calorimetric expression yields the side-chain entropy term $T\Delta S_c$. We will show that these back-calculated entropy values are in excellent agreement with the structure-based entropy scale of Pickett and Sternberg (1993). Finally, we study the validity of the assumed van der Waals cancellation, and show that near-cancellation of van der Waals interactions can be achieved by selecting a nonpolar or partially nonpolar liquid [e.g., hydrocarbon or octanol (see Vajda et al., 1995)] as the reference medium.

Results

Calorimetric data

Based on the calorimetric observation of temperature-induced protein unfolding, Freire and co-workers (Murphy & Freire, 1992; Xie & Freire, 1994a) expressed the free energy change in terms of the changes in apolar (ΔA_{apol}) and polar (ΔA_{pol}) solvent-accessible surface areas. At $T = 25\,^{\circ}\mathrm{C}$, this expression is given by

$$\Delta\tilde{G}_c^* = 49.6\Delta A_{apol} - 19.1\Delta A_{pol} \tag{2}$$

(see Materials and methods for details). The subscript c in $\Delta\tilde{G}_c^*$ indicates that the free energy expression given by Equation 2 is based on calorimetric measurements. The tilde shows that protonation/deprotonation effects have been removed by performing the calorimetric analysis in buffer solutions, thereby compensating for any heats of ionization of protein groups upon unfolding (Murphy et al., 1993). Freire and co-workers also assumed that Equation 2 does not include any change in the conformational entropy. This claim was based on the assumptions that the conformational entropy is the sole contributor to the residual unfolding entropy at the isoentropic temperature of 385.15 K, where the specific folding entropies of various proteins tend to converge to a single value, and that the temperature dependence of the conformational entropy is relatively weak in the temperature interval considered. Due to these assumptions, using $S(T_{ref}) = 385.15$ K as the reference state implies that the relative entropy $S(T) - S(T_{ref})$ does not include any conformational contribution. As we will show, this conclusion is strongly supported by our results. The superscript $*$ in $\Delta\tilde{G}_c^*$ is used to emphasize that the equation does not include conformational entropy, and hence is not a complete free energy expression.

Free energy in terms of surface area changes

The goal of this section is to study the relationship between the effective free energy potential given by Equation 1, and the calorimetric expression given by Equation 2. In order to avoid the need for modeling the unfolded state of proteins, we calculate the binding free energy for a number of protein–protein complexes, and express the results in terms of the changes in apolar and polar solvent-accessible areas. Because Equation 2 excludes conformational entropy (Murphy et al., 1993), for the comparison we calculate only the first two components of Equation 1, i.e.,

$$\Delta\tilde{G}^* = \tilde{E}_{el}^{r\text{-}l} + \Delta\tilde{G}_d. \tag{3}$$

As before, the superscript $*$ and the tilde indicate that both conformational entropy and protonation/deprotonation effects are ex-

cluded. In order to compare the calculated free energies to unfolding data, we go one step further, and consider all ionizable side chains neutral. As will be shown, the removal of charges is necessary when comparing a free energy expression, based on folding, to another expression based on binding.

The calculation of the desolvation free energy ΔG_d is based on the atomic solvation parameter (ASP) model. Because the side chains are now assumed neutral, the solvation model given by Equation 17 in Materials and methods is reduced to the expression $\Delta\tilde{G}_d = \sigma_{apol}\Delta A_{apol} + \sigma_{pol}\Delta A_{pol}$, where ΔA_{apol} and ΔA_{pol} denote the changes in apolar and polar solvent-accessible surface areas, and the σ's are the corresponding solvation parameters. Because the apolar area, ΔA_{apol}, is essentially determined by the exposed area of C atoms (the small contribution of S is neglected), $\sigma_{apol} = \sigma_C = 30.5 \pm 1.2$ cal/mol/Å^2, where 1.2 cal/mol/Å^2 is the standard deviation for the estimate of the solvation parameter (Vajda et al., 1994). The polar area, ΔA_{pol}, is due to the exposed N/O atoms, and thus $\sigma_{pol} = \sigma_{N/O} = -0.9 \pm 2.5$ cal/mol/Å^2 (Vajda et al., 1994). Therefore, restricting consideration to nonionized side chains, the desolvation free energy is given by

$$\Delta\tilde{G}_d = (30.5 \pm 1.2)\Delta A_{apol} - (0.9 \pm 2.5)\Delta A_{pol}. \quad (4)$$

We calculate the electrostatic interaction energy $\tilde{E}_{el}^{r\text{-}l}$ for a number of protein–protein complexes, and express the results in terms of ΔA_{apol} and ΔA_{pol}. Although 20 complexes will be used in this paper (Table 1), at this point the analysis is restricted to the first 15 complexes that satisfy the condition of rigid-body association relatively well. The last five complexes are left out for various reasons: trypsinogen in 2TGP changes conformation when it complexes with BPTI (Huber & Bode, 1978); the ligands are small and flexible in 3CPA and IV; the insulin dimer 4INS consists of relatively small and flexible molecules; and in 4CPA, the binding site includes Zn^{2+} ions that affect the electrostatic interactions significantly.

Table 2 shows, for the first 15 complexes, the changes in apolar solvent-accessible area ΔA_{apol}, the changes in polar solvent-accessible area ΔA_{pol}, the desolvation free energy ΔG_d, the electrostatic interactions energy $E_{el}^{r\text{-}l}$, the electrostatic interactions energy $\tilde{E}_{el}^{r\text{-}l}$ calculated with nonionized side chains, the conformational entropy change term $T\Delta S_c$, the calculated binding free energy ΔG based on Equation 1, and the measured value ΔG_{meas} of the binding free energy. The origin of the constant term, $\Delta G_{other} = 11.2$, will be discussed further in the paper. A least-squares fit to the $\tilde{E}_{el}^{r\text{-}l}$ values yields the expression

$$\tilde{E}_{el}^{r\text{-}l} = (16.6 \pm 1.4)\Delta A_{apol} - (14.4 \pm 2.5)\Delta A_{pol}. \quad (5)$$

As shown in Figure 1, apart from complexes 1PPF and 2IFF, the values provided by Equation 5 are in good agreement with the original (coulombic) interaction energies. The correlation coefficient for all the 15 points is $r = 0.94$. Adding Equations 4 and 5 yields

$$\Delta\tilde{G}^* = (47.1 \pm 1.8)\Delta A_{apol} - (15.3 \pm 3.5)\Delta A_{pol}. \quad (6)$$

Although the numbers in Equation 2 are slightly outside the error bars of the coefficients in Equation 6, the result shows a good agreement between our free energy model and the calorimetry-based expression.

