Energetics of Cyclic Dipeptide Crystal Packing and Solvation

G. Patrick Brady and Kim A. Sharp
Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059, USA

ABSTRACT Calculations of the thermodynamics of transfer of the cyclic alanine-alanine (cAA) and glycine-glycine (cGG) dipeptides between the gas, water, and crystal phases were carried out using a combination of molecular mechanics, normal mode analysis, and continuum electrostatics. The experimental gas-to-water solvation free energy and the enthalpy of gas-to-crystal transfer of cGG are accurately reproduced by the calculations. The enthalpies of cGG and cAA crystal-to-water transfer are also close to the experimental values. A combination of experimental data and normal mode analysis of cGG provides an accurate estimate of the association entropy penalty (loss of rotational and translational entropy and gain in vibrational entropy) for "binding" in the crystalline phase of −14.1 cal/mol/K. This is a smaller number than most previous theoretical estimates, but it is similar to previous experimental estimates. Calculated entropies of the crystal phase underestimate the experimental entropy by about 15 cal/mol/K because of neglect of longe-range lattice motions. Comparison of the intermolecular interactions in the crystals of cGG and cAA provides a possible explanation of the puzzling decrease in enthalpy, with increasing hydrophobicity seen previously for both cyclic dipeptide dissolution and protein unfolding. This decrease arises from a favorable long-range electrostatic interaction between dipeptide molecules in the crystals, which is attenuated by the more hydrophobic side chains.

INTRODUCTION

Protein folding and protein-protein binding equilibria are determined by a balance between various interactions, which usually results in a net free energy in the range of several to tens of kilocalories per mole. Useful insights into folding and binding energetics can be obtained by analyzing this balance, both in terms of entropic versus enthalpic interactions, and in terms of protein-protein interactions versus protein-solvent interactions. In addition, binding involves an entropically unfavorable loss of translational and rotational freedom for the two reactants (in a binary binding reaction), because the relative positions and orientations of the two molecules become greatly restricted with respect to each other in the complex. This net loss of three translational and three rotational degrees of freedom along with the concomitant gain in six degrees of freedom describing relative motions of the binding partners within the complex is referred to here as the association entropy contribution. Determination of the association term is a prerequisite for any attempt to calculate absolute binding energies.

The theoretical framework for describing the association entropy contribution is well established (Gilson et al., 1996; Hill, 1985; McQuarrie, 1976). The magnitude of this contribution in particular binding reactions is, however, difficult to determine accurately. Various theoretical and experimental estimates (Steinberg and Scheraga, 1963; Page and Jencks, 1971; Chothia et al., 1976; Janin and Chothia, 1978; Finkelstein and Janin, 1989; Dwyer and Bloomfield, 1981; Ben-Shaul et al., 1996; Tidor and Karplus, 1994; Horton and Lewis, 1992; Peitzsch and McLaughlin, 1993; Searle and Williams, 1992; Murphy et al., 1994; Novotny et al., 1989) show a wide range of values. Janin and Chothia (1978) provide a high-end theoretical estimate of the association entropy penalty of −57 to −74 e.u. for a medium-sized protein. The estimate from Horton and Lewis (1992), based on analysis of binding data, provides a low-end estimate of about −20 e.u. For binding to a membrane, where the ligand is less conformationally restrained, the association entropy loss is believed to be lower, with a recent theoretical estimate for α-helix/membrane binding of about −12 e.u. (Ben-Shaul et al., 1996). Experimental data on binding of a fatty acid to a lipid membrane provides an even lower estimate, of >-3 e.u. (Peitzsch and McLaughlin, 1993).

In spite of the studies summarized above, accurate estimates of the association entropy in a thermodynamically well-characterized system are almost nonexistent. The cyclic dipeptides, however, provide a suitable system for the study of this. Measurement of the thermodynamics of dissolution of a homologous series of cyclic dipeptides with aliphatic side chains (glycine, alanine, valine, leucine, and isoleucine) has been carried out by Murphy and co-workers (Murphy and Gill, 1990a,b, 1991; Murphy et al., 1990), providing detailed information on the energetics of transfer of dipeptides from crystal to water. The aliphatic cyclic dipeptides crystallize in the anhydrous state, and the structures of two of the dipeptides studied by Murphy et al., cyclic dialanine (cAA) and cyclic diglycine (cGG), are known to very high resolution. In addition, the gas-to-crystal equilibrium has been studied for cGG (Seki et al., 1956). For this molecule, the existence of data for the gas-to-crystal, crystal-to-water, and, by extension, gas-to-water equilibrium allows one to separate the association

Received for publication 5 July 1996 and in final form 12 November 1996.
Address reprint requests to Dr. Kim A. Sharp, Department of Biochemistry and Biophysics, University of Pennsylvania, 37th and Hamilton Walk, Philadelphia, PA 19104-6059. Tel.: 215-573-3506; 215-898-4217; E-mail: sharp@crystal.med.upenn.edu.

0006-3495/97/02/913/15 $2.00

process by means of a thermodynamic cycle into a solvation term (gas-to-water) and an intermolecular term (gas-to-crystal). The transfer enthalpies and entropies are also known for cGG, allowing one to analyze the energetics from this angle as well. Both gas-to-crystal and water-to-crystal transfer include an association entropy term due to restriction of translational and rotational motion in the crystal, which is analogous to the asssociation entropy loss that occurs in binding. A further advantage of cyclic dipeptides is that they are small and rigid. Thus conformational entropy changes arise from vibrational motions. In larger peptides and proteins, significant entropy changes arise from conformational flexibility, hindered rotational motions, chain entropy, etc. All of these factors suggest that a combination of experimental data and theoretical calculations can be used to study in detail the thermodynamics of solvation and association in cyclic dipeptides.