Side-chain entropy loss from binding data

To a good first approximation, the proteins in the first 15 complexes of Table 1 do not change backbone geometry on complex-

Table 1. *Protein complexes studied*[a]

PDB Code	Resolution (Å)	Protein complexes
1CHO	1.8	Alpha–chymotrypsin complex with turkey ovomucoid third domain (OMTKY3)
1CSE	1.2	Subtilisin carlsberg complex with eglin-c
1NCA	2.5	N9 neuraminidase–NC41 complex with Fab
1PPF	1.8	Human leukocyte elastase complex with OMTKY3
1TEC	2.2	Thermitase complex with eglin-c
2HFL	2.65	IgG1 Fab fragment (HyHel-5) complex with lysozyme
2IFF	2.65	IgG1 Fab fragment (HyHel-5) complex with lysozyme mutant (R68K)
2KAI	2.5	Kallikrein A complex with bowine pancreatic trypsin inhibitor
2PTC	1.9	beta-trypsin complex with pancreatic trypsin inhibitor
2SEC	1.8	Subtilisin carlsberg complex with genetically-engineered N-acetyl eglin-c
2SNI	2.1	Subtilisin novo complex with chymotrypsin inhibitor 2 (CI-2)
3SGB	1.8	Proteinase B from streptomyces griseus complex with OMTKY3
3TPI	1.9	Trypsin complex with BPTI and Ile-Val
4SGB	2.1	Serine proteinase B complex with the potato inhibitor PCI-1
4TPI	2.2	Trypsin complex with the Arg 15 analogue of BPTI
2TGP	1.9	Trypsinogen complex with bovine pancreatic trypsin inhibitor (BPTI)
3CPA	2.0	Carboxypeptidase alpha complex with glycyl-L-tyrosine
IV		Same as 3TPI but the binding is calculated between trypsinogen and Ile-Val
4INS	1.5	Insulin dimer
4CPA	2.5	Carboxypeptidase alpha complex with potato carboxypeptidase A inhibitor

[a]Complexes 2TGP, 3CPA, 4CPA, 4INS, and IV were not used in the binding free energy calculation because 2TGP, 3CPA, 4INS, and IV do not obey rigid-body assumption, whereas 4CPA has Zn^{2+} ions in the binding site. Complexes 2HFL, 2IFF, and 1NCA were not used in the vdW calculation because the crystal structures have low resolution.

Table 2. *Free energy calculations of 15 complexes*[a]

PDB code	ΔG_d	$\tilde{E}_{el}^{r,l}$	$E_{el}^{r,l}$	$-T\Delta S_c$	ΔG	ΔG_{meas}
1CHO	−25.7	−9.9	−17.6	19.0	−15.3	−14.4
1CSE	−25.4	−9.6	−17.5	17.4	−16.5	−13.1
1CNA	−28.5	−5.0	−26.5	32.8	−13.2	−9.7
1PPF	−27.9	−8.6	−6.3	14.7	−10.5	−13.5
1TEC	−27.5	−8.6	−15.0	17.9	−15.6	−14.0
2HFL	−22.7	−5.6	−27.7	26.0	−15.4	−14.2
2IFF	−23.3	−0.6	−23.8	26.5	−11.6	−11.1
2KAI	−21.3	−8.2	−22.3	18.6	−16.0	−12.5
2PTC	−19.4	−6.5	−24.0	17.1	−17.3	−18.1
2SEC	−26.3	−7.8	−16.0	17.3	−16.0	−14.0
2SNI	−30.8	−11.0	−15.5	21.0	−16.3	−15.8
3SGB	−21.4	−8.0	−12.5	13.2	−11.7	−12.7
3TPI	−19.3	−7.0	−27.5	16.6	−21.2	−17.3
4SGB	−23.6	−9.4	−8.5	9.4	−13.7	−11.7
4TPI	−19.4	−6.2	−25.5	16.5	−19.4	−17.3

[a]All energies are in kcal/mol.

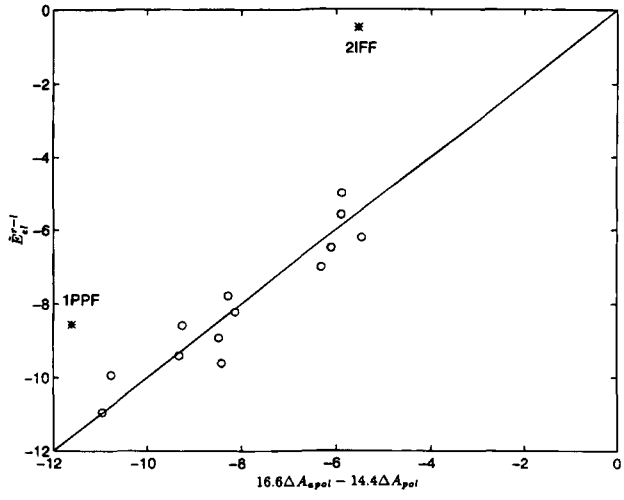

Fig. 1. Electrostatic interaction energy $\tilde{E}_{el}^{r,l}$, calculated with nonionized side chains, versus the linear approximation of $\tilde{E}_{el}^{r,l}$ given by $16.6\Delta A_{apol} - 14.4\Delta A_{pol}$. Points are calculated for the first 15 complexes of Table 1. All values are in kcal/mol.

ation. However, the surface side chains in the free reactants may have conformational degrees of freedom that are lost upon association, resulting in a nonzero conformational entropy term $T\Delta S_c$ in the binding free energy. In the free energy function given by Equation 1, this term has been calculated using the empirical entropy scale derived by Pickett and Sternberg (1993) from the observed rotamer frequencies of surface side chains.