Another feature of the cyclic dipeptide systems that makes them of interest is their heat capacity behavior upon crystal dissolution into water. A plot of the enthalpy of dissolution versus the molar heat capacity change (ΔC_p) for the cyclic dipeptides with aliphatic side chains shows a large negative slope of about −46K (Murphy et al., 1990). In this behavior, dipeptide dissolution resembles protein unfolding, which has a slope of −63K (Murphy et al., 1990). This similarity is of interest for two reasons. The first is that it suggests that, thermodynamically, the packing in the protein interior resembles that of solid cyclic dipeptides. Secondly, the protein and cyclic dipeptide enthalpy/heat capacity behavior is surprising in that it more closely resembles the aqueous dissolution behavior of a hydrophobic gas than a hydrophobic liquid or solid. In particular, hydrophobic liquid or solid dissolution in water gives a very small or positive enthalpy/heat capacity slope, whereas hydrophobic gas dissolution gives a large negative slope. This observation has been the subject of some discussion (Lee, 1991; Murphy and Gill, 1991; Yang et al., 1992). The behavior of the hydrophobic gas can be attributed to increasingly favorable enthalpic interactions with water as solute size, hydrophobicity, and, hence, ΔC_p increase. For the liquid case, the enthalpic interactions in the liquid phase and in water approximately balance out, regardless of solute size (Yang et al., 1992). To explain why proteins have increasingly favorable enthalpies on interaction with water as the hydrophobic area exposed to water upon unfolding increases, we previously suggested that there was a concomitant increase in unfavorable desolvation of polar groups in the folded protein state that always occurs upon burial of hydrophobic groups (Yang et al., 1992). Calculation of the polar desolvation free energy supports this explanation. Detailed calculations on the cyclic dipeptide systems, which have the same behavior, provide a chance to further test this idea, but in a system that is better characterized. In addition, calculation of the interactions in the dipeptide crystal phase may provide general insight into the interactions in the protein interior.

The goal of this paper was to perform detailed calculations of the water/gas/crystal transfer energetics for cyclic dipeptides using current classical empirical potential functions and to compare the results to experimental data, both to understand the balance of interactions involved and to provide a test of the abilities of current computational methods. Two aspects were of particular interest: 1) Obtaining an accurate estimate of the association entropy for the cyclic dipeptide immobilization in the crystal phase, either from calculations or by a combination of calculation and experimental data. ii) Determining the balance of intermolecular and solvation forces in the crystal-water transfer process, not only for understanding the cyclic dipeptide dissolution data, but for obtaining clues about the corresponding interactions in protein folding.

THEORY

Consider the transfer of the cyclic dipeptides cyclic diglycine (cGG) and cyclic dialanine (cAA) between the anhydrous crystal phase, the gas phase, and the aqueous solution phase. The condition for equilibrium is that the dipeptide chemical potential be the same in each phase:

$$\mu_g = \mu_x = \mu_s, \tag{1}$$

where the subscripts g, x, and s denote the gas, crystal, and aqueous solution phases, respectively. To compute the free energy of transfer between any two of these phases for comparison with experimental data, we require expressions for the chemical potential in each phase. For the gas phase in the dilute ideal gas limit (Hill, 1985, 1986),

$$\begin{aligned}\mu_g &= -kT \ln(q^{rot} q^{trans} q_g^{vib} q_g^{int}) \\ &= -kT \ln\left(\frac{q^{rot} q_g^{vib} q_g^{int}}{\Lambda^3}\right) + kT \ln \rho_g = \mu_g^o + kT \ln \rho_g,\end{aligned} \tag{2}$$

where q^{rot}, q_g^{vib}, and q_g^{int} are the rotational, vibrational, and internal energy partition functions, respectively. The internal energy partition function, q_g^{int}, is the Boltzmann factor of the potential energy of the molecule at its gas phase minimum, $q^{int} = \exp(-U_g^{min}/kT)$, and the vibrational partition function accounts for thermal fluctuations of the molecule about this minimum. This contribution arises from intramolecular, primarily bonded, interactions. The translational partition function is given by $q^{trans} = 1/(\rho_g \Lambda^3)$, where ρ_g is the number density of dipeptide in the gas phase, and $\Lambda = \sqrt{h^2/2\pi mkT}$ is the thermal de Broglie wavelength, where h and k are Planck's and Boltzmann's constants, respectively; T is the absolute temperature; and m is the mass of the cyclic dipeptide. μ_g^o, which contains all of the concentration-independent terms, is the standard state ideal gas chemical potential. Note that the arguments of both logarithms in Eq. 2 have units of concentration and, hence, ρ_g and $1/\Lambda^3$ must be expressed in the same units. For the dipeptide in dilute

aqueous solution (Ben-Naim and Marcus, 1984; McQuarrie, 1976),

$$\mu_s = \mu_g^\circ + \Delta G^{g\to s} + kT \ln \rho_s, \quad (3)$$

where ρ_s is the concentration of peptide in the aqueous phase and $\Delta G^{g\to s}$ is the chemical coupling (solvation) work to introduce one solute molecule into a fixed position in water, comprising electrostatic and nonelectrostatic (hydrophobic and cavity) contributions. As is customary, Eq. 3 accounts for any changes in internal, bonded interactions due to interactions with the solvent as part of the solvation term (Ben-Naim and Marcus, 1984; Holtzer, 1995). For the cyclic dipeptide in the pure crystal phase, the chemical potential is just the molar Gibbs free energy, obtained from standard thermodynamic expressions (Hill, 1986; McQuarrie, 1976) as

$$\mu_x = \frac{G_x}{N} = -kT \ln(q_x^{vib} q_x^{int} q^{lat} e^{-U_{ext}/kT}) + P\bar{V}_x, \quad (4)$$

where q^{lat} is the crystal lattice vibrational partition function and $e^{-U_{ext}/kT}$ is the contribution from intermolecular interactions in the crystal phase, where U_{ext} is the intermolecular interaction energy of a molecule at its minimum energy state in the crystal. Both electrostatic and van der Waals (VDW) nonbonded interactions contribute to U_{ext}, which by definition is zero in the gas phase. P is the external pressure, and $\bar{V}_x$ is the molar volume in the crystal. At 1 atm, the PV term is negligible and will be ignored hereafter. The expression for the chemical potential in the crystal, in contrast to the solution and gas phase expressions, contains no concentration term, because each molecule is fixed at a distinguishable lattice site. For exactly this reason also, the gas and solution phase translational and rotational terms are replaced by a lattice vibrational term in the crystal.