According to Murphy et al. (1993), the calorimetric free energy expression given by Equation 2 does not include any change in conformational entropy. Therefore, subtracting $\Delta\tilde{G}_c^*$ from the measured binding free energy yields the relationship

$$\Delta\tilde{G} - \Delta\tilde{G}_c^* = -T\Delta S_c + \Delta G_{other}. \quad (7)$$

where $\Delta\tilde{G}$ denotes the binding free energy with neutral side chains. This value can be obtained by adding $(\tilde{E}_{el}^{r,l} - E_{el}^{r,l}) + (\Delta\tilde{G}_d - \Delta G_d)$ to the experimentally determined binding free energy ΔG, and thereby correcting for the electrostatic and solvation effects of side-chain charges. The side-chain entropy term, $T\Delta S_c$, can also be calculated using the empirical entropy scale of Pickett and Sternberg (1993), based on the observed frequencies of surface side-chain rotamers (see Materials and methods). As shown in Figure 2, $(\Delta\tilde{G} - \Delta\tilde{G}_c^*)$ correlates very well ($r = 0.97$) with these values of the side-chain entropy loss term $T\Delta S_c$. The line shown in Figure 2 is the result of a least-squares fit in which the slope is fixed at 1.0, and only the intercept b is a free parameter. The fit yields $b = 11.2 \pm 1.8$ kcal/mol which, according to Equation 6, is an estimate of the constant free energy term ΔG_{other}.

Reference medium and van der Waals interactions

The desolvation term ΔG_d to the free energy is defined in terms of the free energies of transferring the molecule from water into a reference state (Vajda et al., 1994). If the reference state is the ideal gas phase, this concept of transfer free energy agrees with the usual definition of the net solvation free energy (Ben-Naim, 1994; Makhatadze & Privalov, 1995). In practice, the transfer free energies are determined from the partition of small organic molecules between their vapor phase and aqueous solution (Makhatadze & Privalov, 1993; Privalov & Makhatadze, 1993) and we will use the notation $\Delta G_d(vw)$ to indicate that the desolvation term is defined using vapor as the reference medium.

In most cases, we used Equation 1 with a nonpolar liquid as the reference state (Vajda et al., 1994). This has clear relevance to protein folding or binding, because desolvation in these processes is due to moving atoms from water into the largely nonpolar environment of the protein interior. The use of a liquid for reference state yields a differential rather than complete hydrophobic contribution. In fact, the vdW interactions between the solute and a reference liquid are *almost* the same as between the solute and the water. Therefore, the desolvation free energy $\Delta G_d(lw)$, defined in

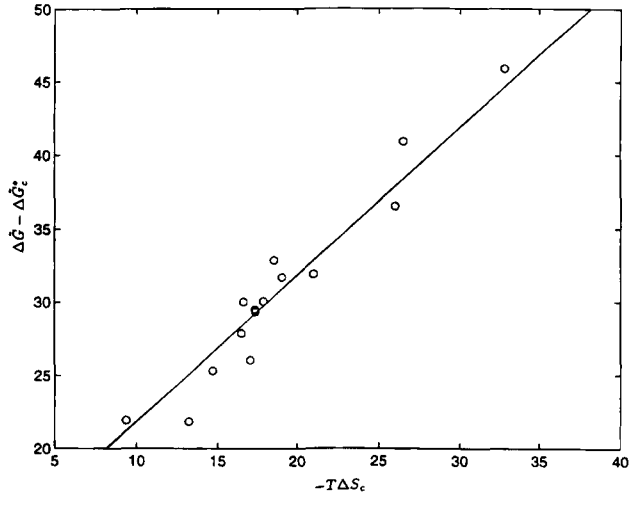

Fig. 2. $\Delta\tilde{G} - \Delta\tilde{G}_c^*$ versus the entropic term $(-T\Delta S_c)$, where $\Delta\tilde{G}$ is the experimentally determined binding free energy corrected for neutral side chains, $\Delta\tilde{G}_c^*$ is the free energy from the calorimetric expression by Murphy and Freire (1992), and ΔS_c is the side-chain entropy loss calculated according to Vajda et al. (1994). All values are in kcal/mol.

terms of the liquid-to-water transfer free energies, includes only a small differential vdW component. As a first-order approximation, we neglected this contribution, and assumed that none of the terms in Equation 1 includes vdW contributions. The removal of vdW interactions reduces substantially the sensitivity of the free energy function to atomic coordinates, resulting in a relatively "smooth" effective potential (Vajda et al., 1994).

If we use a liquid-based desolvation term $\Delta G_d(lw)$, and assume the cancellation of van der Waals interactions, then the energy change ΔE reduces to the electrostatic interaction energy $E_{el}^{r\text{-}l}$, and the free energy is given by $\Delta G = E_{el}^{r\text{-}l} + \Delta G_d(lw) - T\Delta S_c + \Delta G_{other}$. By contrast, using a vapor-based desolvation term $\Delta G_d(vw)$, the binding free energy is given by $\Delta G = E_{el}^{r\text{-}l} + E_{vdW}^{r\text{-}l} + \Delta G_d(vw) - T\Delta S_c + \Delta G_{other}$, where $E_{vdW}^{r\text{-}l}$ is the vdW component of the receptor–ligand interaction energy. Because ΔG should be independent of the reference medium, these two expressions for binding free energy imply that

$$\Delta G_d(lw) - \Delta G_d(vw) = E_{vdW}^{r\text{-}l}. \quad (8)$$

The errors due to assuming the cancellation of vdW interactions can be studied by calculating the terms in Equation 8. Because this relationship does not assume rigid-body association, we can extend the analysis to the last five complexes in Table 1. However, the resolution of the *X*-ray structure is critical when vdW interactions are taken into account, and we leave out the lowest-resolution structures, 1NCA, 2HFL, and 2IFF. Due to its sensitivity to the atomic coordinates, for each complex, the calculation of $E_{vdW}^{r\text{-}l}$ is preceded by 200 steps of unconstrained minimization of the CHARMm potential. This protocol helps to remove the initial differences in resolution and in geometric parameters used by the crystallographers. The desolvation terms $\Delta G_d(lw)$ and $\Delta G_d(vw)$ are calculated using the ASP values listed in Table 3. Figure 3 shows $\Delta G_d(lw) - \Delta G_d(vw)$ as a function of $E_{vdW}^{r\text{-}l}$. Although some points are relatively far from the 45° line shown in the figure, the correlation ($r = 0.94$) suggests that the cancellation of van der Waals effects is a viable first approximation.

Table 3. *Atomic solvation parameters*[a]

Parameters	Reference states	
	Liquid[b]	Vapor[c]
σ_C	30.5 ± 1.2	9.6 ± 3.0
$\sigma_{N/O}$	−0.9 ± 2.5	−119.4 ± 13.1
σ_S	10.0 ± 9.4	45.2 ± 23.1
σ_O	−15.0 ± 7.3	−187.8 ± 38.5
$\sigma_{N'}$	−38.5 ± 4.5	−178.4 ± 30.0

[a] All parameters are in cal/(mol Å^2).
[b] σ_C is determined from hydrocarbon (liquid) to water transfer free energies. All other parameters are from octonal to water transfer free energies of *N*-acetyl amino acid amides (Fauchere & Pliska, 1983).
[c] σ_C is determined from hydrocarbon (vapor) to water transfer free energies. All other parameters are from vapor to water transfer free energies of amino acid side-chain analogues (Wolfenden et al., 1981).

Discussion

Overall agreement

We considered a free energy function that includes molecular mechanics interaction energy and empirical terms representing desolvation and conformational entropy, and compared the function to a simple expression describing calorimetric data on protein unfolding. As described in Materials and methods, the calorimetric data include neither protonation/deprotonation effects nor changes in conformational entropy (Murphy et al., 1993; Xie & Freire, 1994a). We have shown that, under appropriate conditions, the effective free energy function given by Equation 1 agrees with the calorimetric data. The agreement confirms that very simple theoretical models can be used to estimate the free energy in complex systems. This may not seem to be surprising, because many of the effects involved in the calculations scale reasonably well with polar and apolar areas buried. However, we have much stronger results. In fact, transforming Equation 1 into the form of Equation 2, we were able to show not only that the free energy correlates with the surface areas, but also that the two expressions have almost identical coefficients.

The binding free energy model agrees with the calorimetric expression based on unfolding data if and only if all ionizable side chains are considered neutral in the binding free energy calculation. It is well known that there are not too many charged groups in the interior of proteins, and that the few salt bridges do not play an important role in protein folding (Dill, 1990; Tidor & Karplus, 1994; Honig & Nicholls, 1995). By contrast, salt bridges are seen frequently and can be very important in the contact region of complexes. It has been assumed that electrostatic interactions may help to orient the two molecules favorably, thereby substantially increasing the rate of association (Schreiber & Fersht, 1996).

Equation 1 includes the molecular mechanics energy change ΔE. In the binding free energy calculation, we assumed that both the receptor and the ligand are rigid, and hence the internal energy terms (bond stretching, angle bending, torsional, and improper) do not change upon the receptor–ligand association. According to the results presented here, the same free energy function adequately describes protein unfolding, indicating that, in spite of the major

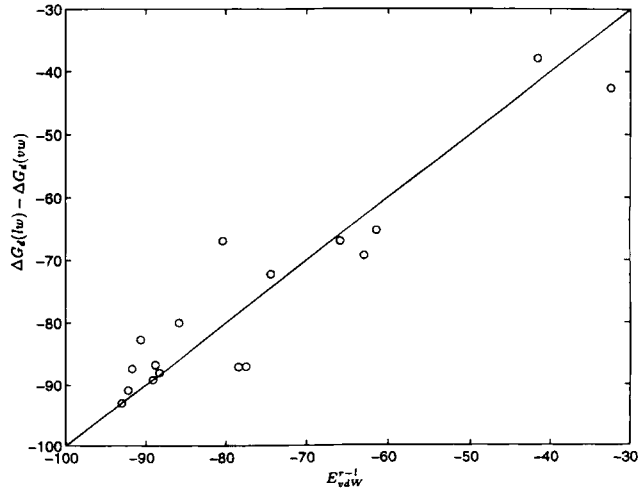

Fig. 3. The difference $\Delta G_d(lw) - \Delta G_d(vw)$ of desolvation free energies, based on liquid-to-water and vapor-to-water transfer data, respectively, as a function of the van der Waals interaction energy $E_{vdW}^{r\text{-}l}$, calculated for the complexes of Table 1. (The the relatively low-resolution structures 1NCA, 2HFL, and 2IFF are omitted from the calculation.) All values are in kcal/mol.

conformational change, the energy terms associated with local deformations (i.e., bond stretching, angle bending, torsional, and improper energy terms) do not contribute substantially to the process. Indeed, it is likely that all potential "hot spots" (i.e., local distortions) are removed in equilibrium, in both the folded and the unfolded states.

Side-chain entropy change

Freire and co-workers (Murphy & Freire, 1992) assumed that the calorimetric free energy expression given by Equation 2 does not include any change in conformational entropy. Our results strongly support this assumption (Fig. 2). If $\Delta\tilde{G}_c^*$ does not include conformational entropy, then subtracting it from the experimentally determined binding free energy ΔG, after correcting for the effect of ionized side chains, should correlate with the side-chain entropy loss $T\Delta S_c$. Figure 2 not only confirms such correlation, but the resulting slope of 1.0 shows that the absolute scale of the side-chain entropies, adopted from the structure-based statistics of Pickett and Sternberg (1993), is also correct.

Fitting a line to the data in Figure 2 with the slope of 1.0 yields the intercept of 11.2 kcal/mol, which is the estimate of the constant free energy term ΔG_{other} in Equation 1. In the most general case of receptor–ligand association, ΔG_{other} includes translational, rotational, vibrational, cratic, and protonation/deprotonation contributions (Vajda et al., 1994). Assuming an ideal gas phase, the loss of rotational and translational entropy and energy can be calculated using simple kinetic arguments (Finkelstein & Janin, 1989). For complexes of proteases with protein inhibitors, the resulting entropy loss is around 16 kcal/mol, with a weak dependence on size (K. Gulukota, pers. comm.). However, about half of the rotational/translational entropy loss will go toward increasing the vibrational entropy upon binding (Finkelstein & Janin, 1989). The remaining 8 kcal/mol net entropy loss is in good agreement with the range of 7–11 kcal/mol, estimated for enzymatic reactions in the liquid phase (Page & Jencks, 1971; Jencks, 1981). In addition, the loss of dilutional or cratic entropy adds about 2 kcal/mol (Kauzmann, 1959). ΔG_{other} also includes 1.8 kcal/mol change in energy due to the loss of three translational and three rotational degrees of freedom in the process of bimolecular association. Adding these contributions yields $\Delta G_{other} \approx 11.8$ kcal/mol. The intercept in Figure 2, 11.2 kcal/mol, is clearly in excellent agreement with this estimate.

In contrast to the other terms in Equation 1, the side-chain entropy does not scale with polar and apolar areas buried. Therefore, the free energy can be described by a linear expression of the form of Equation 2 if and only if all conformational entropy contributions are removed. Furthermore, if we allow for ionizable side chains but remove conformational entropy, the binding free energy can be expressed as a linear combination of the solvent-accessible surface areas of five atomic groups, C, N/O, S, N^+, and O^- (Fig. 4). However, this is not the case if the side-chain entropy change is included, and a linear model fitted to the binding free energies calculated for 15 complexes (Table 2) yields more than 15% average error.