Gas-to-solution transfer

Equating the gas and solution chemical potentials yields

$$kT \ln\left(\frac{\rho_g}{\rho_s}\right) = \Delta G^{g\to s}, \quad (5)$$

where the experimentally measurable quantities are collected on the left-hand side. The concentration-independent part of the translation partition function, Λ, has dropped out, leaving only concentration-dependent translation terms on the left-hand side. To calculate the net transfer free energy (step 1 of Fig. 1), $\Delta G^{g\to s}$, the transfer of a solute from gas-to-water, is treated as a three-step process: 2) discharge of the solute in the vapor phase; 3) transfer of the resulting nonpolar solute to water; 4) subsequent recharging of the solute in the presence of solvent. The sum of the steps 2 and 4 gives the electrostatic contribution, and step 3 gives the nonpolar, or hydrophobic term, which is computed separately, as described in Methods below. In light of the ring structure of the cyclic dipeptides, we expect little change in molecular geometry upon solvation and intramolecular interactions to contribute negligibly to the solvation free energy. That is, we assume that q^{int} is identical in the gas and solution phases. Thus, solute-solvent electrostatic and hydrophobics interactions dominate the thermodynamics of solvation, and we write

$$\Delta G^{g\to s} \equiv \Delta G^{g\to s}_{elec} + \Delta G^{g\to s}_{hydrophobic}. \quad (6)$$

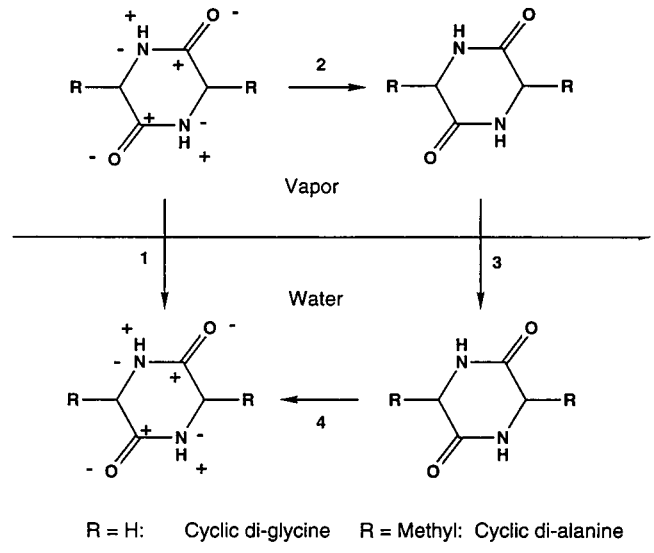

FIGURE 1 Schematic of the structure of the cyclic dipeptides and illustration of the three-step cycle used to calculate solvation free energies. Step 1: Transfer of molecule from gas to water. Step 2: Discharge of molecule in gas phase. Step 3: Transfer of discharged (apolar) molecule into water. Step 4: Recharging of molecule in water.

Gas-to-crystal transfer

Equating the gas and crystal chemical potentials gives

$$\ln \rho_g = \ln\left(\frac{q^{rot} q_g^{vib} q_g^{int} e^{U_{ext}/kT}}{\Lambda^3 q^{lat} q_x^{vib} q_x^{int}}\right), \quad (7)$$

where again, the experimentally measurable quantity appears on the left-hand side. In this case, both concentration-dependent and -independent parts of the translation partition function remain. The term on the right-hand side represents the concentration-independent contribution to the free energy of transferring the molecule from the gas phase to the crystal. The net transfer free energy depends on the molecule's concentration in the gas phase. The transfer free energy must therefore be specified with respect to some reference concentration, ρ_{ref}, as

$$\Delta G^{g\to x} = kT \ln\left(\frac{q^{rot} q_g^{vib} q_g^{int} e^{U_{ext}/kT}}{\rho_{ref} \Lambda^3 q^{lat} q_x^{vib} q_x^{int}}\right) = kT \ln\left(\frac{\rho_g}{\rho_{ref}}\right). \quad (8)$$

Crystal-to-solution transfer

Thermodynamic expressions for the crystal-to-solution transfer are obtained by subtracting the gas-to-crystal con-

tributions (Eq. 8) from the gas-to-solution contributions (Eq. 5), which yields

$$\ln \rho_s = \ln\left(\frac{q^{rot} q_g^{vib} q_g^{int} e^{U_{ext}/kT}}{\Lambda^3 q^{lat} q_x^{vib} q_x^{int}}\right) - \Delta G^{g\to s}/kT. \quad (9)$$

The corresponding crystal-to-water transfer free energy is defined as

$$\Delta G^{x\to s} = \Delta G^{g\to s} - \Delta G^{g\to x} = kT \ln\left(\frac{\rho_{ref}}{\rho_s}\right) \quad (10)$$

with respect to the reference concentration ρ_{ref}.

Free energy, enthalpy, and entropy contributions to transfer from gas to crystal

Equations 5, 7, and 9 provide expressions for calculating the experimentally measurable quantities (i.e. concentrations) in terms of internal and external interactions in the gas, solution, and crystal phases. Expressions for the different free energy contributions are obtained via the usual statistical mechanical expressions (McQuarrie, 1976). These are summarized below, where, following the notation of Tidor and Karplus (1994), the enthalpic and entropic contributions are indicated by the expressions enclosed in square brackets [ ] and curly braces { }, respectively.

For the internal potential energy contribution to the free energy, we have

$$kT \ln\left(\frac{q_g^{int}}{q_x^{int}}\right) = [\Delta U_{int}^{g\to x}], \quad (11)$$

where $\Delta U_{int}^{g\to x}$ is the difference in mean intramolecular potential energy in the gas and crystal phases. Similarly, the external, or intermolecular, contribution is given by

$$[U_{ext}], \quad (12)$$

which includes intermolecular VDW and electrostatic interactions in the crystal. Neither potential energy contribution has an entropic component.

The translational contribution to free energy upon transfer from a gas at the reference concentration ρ_{ref} to the crystal phase is

$$\Delta G^{trans} = \left[\frac{5kT}{2}\right] - \left\{kT \ln\left(\frac{e^{5/2}}{\rho_{ref}\Lambda^3}\right)\right\}. \quad (13)$$

The volume change upon vapor-to-crystal transfer contributes a term kT to the enthalpy and is included here in thetranslational term. The translational partition function depends on concentration, so the free energy and entropy are specified with respect to the reference concentration ρ_{ref}.

The rotational partition function contribution to the free energy is given by

$$\Delta G^{rot} = -kT \ln q^{rot} \quad (14)$$
$$= [3kT/2] - \left\{kT \ln\left(\left(\frac{2\pi kTe}{h^2}\right)^{3/2} \frac{8\pi^2 (I_x I_y I_z)^{1/2}}{\sigma}\right)\right\},$$

where I_x, I_y, and I_z are the principal moments of inertia, and σ is the molecular symmetry number (2 for the cyclic dipeptides).