Van der Waals cancellation

Because the empirical model does not consider solvent explicitly, simplifying assumptions were used to describe the interaction between the protein and the surrounding solvent. We assumed that

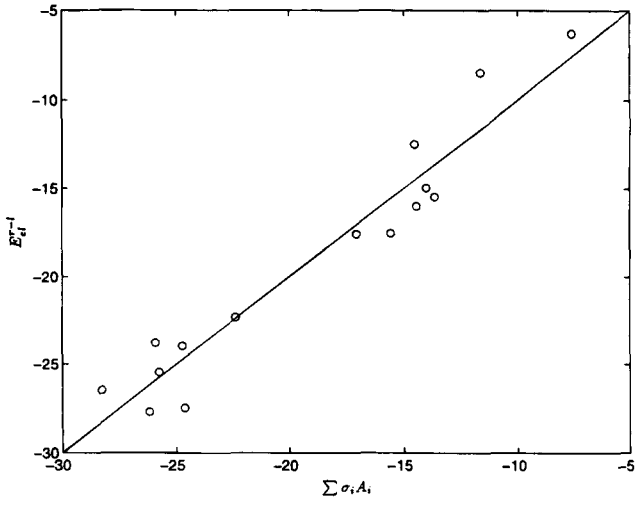

Fig. 4. Electrostatic interaction energy E_{el}^{r-l}, calculated by a coulombic expression, versus the electrostatic interaction energy approximated by the linear expression sum $\sum \sigma_i A_i$, where A_i denotes the contact area of the ith atom type. All energy values are in kcal/mol. From the best fit to the E_{el}^{r-l} data, the parameters (in cal/mol/Å^2) of the linear expression are as follows: σ(C) = −8.5 ± 7.4, σ(N/O) = 35.3 ± 17.2, $\sigma(O^-)$ = 51.9 ± 32.1, $\sigma(N^+)$ = 85.6 ± 72.9, and σ(S) = 81.0 ± 80.5. The standard deviations of the parameters are large because we estimated five parameters from 15 data points.

the solute–solute and solute–solvent interfaces are equally well packed, and hence the intermolecular vdW interactions in the bound state are balanced by interactions with the solvent in the free state.

Although this assumption of vdW cancellation works well for complexes with known X-ray structure (Adamson, 1982; Novotny et al., 1989; Nicholls et al., 1991; Horton & Lewis, 1992; Krystek et al., 1993; Jackson & Sternberg, 1995; Nauchitel et al., 1995), it limits both the accuracy and the range of applicability of the free energy evaluation method. First, although empirical free energy would be an ideal target function in docking and conformational search calculations, the lack of an excluded volume term prevents its direct minimization (Sezerman et al., 1996). Second, a protein–protein interface is not always well packed. In fact, unfavorable vdW interactions or cavities reduce binding, but do not affect the free energy function if vdW cancellation is assumed (Eriksson et al., 1992; Vajda et al., 1994).

In principle, the vdW effects can be taken into account by using desolvation terms based on vapor-to-water transfer free energies, which include vdW interactions between solute and solvent. If ΔG_d includes solute–solvent vdW effects, then the energy term ΔE should include solute–solute vdW interactions. This model has been used in the past, sometimes with limited success. For example, Wilson et al. (1993) found that the protein–protein vdW contributions to the free energy were extremely large and tended to dominate the entire expression. Acceptable agreement with experimentally observed free energies required the use of a small scaling factor to reduce the contribution of vdW effects. However, the magnitude of vdW interactions may have been overestimated due to inadequate sampling and minimization (Wilson et al., 1993). In fact, according to the results presented in Figure 3, the values of receptor–ligand vdW interaction are consistent with the other free energy terms, and can be obtained as differences between vapor-based and liquid-based desolvation terms.

Equation 7 implies that the ASP values should depend heavily on the reference medium (see Table 3). For example, $\sigma_C = 30.5 \pm 1.2$ cal/mol/Å^2 if a nonpolar liquid is used as reference. This value is extracted from liquid-to-water transfer free energies that include only a small differential vdW component (Vajda et al., 1995). By contrast, the use of vapor-to-liquid transfer free energies, which include solute–solvent vdW effects, yield $\sigma_C = 9.6 \pm 3.0$ cal/mol/Å^2 (Wolfenden et al., 1981; Wesson & Eisenberg, 1992). It is frequently overlooked that the two values, 30.5 ± 1.2 cal/mol/Å^2 and 9.6 ± 3.0 cal/mol/Å^2, are not different estimates of the same quantity, but represent very different physical interactions (Cummings et al., 1995; Juffer et al., 1995). Similarly, $\sigma_{N^+} = -38.5 \pm 4.5$ cal/mol/Å^2 for liquid reference state, and $\sigma_{N^+} = -178.4 \pm 30.0$ cal/mol/Å^2 for vapor (see Table 3).

The reference medium, and thus the definition of the solvation free energy, also affects the perceived importance of various factors in protein folding and association. For example, $\sigma_{N/O} = -0.9 \pm 2.5$ cal/mol/Å^2 with liquid reference state. If the reference is vapor, $\sigma_{N/O} = -119.4 \pm 13.1$ cal/mol/Å^2. Therefore, considering a liquid reference state, one might conclude that the polar groups play a relatively minor role compared to the "hydrophobic" effect, i.e., to the contribution of apolar groups (Murphy & Freire, 1992). By contrast, the use of vapor as the reference state gives a huge importance to polar and charged groups relative to the hydrophobic effect (Makhatadze & Privalov, 1995).

Calorimetric studies of protein unfolding

The good agreement between a free energy function and protein unfolding thermodynamics is a potentially important finding. However, we should note that the experimental component in this comparison, i.e., the calorimetry-based free energy expression, is itself subject to some uncertainty. Differential scanning calorimetry of protein unfolding yields the heat capacities of denatured and folded states as functions of the temperature, as well as the enthalpy of transition. All other thermodynamic quantities are determined from these data. The heat capacity, ΔC_p, is the difference between the heat capacities of denatured and folded states, and its determination requires extrapolation to the transition temperature. Although extrapolation is generally not a very reliable procedure, different groups report similar values for the parameters of the linear model $\Delta C_p = \sigma_{apol}\Delta A_{apol} + \sigma_{pol}\Delta A_{pol}$, at least at $T = 25$ °C (Table 4). The parameters obtained by Murphy and Freire (1992) and Spolar et al. (1992) are based on liquid hydrocarbon/amid transfer data, whereas we derived the last row in Table 4 by fitting the linear expression to the ΔC_p values reported for 18 single-domain proteins.