The presence of intermolecular forces in the crystal phase potentially alters the intramolecular modes of vibration of the peptides relative to the gas phase. For instance, internal torsional freedom may be sterically restricted upon binding. This contribution to the partition function comes from $3n - 6$ intramolecular "normal" modes of vibration, where n is the number of atoms (14 for cGG, 20 for cAA). In the harmonic oscillator approximation, the free energy can be written (Tidor and Karplus, 1994)

$$\Delta G^{vib} = -kT \ln\left(\frac{q_x^{vib}}{q_g^{vib}}\right)$$
$$= \left[\sum_{i=1}^{3N-6}\left(\frac{h\nu_{i,x}}{2}\coth\left(\frac{h\nu_{i,x}}{2kT}\right) - \frac{h\nu_{i,g}}{2}\coth\left(\frac{h\nu_{i,g}}{2kT}\right)\right)\right]$$
$$- \left\{\sum_{i=1}^{3N-6}\frac{h\nu_{i,x}}{2}\coth\left(\frac{h\nu_{i,x}}{2kT}\right)\right. \quad (15)$$
$$- kT \ln\left(2\sinh\left(\frac{h\nu_{i,x}}{2kT}\right)\right) - \frac{h\nu_{i,g}}{2}\coth\left(\frac{h\nu_{i,g}}{2kT}\right)$$
$$\left. + kT \ln\left(2\sinh\left(\frac{h\nu_{i,g}}{2kT}\right)\right)\right\},$$

where $\nu_{i,g}$ and $\nu_{i,x}$ are the frequencies of the ith internal vibrational modes in the gas and crystal phases, respectively.

In analyzing the transfer of a molecule from gas to crystal, one must account for whole-body (center of mass translational and rigid-body rotational) motions of the molecule within the crystal. The simplest way to model these motions, the Einstein model, is to regard neighboring molecules as establishing a mean harmonic potential in which the molecule vibrates and librates within its unit cell (Hill, 1986), giving a lattice vibration contribution

$$G^{lat} = -kT \ln q^{lat} = \left[\sum_{i=1}^{6}\frac{h\nu_i}{2}\coth\left(\frac{h\nu_i}{2kT}\right)\right] \quad (16)$$
$$- \left\{\sum_{i=1}^{6}\frac{h\nu_i}{2}\coth\left(\frac{h\nu_i}{2kT}\right) - kT \ln\left(2\sinh\left(\frac{h\nu_i}{2kT}\right)\right)\right\},$$

where υ_i is the frequency of the ith lattice vibration.

The net free energy, enthalpy, and entropy for the gas-to-crystal transfer are given by

$$\Delta G^{g\to x} = \Delta U_{int}^{g\to x} + U_{ext} + \Delta G^{vib} + \Delta G^{lat} - \Delta G^{trans} - \Delta G^{rot} \quad (17)$$

$$\Delta H^{g\to x} = \Delta U_{int}^{g\to x} + U_{ext} + \Delta H^{vib} + \Delta H^{lat} - \Delta H^{trans} - \Delta H^{rot} \quad (18)$$

$$\Delta S^{g\to x} = \Delta S^{vib} + \Delta S^{lat} - \Delta S^{trans} - \Delta S^{rot}, \quad (19)$$

respectively. In the high temperature limit, each vibrational degree of freedom contributes kT to the enthalpy, so the internal vibrations make no net contribution and the lattice terms give $6kT$, and the loss of gas phase translational and rotational freedom contributes $-4kT$. In this limit,

$$\Delta H^{g\to x} \approx U_{int}^{g\to x} + U_{ext} + 2kT \quad (20)$$

$$\text{for } \frac{kT}{h} > \nu_i, \quad \text{for all i.}$$

The high temperature limit is useful for estimating the expected enthalpy change (Hagler et al., 1974), but in the work described here the exact expression (Eq. 18) was used throughout.

One contribution of particular interest is the cost of immobilizing a molecule in the crystal phase, which is completely analogous to the relative immobilization of one molecule with respect to another when forming a binary complex in solution. The last three terms in each of Eqs. 17–19 represent this immobilization contribution, referred to here as the "association" term. The major contribution to this term is the entropically unfavorable conversion of three translational and three rotational vapor-phase degrees of freedom into six new vibrational degrees of freedom in the crystal. In the high temperature limit, approximately $2kT$ of the association term is enthalpic, as described by Eq. 20, so association entropy makes the predominant association contribution to crystallization and poses the extensively discussed entropic barrier to binding.

Heat capacity contributions to transfer from gas to crystal

Heat capacity contributions arise from the translational, rotational, and vibrational terms, and from the temperature dependence of the enthalpic contributions in Eqs. 13–16. These are given by $5k/2$ and $3k/2$ for translation and rotation, respectively, and

$$C_p^{vib} = \frac{(h\nu_i/2kT)^2}{\sinh^2(h\nu_i/2kT)} \quad (21)$$

for each internal or lattice vibrational frequency, ν_i.

METHODS

Extraction of transfer thermodynamic data from experiments

The gas-to-water equilibrium for both cGG and cAA has been studied experimentally by Murphy and Gill (1990a,b). The enthalpies of dissolution were taken directly from their data as 6.27 kcal/mol and 3.28 kcal/mol for cGG and cAA, respectively, at 298K. Their solubility data also give the concentration of peptide in water that is in equilibrium with the crystal phase at 298K as ρ_s = 0.145 M and ρ_s = 0.187 M for cGG and cAA, respectively. In their previous analysis of this data, Murphy and Gill also derived the dissolution free energy and entropy using a mole fraction standard state (Murphy and Gill, 1990a,b). Use of the mole fraction scale, however, is unnecessary (Holtzer, 1995), and as Eqs. 9–10 make clear, the thermodynamic quantity required for this analysis is the number density (i.e., molar concentration). With crystal concentrations of cGG and cAA of 14 M and 9 M, respectively, the solubility data give experimental crystal-to-water transfer free energies of 2.7 and 2.3 kcal/mol for cGG and cAA, respectively, and corresponding entropies of 12.0 and 3.3 cal/mol/K, respectively.