Table 4. *Parameters in linear expressions for ΔC_p of protein unfolding*[a]

References	σ_{apol}	σ_{pol}
Murphy and Freire (1992)[b]	0.45 ± 0.04	−0.26 ± 0.07
Spolar et al. (1992)[c]	0.33 ± 0.04	−0.14 ± 0.04
Makhatadze and Privalov (1995)[d]	0.38 ± 0.08	−0.13 ± 0.16

[a] $\Delta C_p = \sigma_{apol}\Delta A_{apol} + \sigma_{pol}\Delta A_{pol}$. Sign convention is as used by Murphy and Freire (1992), i.e., $\Delta A_{pol} > 0$ in protein denaturation.

[b] See Materials and methods for determination of σ_{apol} and σ_{pol}.

[c] σ_{apol} was derived from liquid hydrocarbon transfer data and σ_{pol} was derived from organic amide transfer data.

[d] Direct fit to ΔC_p values, given by Makhatadze and Privalov (1995) for $T = 25$ °C.

The calculation of other thermodynamic quantities requires assumptions that generally differ from one group to another. For example, not all groups accept that ΔC_p can be expressed in terms of polar and nonpolar surface areas with temperature-independent parameters (Makhatadze & Privalov, 1995). Because the enthalpy ΔH_T of unfolding at temperature T is given by $\Delta H_T = \Delta H_{T_{ref}} + \int_{T_{ref}}^{T} \Delta C_p(T)\, dT$, the assumed temperature dependence of ΔC_p has a substantial effect on ΔH_T. The reference temperature T_{ref} is an additional parameter. Freire and co-workers (Murphy & Freire, 1992) used $T_{ref} = 60$ °C in the enthalpy calculation, and assumed that ΔC_p is independent of the temperature. By contrast, Makhatadze and Privalov (1995) considered a different T_{ref} value for each protein (i.e., the transition temperature), assumed a significant dependence of ΔC_p on the temperature, and strongly suggested partitioning ΔC_p in terms of residue contributions rather than in terms of polar and nonpolar surface areas.

We emphasize that the free energies calculated by Equation 1 can be compared to the calorimetric expression developed by Freire and co-workers, but *cannot* be compared to the free energies of protein unfolding given by Makhatadze and Privalov (1995). The reason is that the latter include substantial contributions due to the loss of both side-chain and backbone conformational entropy. The binding free energy function allows for side-chain entropy loss, but assumes that the backbones are essentially unchanged upon receptor–ligand association. Thus, without introducing a backbone entropy term, we have to restrict consideration to the calorimetric results by Freire and co-workers (Murphy & Freire, 1992; Murphy et al., 1993; Xie & Freire, 1994a), who tried to remove all conformational entropy contributions by selecting the isoentropic temperature $T^*_{ref} = 385.15$ K as reference in the entropy calculation. Our results strongly support the assumed removal of all conformational entropy. This point is very important, because we have shown that the side-chain conformational entropy change does not scale with polar and apolar surface areas buried. Thus, if conformational entropy contributions are included, then the free energy of unfolding can neither be expressed in terms of solvent-accessible surface areas, nor compared to a potential that has been developed for binding free energy calculation. As shown by Privalov and Makhatadze (1993), the simplest model should include residue-specific solvation coefficients rather than a simple partition into apolar and polar contributions.

Materials and methods

Calorimetry of protein unfolding

Freire and co-workers (Murphy & Freire, 1992; Xie & Freire, 1994a) expressed the heat capacity difference, ΔC_p, between folded and unfolded states of a protein in terms of the changes in apolar (ΔA_{apol}) and polar (ΔA_{pol}) solvent-accessible surface areas. If the areas are in Å^2, ΔC_p (in cal/mol) is given by

$$\Delta C_p = (0.45 \pm 0.04)\Delta A_{apol} - (0.26 \pm 0.07)\Delta A_{pol}. \quad (9)$$

The apolar contribution coefficient, 0.45 cal/mol/Å^2, is based on the measured heat capacity increments of apolar atoms from crystalline amino acids into water (Murphy et al., 1990). The polar contribution coefficient, −0.26 cal/mol/Å^2, is based on the thermodynamics of aqueous dissolution of solid cyclic dipeptides (Mur-

phy et al., 1990). The analysis of the specific heat capacity increments upon protein unfolding yields similar contribution coefficients (Ooi et al., 1987; Spolar et al., 1992; Khechinashvili et al., 1995; Makhatadze & Privalov, 1995).

The enthalpy change for protein folding/unfolding has also been expressed as a linear combination of the changes in apolar and polar solvent-accessible areas (Murphy & Freire, 1992; Xie & Freire, 1994a). An analysis of calorimetric data for a variety of proteins, extrapolating all values to $T = 60\,°C$, yields

$$\Delta H_{60} = 31.4\Delta A_{apol} - 8.44\Delta A_{pol}. \quad (10)$$

Equation 10 provides estimates of ΔH_{60}, with an average error of 6% (Xie & Freire, 1994a, 1994b). Equations 9 and 10 permit calculation of the enthalpy change at any other temperature by means of the standard thermodynamic relationship:

$$\Delta H_T = \Delta H_{60} + \Delta C_p(T - 60). \quad (11)$$

Taking into account that in condensed phases the enthalpy and conformational energy are essentially identical, at $T = 25\,°C$:

$$\Delta E = 15.6\Delta A_{apol} + 0.7\Delta A_{pol}. \quad (12)$$

The entropy change in protein folding/unfolding can be written as

$$\Delta S = \Delta S_d + \Delta S_c + \Delta S_{other}, \quad (13)$$

where ΔS_d and ΔS_c denote the desolvation entropy and the conformational entropy, respectively (Vajda et al., 1994). These two terms contribute more than 95% of the total entropy of unfolding (Murphy & Freire, 1992), with a remaining small term, ΔS_{other} due to some vibrational and cratic effects (Murphy & Freire, 1992; Xie & Freire, 1994a). The conformational entropy change, ΔS_c, is due to the change in the conformational degrees of freedom for the side chains and possibly for the the backbone (Creamer & Rose, 1992; Murphy et al., 1993; Pickett & Sternberg, 1993; Lee et al., 1994; Stites & Pranata, 1995). Based on the general expression

$$\Delta S_T = \Delta S_{T^*_{ref}} + \int_{T_{ref}}^{T} \Delta C_p/T\, dT, \quad (14)$$