The gas-to-crystal equilibrium for cGG has been studied by Seki et al. (1956). The vapor pressure of the crystalline dipeptide as function of temperature over the range 414–449K was taken directly from this study. Unfortunately, corresponding data for cAA appear to be unavailable. $\Delta H^{g\to x}$ is obtained from the vapor pressure versus temperature data via the Clausius-Clapeyron equation

$$\frac{\Delta H^{g\to x}}{k} = \frac{\partial \ln P_g}{\partial(1/T)}, \quad (22)$$

where P_g is the dipeptide vapor pressure. A least-squares fit of the data gives an enthalpy of −24.9 kcal/mol from the slope, with an intercept of 9.65 (when pressure is expressed in units of mmHg), which is indistinguishable from the previously published value. From this fit the vapor pressure at 298K is 2.83×10^{-9} mmHg, which gives the concentration of vapor in equilibrium with both the crystal and saturated solution at 298K as $\rho_g = 1.66 \times 10^{-13}$M. The ratio of gas and saturated solution concentrations provides the experimental estimate for $\Delta G^{g\to s}$, using Eq. 5, of −16.3 kcal/mol. Because, however, the gas/crystal equilibrium experiments must be performed at a rather high temperature (>400K) to produce a measurable vapor pressure, a long extrapolation of the vapor pressure to 298K is required. Because there will be some heat capacity change from gas to crystal, some curvature of the plot of ln P versus $1/T$ is expected. In this case, the experimental enthalpy of gas-to-crystal transfer is given by

$$\Delta H^{g\to x} = \Delta H_o^{g\to x} + \Delta C_p^{g\to x}(T - T_o), \quad (23)$$

where $\Delta H_o^{g\to x}$ is the enthalpy at some reference temperature T_o, and $\Delta C_p^{g\to x}$ is the gas-to-crystal heat capacity change, which is assumed to be temperature-invariant over this range. Substituting Eq. 23 into Eq. 22 and integrating from T_o to T gives an expression for the vapor pressure at T:

$$k \ln P_g(T) = k \ln P_g(T_o) + (\Delta H_o^{g\to x} - \Delta C_p^{g\to x} T_o)\left(\frac{1}{T} - \frac{1}{T_o}\right) + \Delta C_p^{g\to x} \ln\left(\frac{T_o}{T}\right), \quad (24)$$

where $P_g(T_o)$ is the vapor pressure at the reference temperature. If the high temperature approximation for the heat capacity change from gas to crystal is used (i.e., using the temperature derivative of Eq. 20), then $\Delta C_p^{g\to x} = 2k$. Taking T_o to be 430K, the middle of the experimental temperature range, then ln $P_g(T_o)$ = −6.71, $\Delta H_o^{g\to x}$ = −24.9 kcal/mol. Equation 24 yields −25.4 kcal/mol and 2.29×10^{-9} mmHg for the enthalpy and vapor pressure at 298K, respectively. Using this corrected vapor pressure value, $\Delta G^{g\to s}$ = −16.4 kcal/mol. The high temperature heat capacity limit has been used to analyze the experimental enthalpy and vapor pressure data,

neglecting the possible contribution of intramolecular vibrations. In fact, the more detailed calculations will themselves provide a better estimate of the heat capacity change (see Results), which can then be used to analyze the experimental enthalpy and vapor pressure value in a self-consistent manner. The size of the correction seen in the high temperature limit already shows, however, that the additional correction to enthalpy and free energy values will be on the order of tenths of a kilocalorie per mole or less.

The crystal structures of cGG and cAA were obtained from the Cambridge small molecule database. cGG crystallizes in the $P21_c$ space group with unit cell parameters of a = 5.23 Å, b = 11.55 Å, c = 3.98 Å, and α = 90°, β = 97.98°, γ = 90°, with two molecules in a cell of volume 238 Å^3 (Degeilh and Marsh, 1959). cAA crystallizes in the P1 space group with unit cell parameters of a = 5.16 Å, b = 8.06 Å, c = 4.67 Å, and α = 103.16°, β = 103.68°, γ = 97.59°, with one molecule in a cell of volume 180 Å^3 (Sletten, 1970).

Calculation of gas-to-water transfer thermodynamics

The free energy of transferring a molecule from gas to aqueous solution, $\Delta G^{g\to s}$, is divided into a hydrophobic and an electrostatic component (Eq. 6) by means of the thermodynamic cycle shown in Fig. 1. The electrostatic term is computed using the continuum finite-difference Poisson-Boltzmann (FDPB) method and PARSE atomic radii/charge parameters combined with a surface area treatment of the nonelectrostatic term as implemented by Sitkoff et al. (1994). All calculations were performed using crystal structure coordinates at zero ionic strength. To account for electronic polarizability, the dipeptide solutes were assigned a dielectric constant of 2 (Sharp et al., 1992). The solvent was assigned a dielectric constant of 1 (vacuum) for the discharge step and 80 (water) for the charging step. Only solvent-solute reaction-field energies contribute to $\Delta G^{g\to s}_{elec}$, because the intramolecular solute-solute Coulomb energy does not change upon solvation. FDPB calculations were run with the same grid size and scaling parameters as described previously (Sitkoff et al., 1994). The hydrophobic contribution to dipeptide solvation was calculated using

$$\Delta G^{g\to s}_{hydrophobic} = \gamma A + b, \qquad (25)$$

where A is the solvent-accessible surface area of the solute, γ = 5.0 ± 0.5 cal/mol/Å^2, and b = 0.86 ± 0.10 kcal/mol (Sitkoff et al., 1994). Accessible surface areas of the dipeptides were computed from the crystal structure using the vertex algorithm (Sridharan et al., 1992), with a 1.4-Å solvent probe radius.

Gas-to-crystal transfer

Translational and rotational contributions were calculated directly from Eqs. 13 and 14, using the known dipeptide mass and the moments of inertia computed from the crystal structure.

The internal energy contribution ΔG^{int} denotes the change in covalent, intramolecular electrostatic, and intramolecular (VDW) internal energies, or the strain energy introduced upon transfer of a molecule from the gas to the crystal lattice. The molecular mechanics package Discover (Biosym Corporation) was used with either the CVFF (Hagler et al., 1974) or AMBER (Weiner et al., 1986) empirical energy functions to compute this energy. Partial charges in these potential functions were parameterized using a dielectric constant of 1. Consequently, a dielectric of 1 was used for both gas and crystal phase minimizations. The peptide was minimized in both the gas and crystal phases, and the difference in internal energy was then obtained directly from the final minimized structures. Minimization was performed until the maximum gradient was less than 0.001 kcal/mol/Å. No nonbonded interaction cutoff was used in the gas phase. The crystal-phase minimizations were performed, including the crystal symmetry with periodic boundary conditions in the explicit image convention. To check that the nonbonded cutoff was sufficiently large, the crystal-phase energy was evaluated by using cutoffs lengths from 10 to 30 Å and then comparing these results to an infinite-cutoff Ewald summation calculation. Discover currently does not support Ewald summation evaluations for normal mode analysis. However, ensuring that our cutoff was sufficiently large for the minimization analysis before proceeding enabled us to use direct truncation in the normal mode analysis with little error. Comparison of the truncation and Ewald summation methods shows that the nonbonded interactions converge at cutoffs of 20 Å or greater, so a cutoff of at least this value was used in all calculations. The crystal-phase energy was also minimized with respect to unit cell lengths to account for the possibility of mismatch between force-field parameters (particularly the VDW radii) and the experimental unit cell parameters. This minimization lowered the crystal-phase energy by about an additional 2 kcal/mol while preserving the unit cell volume within 5%.

Alterations in the internal vibrational modes of the peptide by intermolecular crystal forces were calculated within the harmonic approximation by using Eq. 15, which requires the internal vibrational frequencies. The frequencies were obtained using the Normal Mode Analysis (NMA) facility within Discover, using both the CVFF and AMBER force fields. The crystal-phase NMA was performed using explicit periodic images with a nonbond cutoff of 20 Å for both diglycine and dialanine.

The contribution of lattice vibrations was obtained within the Einstein model, by using Eq. 16. Two methods were used to determine the Einstein frequencies, ν_i. In the first, Discover was used to perform NMA with a 20-Å nonbond cutoff and periodic boundary conditions in the explicit image convention. For the purpose of these calculations, the six lowest frequencies were assigned to lattice motions, with the remaining $3n - 6$ frequencies assigned to intramolecular vibrations. Although the assignment of modes does not affect the total calculated entropy, it should be noted that because of extensive coupling of low frequency modes, the six lowest frequency motions will involve more than pure center-of-mass motions. The lattice frequencies were also estimated using a form of Hooke's law analysis by explicitly probing the forces opposing the six orthogonal motions: a small translational or rotational perturbation was applied to a single molecule relative to the surrounding lattice and the resulting energy change was plotted versus the square of the perturbation. The force constant is twice the slope of this plot. From the force constants, the total molecular mass (M_{total}) and moments of inertia (I_i), the lattice vibrational thermodynamics were calculated by using Eq. 16. The CVFF forcefield was used in all lattice frequency calculations. These vibrational frequencies were also used to calculate contributions to the heat capacity by using Eq. 21.

In obtaining a free energy or entropy for transfer into the crystal phase, either from experimental data or from calculations, the reference concentration (ρ_{ref}) in the originating phase (gas or aqueous solution) must be specified, because the translational contribution, unlike the other contributions, depends on concentration. The usual choice of reference concentration in biological applications is 1 M, because concentrations are usually expressed in molarity units. However, two other natural concentration units suggest themselves. The first, although somewhat unfamiliar, is units of $1/\Lambda^3$, the "volume" defined by the thermal de Broglie wavelength. Physically, the de Broglie volume is the smallest volume in which one can confine the center of mass of the molecule, because of the Heisenberg uncertainty principle. When the crystal phase is involved, the second concentration unit that suggests itself is $1/V_x$, defined by the molar volume in the crystal. Results are presented using the 1 M reference state because of its common usage. We also present results obtained by using the crystal molar volume reference state for gas and solution transfer, because this effectively defines a transfer free energy or entropy from the crystal phase in which any purely dilutional or gas expansion contribution is removed.

Numerical uncertainties

Uncertainties can arise from empirical parameters, the physical model itself, and uncertainty in the structure. Uncertainties were estimated for the calculated solvation free energy and for key gas-to-crystal transfer quantities (Tables 1 and 2). The contribution of structural uncertainty to the solvation free energy was estimated by running molecular dynamics on the

TABLE 1 Dipeptide solvation free energies*

Contribution	cGG	cAA
Electrostatic	−18.5 ± 1.0	−17.8 ± 0.6
Hydrophobic	2.1 ± 0.3	2.4 ± 0.3
$\Delta G^{g\to s}$ calculated	−16.4 ± 1.3	−15.4 ± 0.9
$\Delta G^{g\to s}$ experimental	−16.4	NA#

*All free energies in kcal/mol, T = 298K. Discover CVFF-minimized gas phase structures used.
#Not available because of absence of sublimation data for cAA.

dipeptide in the gas phase at 300K, taking 10 randomly selected snapshots, and recomputing the solvation free energy. The standard deviation of these calculations is shown in Table 1. Uncertainties in gas-to-crystal thermodynamic quantities are likewise a sum of individual terms. The translational term has no error and the rotational term essentially none, because the moments of inertia were found to vary negligibly during dynamics. Thus, gas-to-crystal uncertainties are due principally to uncertainties in the potential energy minimum of the system, intramolecular vibrations, and lattice vibrations. We attribute the major uncertainty in energy minima (Tables 2 and 3) to the error introduced by cutting off nonbond interactions at 20 Å. The crystal-phase minima adequately stabilize at 20 Å cutoff, as evidenced by the observation that the minimum energy drops by less than 0.1 kcal/mol upon increasing the cutoff to 25 Å and, in cAA, by another 0.1 kcal/mol upon "infinite-cutoff" Ewald summation. We consider investigation of the sensitivity of our energy minima with respect to force-field parameters to be beyond the scope of this work. However, internal energy calculations were performed with both the CVFF and AMBER parameters. The intramolecular vibrational uncertainties were estimated from the differences in the CVFF and Amber gas-phase NMA results for cAA and cGG. The uncertainty in our determination of lattice vibrational thermodynamics is considered later in the Results and Discussion sections. We combine our various error estimates to obtain uncertainties in the net gas-to-crystal thermodynamic energies (Table 2).

RESULTS

The results for diglycine and dialanine solvation thermodynamics are shown in Table 1. The net solvation free energy for cGG (−16.4 kcal/mol) has a small, unfavorable contribution (2.2 kcal/mol) due to the ordering of water around the nonpolar cavity, plus a large, favorable contribution (−18.6 kcal/mol) arising via solute-water electrostatics. cAA has a larger cavity term due to the methyl side chains. This also results in a smaller electrostatic solvation interaction, because the solvent can approach the polar peptide group less closely.