Freire and co-workers (Murphy & Freire, 1992; Xie & Freire, 1994a) selected the isoentropic temperature, 385.15 K, as T_{ref} for the entropy. At the isoentropic temperature, the specific folding entropies of various proteins converge to a single value. They assumed that, at 385.15 K, after correction for protonation effects, the total entropy of protein unfolding is very close to the conformational entropy change, and thus have $\Delta S_d \approx 0$ (Murphy & Freire, 1992; Xie & Freire, 1994a). They also assumed that the temperature dependence of the conformational entropy can be neglected. Based on Equation 13 and neglecting ΔS_{other}, the two assumptions imply that $\Delta S - \Delta S_{T_{ref}}$ is the desolvation entropy. Thus, by Equation 14

$$\Delta S_d = \Delta C_p \ln(T/385.15). \quad (15)$$

Using Equation 9 to express ΔC_p in terms of surface areas, at $T = 25\,°C$, Equation 15 yields

$$-T\Delta S_d = 34.3\Delta A_{apol} - 19.8\Delta A_{pol}. \quad (16)$$

Finally, adding Equations 12 and 16, we have the free energy expression given as Equation 2 earlier in the paper.

Free energy calculation

Assuming rigid-body association and the cancellation of van der Waals interactions, the energy change ΔE in Equation 1 is reduced to the electrostatic interaction energy, $E_{el}^{r,l}$ (Vajda et al., 1994). The $E_{el}^{r,l}$ values were calculated using version 19 of the CHARMm force field (Brooks et al., 1983) with a distance-dependent dielectric coefficient $\epsilon = 4r$, and nonbonded cutoff of 17 Å. Only polar hydrogens were used. To refine the complexes before free energy evaluation (i.e., to remove potential van der Waals clashes or substantial flexible deformations), we performed 200 steps of minimization using the CHARMm potential with mass-weighted harmonic constraints (force constant, 20 kcal/mol/Å^2), applied to all non-hydrogen atoms.

The desolvation free energy ΔG_d in Equation 1 was obtained from the free energies of transferring the reactants and the products of the binding reaction from water into a reference medium (Vajda et al., 1994). Each transfer free energy was calculated using the simple ASP model:

$$\Delta G_{tr} = \sum \sigma_i \Delta A_i, \quad (17)$$

where ΔA_i denotes the solvent-accessible surface area of the ith atomic group, and σ_i is the corresponding ASP (Eisenberg & McLachlan, 1986). The ASP values given in Table 3 were used (Vajda et al., 1994).

The maximum conformational entropy S_c of each side chain was calculated by the classical expression $S_c = -R\sum_i p_i \ln(p_i)$, where p_i denotes the probability of the ith rotamer (Pickett & Sternberg, 1993). In the binding free energy calculation by Equation 1, we assumed that, upon association, the entire side-chain entropy is lost, i.e., $\Delta S_c = S_c$ if the change ΔA_t in the total solvent-accessible surface area of the side chain is more than 60% of the standard side-chain surface area A_t^*. Otherwise, the entropy loss was scaled according to $\Delta S_c = \alpha S_c$, where $\alpha = \Delta A_t/(0.6A_t^*)$. It is important to note that, although the entropy losses depend on changes in side-chain surface surface areas, the proportionality coefficients are side-chain specific, and thus the side-chain entropy loss does not simply scale with polar and apolar areas buried.

Acknowledgments

This research has been supported partially by the Donors of The Petroleum Research Fund, administered by the American Chemical Society, and by grant DBI-9630188 from the National Science Foundation.

References

Adamson AW. 1982. *Physical chemistry of surfaces*. New York: J. Wiley.

Ajay, Murcko MA. 1995. Computational methods to predict binding free energy in ligand–receptor complexes. *J Med Chem 38*:4953–4967.

Ben-Naim A. 1994. Solvation: From small to macro molecules. *Curr Opin Struct Biol 4*:264–268.

Bohm HJ. 1994. The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. *J Comput-Aid Mol Design 8*:243–256.

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem 4*:187–217.

Connelly PR. 1994. Acquisition and use of calorimetric data for prediction of the thermodynamics of ligand-binding and folding reactions of proteins. *Curr Opin Biotech 5*:381–388.

Creamer TP, Rose GD. 1992. Side-chain entropy opposes alpha-helix formation but rationalizes experimentally determined helix-forming propensities. *Proc Natl Acad Sci USA 89*:5937–5941.