The thermodynamics of gas-to-crystal transfer for cGG and cAA using the CVFF force field are shown in Table 2. Results for cGG and cAA confirm the expectation that entropy opposes crystallization, whereas enthalpy favors it. The bulk of the driving force for cGG crystallization comes from the introduction of intermolecular contacts in the crystal phase, which is reflected in the −27.0 kcal/mol change in CVFF energy minimum upon crystallization. There are additional small enthalpy changes due to translational, rotational, and zero-point vibrational effects, and a 1 kT PV contribution from the large molar volume difference in the gas and crystal states. The net enthalpy change of −24.6 kcal/mol compares favorably with the experimental value of −24.8 kcal/mol. The relative contributions in cAA are similar, although the intermolecular contribution is less favorable, resulting in a smaller net enthalpy decrease upon crystallization. The bulk of the entropy change for both dipeptides arises from loss of translational and rotational freedom, with concomitant partial entropy recovery in the form of lattice vibrations. A small negative entropy contribution comes from the restriction of internal motion. The net entropy change for cGG is significantly larger than that seen experimentally. Results for cAA show similar entropy contributions, although the lattice entropy is somewhat greater.

Table 3 provides a more detailed look at the enthalpic contributions to gas-to-crystal transfer for both cGG and cAA. In general, the change in intramolecular energy upon crystallization is small (<2 kcal/mol) and unfavorable, whereas the creation of new, favorable VDW contacts (−18 kcal/mol) and charge pairings (−8 kcal/mol) provides significant stabilization. To analyze the stability provided by the network of backbone hydrogen-bonding groups that form in the crystal, the crystal-phase intermolecular energy was further divided into VDW, hydrogen-bonding, and long-range electrostatic contributions. The hydrogen bond contribution is given for the CVFF force field by the pure dipole-dipole electrostatic interactions between backbone carbonyl and amide groups of neighboring molecules. The rest of the intermolecular electrostatic energy (i.e., between dipolar groups on molecules not directly H-bonded to each other) is termed "long-range" electrostatics. Hydrogen bonding contributes about −6 kcal/mol to $\Delta G^{g\to x}$ and accounts for most (~70%) of the intermolecular electrostatic stabilization in both crystals. The long-range contribution is nevertheless significant.

Comparison of the cAA results to those of cGG quantifies the energetic effects of replacing the hydrogen side chains of cGG by larger, nonpolar methyl groups. The CVFF results in Table 3 indicate that cAA is about 8 kcal/mol less enthalpically stable in the crystal than cGG. The change in intramolecular energy upon cAA crystallization is larger than that of cGG by 1.4 kcal/mol, reflecting steric strain induced in the crystal due to the bulkier cAA side chains. This effect is geometrically apparent, as the cAA ring is noticeably puckered in its CVFF crystal-minimized configuration and planar in the gas phase. The cGG ring, on the other hand, is planar in both phases. The total intermolecular energy of cAA is identical to that of cGG (−27.2 kcal/mol), although the division of energy contributions comprising this total is markedly different. Backbone hydrogen bonds contribute equally (−6 kcal/mol) in the two crystals; however, intermolecular VDW contacts provide cAA ~1 kcal/mol more binding energy than cGG, whereas cGG has ~1 kcal/mol more favorable long-range electrostatic energy. The methyl side chains in cAA provide additional intermolecular VDW contacts, which account for the additional 1 kcal/mol of VDW binding energy in cAA. The additional long-range electrostatic stabilization in the cGG crystal can be understood by detailed inspection of the crystal packing (Fig. 2), which reveals that neighboring

TABLE 2 Cyclic dipeptide gas-to-crystal transfer thermodynamics*

Contribution	cGG			cAA		
	ΔH	$T\Delta S$	ΔG	ΔH	$T\Delta S$	ΔG
Translation#	−1.5	−8.5	7.0	−1.5	−8.9	7.4
Rotation	−0.9 ± 0.0	−7.9 ± 0.1	7.1 ± 0.1	−0.9 ± 0.0	−8.4 ± 0.1	7.5 ± 0.1
Intramol. vibration	1.3 ± 0.4	−1.6 ± 0.5	2.9 ± 0.7	0.8 ± 0.4	−1.5 ± 0.5	2.3 ± 0.7
Lattice vibration§	3.6	7.7	−4.1	3.6	9.7	−6.1
U	−27.0 ± 0.2	0	−27.0 ± 0.2	−23.9 ± 0.2	0	−23.9 ± 0.2
Net calculated	−24.6 ± 0.6	−10.3 ± 0.6	−14.1 ± 1.0	−21.9 ± 0.6	−9.1 ± 0.6	−12.8 ± 1.0
Experimental	−24.8	−5.8	−19.0	NA	NA	NA

*All energies in kcal/mol, T = 298K. All structures are Discover CVFF-minimized to a maximum derivative of 0.001 kcal/mol/Å. Crystal structures were also optimized with respect to unit cell lengths using a 20-Å nonbond cutoff for both cGG and cAA.
#Crystal concentration used as reference state.
§Calculated using the six lowest frequencies from Discover normal mode vibrational analysis.

highly charged backbone groups are more densely packed in the cGG crystal than in cAA. Given that the net long-range interaction is favorable, this denser packing provides cGG 1.2 kcal/mol more stabilization than cAA. A plot of the non-H-bonding electrostatic interactions between one dipeptide molecule and its neighboring CO and NH dipolar groups as a function of distance is given in Fig. 3. The curves show that there are significant contributions out to 20–30 Å. This plot also emphasizes the necessity for using long cutoffs or Ewald summations in the crystal-phase calculations. Summarizing these results: if the packing contribution is defined as the sum of strain and VDW interactions, there is little difference enthalpically between cAA and cGG in packing and hydrogen bonding. The major difference arises from the long-range electrostatics.