Cummings MD, Hart TN, Read RJ. 1995. Atomic solvation parameters in the analysis of protein–protein docking results. *Protein Sci 4*:2087–2099.
Dill KA. 1990. Dominant forces in protein folding. *Biochemistry 29*:7133–7155.
Eisenberg D, McLachlan AD. 1986. Solvation energy in protein folding and binding. *Nature (Lond) 319*:199–203.
Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Mattews BW. 1992. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science 255*:172–183.
Finkelstein AV, Janin J. 1989. The price of lost freedom: Entropy of biomolecular complex formation. *Protein Eng 3*:1–3.
Fried MG, Bromberg JL. 1996. Changes in protein hydration during thermal denaturation of ribonuclease A (Abstract). *Tenth Symposium of the Protein Society 1*:81.
Gulukota K, Vajda S, DeLisi C. 1996. Peptide docking using dynamic programming. *J Comput Chem 17*:418–428.
Holloway MK, Wai JM, Halgren TA. 1995. A priori predictions of activity for HIV-1 protease inhibitors employing energy minimization in the active site. *J Med Chem 38*:305–317.
Honig B, Nicholls A. 1995. Classical electrostatics in biology and chemistry. *Science 268*:1144–1149.
Horton N, Lewis M. 1992. Calculation of the free energy of association for protein complexes. *Protein Sci 1*:169–181.
Huber R, Bode W. 1978. Structural basis of the activation and action of trypsin. *Acc Chem Res 11*:114–121.
Jackson RM, Sternberg MJ. 1995. A continuum model for protein–protein interactions: Application to the docking problem. *J Mol Biol 250*:258–275.
Jencks WP. 1981. On the attribution and additivity of binding energies. *Proc Natl Acad Sci USA 78*:4046–4050.
Juffer AH, Eisenhaber F, Hubbard SJ, Walther D, Argos P. 1995. Comparison of atomic solvation parametric sets: Applicability and limitations in protein folding and binding. *Protein Sci 4*:2499–2509.
Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv Prot Chem 14*:1–63.
Khechinashvili NN, Janin J, Rodier F. 1995. Thermodynamics of the temperature-induced unfolding of globular proteins. *Protein Sci 4*:1315–1324.
King BL, Vajda S, DeLisi C. 1996. Empirical free energy as a target function in docking and design: Application to HIV-1 protease inhibitors. *FEBS Lett 384*:87–91.
Krystek S, Stouch T, Novotny J. 1993. Affinity and specificity of serine endopeptidase–protein inhibitor interactions. *J Mol Biol 234*:661–679.
Lazaridis T, Archontis G, Karplus M. 1995. Enthalpic contributions to protein stability: Atom-based calculations and statistical mechanics. *Adv Prot Chem 47*:231–306.
Lee KH, Xie D, Freire E, Amzel LM. 1994. Estimation of changes in side chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins Struct Funct Genet 20*:68–84.
Makhatadze GI, Privalov PL. 1993. Contribution of hydration to protein folding thermodynamics I. The enthalpy of hydration. *J Mol Biol 232*:639–659.
Makhatadze GI, Privalov PL. 1995. Energetics of protein structure. *Adv Prot Chem 47*:307–425.
Mezei M, Beveridge DL. 1986. Free energy simulations. *Ann N Y Acad Sci 494*:1–23.
Murphy KP, Freire E. 1992. Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv Prot Chem 43*:313–361.
Murphy KP, Privalov PL, Gill SJ. 1990. Common features of protein unfolding and dissociation of hydrophobic compounds. *Science 247*:559–561.
Murphy KP, Xie D, Garcia KC, Amzel LM, Freire E. 1993. Structural energetics of peptide recognition: Angiotensin-ii/antibody binding. *FEBS Lett 15*:113–120.
Nauchitel V, Villaverde MC, Sussman F. 1995. Solvent accessibility as a predictive tool for the free energy of inhibitor binding to the HIV-1 protease. *Protein Sci 4*:1356–1364.
Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association—Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct Funct Genet 11*:281–296.
Novotny J, Bruccoleri RE, Saul FA. 1989. On the attribution of binding energy in the antigen–antibody complexes mcpc 603, d1.3 and hyhel-5. *Biochemistry 28*:4735–4749.
Novotny J, Rashin AA, Bruccoleri RE. 1988. Criteria that discriminate between native proteins and incorrectly folded models. *Proteins Struct Funct Genet 4*:19–30.
Ooi T, Oobatake M, Nemethy G, Scheraga HA. 1987. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc Natl Acad Sci USA 84*:3086–3090.
Page MI, Jencks WP. 1971. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc Natl Acad Sci USA 68*:1678–1683.
Pickett SD, Sternberg MJ. 1993. Empirical scale of side-chain conformational entropy in protein folding. *J Mol Biol 231*:825–839.
Privalov PL, Makhatadze GI. 1993. Contribution of hydration to protein folding thermodynamics II. The entropy and Gibbs energy of hydration. *J Mol Biol 232*:660–679.
Reynolds CA, King PM, Richards WG. 1992. Free energy calculations in molecular biophysics. *Molecular Physics 76*:251–275.
Schreiber G, Fersht AR. 1996. Rapid, electrostatically assisted association of proteins. *Nature Struct Biol 3*:427–431.
Sezerman OU, Vajda S, DeLisi C. 1996. Free energy mapping of class I MHC molecules and structural determination of bound peptides. *Protein Sci 5*:1272–1281.
Shoichet BK, Kuntz ID. 1991. Protein docking and complementarity. *J Mol Biol 221*:327–346.
Smith KC, Honig B. 1994. Evaluation of the conformational free energies of loops on proteins. *Proteins Struct Funct Genet 18*:119–132.
Spolar RS, Livingstone JR, Record MT 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.
Stites WE, Pranata J. 1995. Empirical evaluation of the influence of side chains on the conformational entropy of the polypeptide backbone. *Proteins Struct Funct Genet 22*:132–140.
Stouten PF, Frommel C, Nakamura H, Sander C. 1993. An effective solvation term based on atomic occupancies for use in protein simulations. *Molecular Simulation 10*:97–120.
Tidor B, Karplus M. 1994. The contribution of vibrational entropy to molecular association. *J Mol Biol 238*:405–414.
Vajda S, Weng Z, DeLisi C. 1995. Extracting hydrophobicity parameters from solute partition and protein mutation/unfolding experiments. *Protein Eng 8*:1081–1092.
Vajda S, Weng Z, Rosenfeld R, DeLisi C. 1994. Effect of conformational flexibility and solvation on receptor–ligand binding free energies. *Biochemistry 33*:13977–13988.
Verkhivker G, Appelt K, Freer ST, Villafranca JE. 1995. Empirical free energy calculations of ligand–protein crystallographic complexes. I. Knowledge-based ligand–protein interaction potentials applied to the prediction of human immunodeficiency virus 1 protease binding affinity. *Protein Eng 8*:677–691.
Wallqvist A, Jernigan RL, Covell DG. 1995. A preference-based free energy parameterization of enzyme–inhibitor binding. Applications to HIV-1-protease inhibitor design. *Protein Sci 4*:1881–1903.
Weng Z, Vajda S, DeLisi C. 1996. Prediction of complexes using empirical free energy functions. *Protein Sci* 5:614–626.
Wesson L, Eisenberg D. 1992. Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci 1*:227–235.
Wilson C, Gregoret LM, Agard DA. 1993. Modeling side-chain conformation for homologous proteins using an energy-based rotamer search. *J Mol Biol 229*:996–1006.
Wilson C, Mace JE, Agard DA. 1991. Computational method for the design of enzymes with altered substrate specificity. *J Mol Biol 220*:495–506.
Wolfenden R, Andersson L, Cullis PM, Southgate CC. 1981. Affinities of amino acid side chains for solvent water. *Biochemistry 20*:849–855.
Xie D, Freire E. 1994a. Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. *Proteins Struct Funct Genet 19*:291–301.
Xie D, Freire E. 1994b. Structure based prediction of protein folding intermediates. *J Mol Biol 242*:62–80.
Zhang T, Koshland DE Jr. 1996. Computational method for relative binding energies of enzyme–substrate complexes. *Protein Sci 5*:348–356.