Table 3 also shows the energy minimization results obtained by using the Amber force field, which illustrates the sensitivity of the results to energy parameterization. The gas-to-crystal transfer enthalpies computed using Amber are 3.1 and 4.7 kcal/mol larger than CVFF for cGG and cAA, respectively. Combining the intermolecular term with the vibrational and translation/rotation enthalpy contributions (ignoring the small zero-point energy change), Amber yields a gas-to-crystal transfer enthalpy of −22.7 kcal/mol, which compares more poorly with the experimental value (−24.8 kcal/mol) than the CVFF result (−24.6 kcal/mol). With Amber, backbone hydrogen bonds apparently account for 3.1–3.7 kcal/mol more binding energy than with CVFF, whereas the VDW contribution shows no difference between cGG and cAA. The CVFF force field has a zero radius for the amide H; hence there is no intermolecular H···O VDW energy and the H bond is purely electrostatic. Amber assigns the amide hydrogen a radius of about 1 Å, and thus the H-bond contribution contains the electrostatic term plus an explicit H-bond (12–10) interaction between the amide hydrogen and carbonyl oxygen. Relative to the CVFF crystal, the Amber-minimized crystal is stabilized about 3.5 kcal/mol more by hydrogen bonding and 6.5 kcal/mol less by intermolecular VDW contacts (Table 3). This difference reflects the different forms of the potential functions. Although the force fields disagree by more than 1 kcal/mol on the long-range electrostatic contributions,

TABLE 3 Cyclic dipeptide energies in gas and crystal phases using CVFF and Amber force fields*

Phase	Contribution	CVFF		Amber	
		cGG	cAA	cGG	cAA
Gas	Internal (U_g^{int})	29.4	36.0	70.5	80.9
Crystal§	Internal (U_X^{int})	29.6	37.6	70.7	82.1
	van der Waals	−18.5	−19.6	−12.8	−12.8
	H-bond	−6.0	−6.1	−9.7	−9.2
	Longe-range electrostatics#	−2.7	−1.5	−1.6	−0.1
	Intermolecular total (U^{ext})	−27.2	−27.2	−24.1	−22.1
	Total	2.4	10.4	46.6	60.0
Crystal-gas	Packing ($\Delta U^{int} + U^{vdw}$)	−18.3	−18.0	−12.6	−11.6
	H bonding	−6.0	−6.1	−9.7	−9.2
	Longe-range electrostatics	−2.7	−1.5	−1.6	−0.1
	Total ($\Delta U^{int} + U^{ext}$)	−27.0 ± 0.2	−25.6 ± 0.2	−23.9	−20.9

*All data are in kcal/mol at 25°C. Conjugate gradient minimization was performed using Discover to a maximum derivative of 0.001 kcal/mol/Å.
#Total intermolecular electrostatic interaction excluding backbone hydrogen-bonding electrostatics.
§Optimization with respect to unit cell lengths performed in each case. Nonbond cutoff of 20 Å for cGG and cAA.

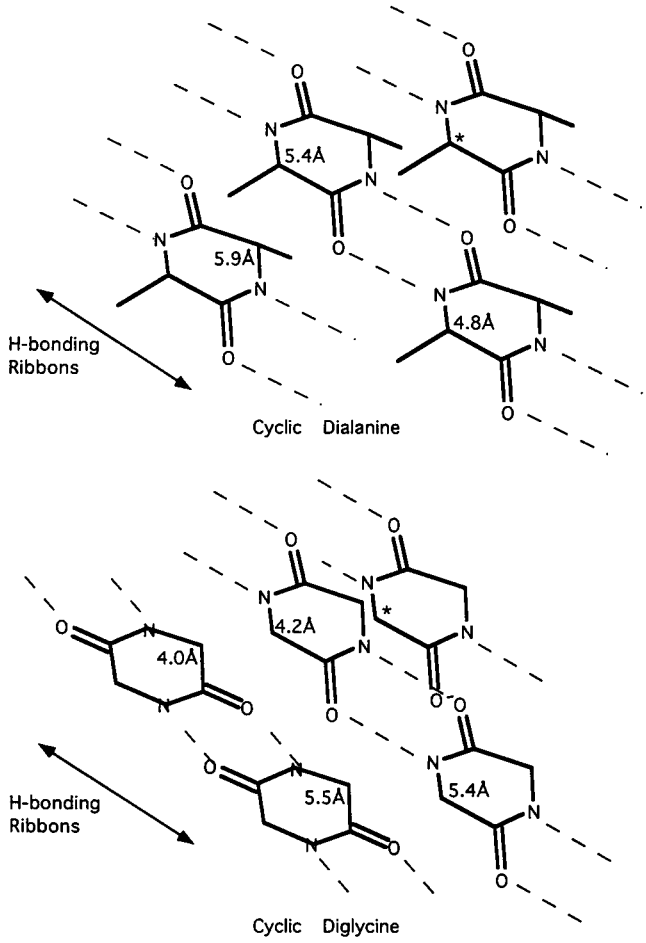

FIGURE 2 Illustration of packing in cGG and cAA crystals. Hydrogen bonding ribbons are illustrated by dotted lines. Distances are indicated from the asterisked C_α to the nearest neighboring C_α atoms.

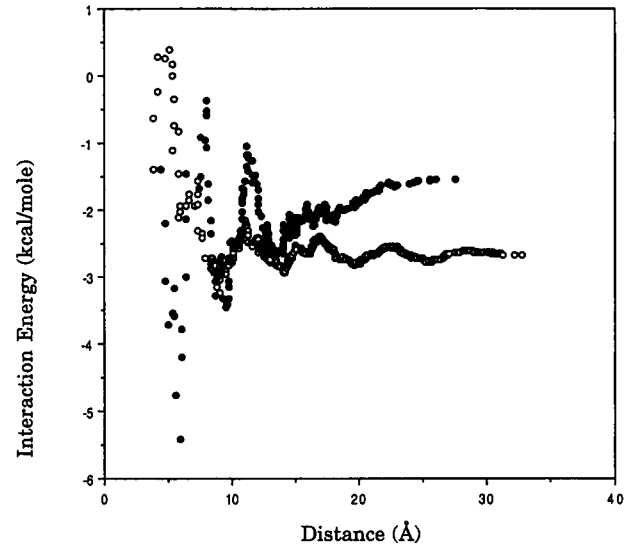

FIGURE 3 Plot of the cumulative electrostatic interaction (excluding the direct hydrogen bonding interaction) as a function of distance for cGG (○) and cAA (●).

they both indicate that cGG has over 1 kcal/mol more long-range electrostatic binding energy than cAA. Backbone hydrogen bond angles differ between force fields by less than 2° for both compounds and are thus energetically negligible.

The association entropy, defined as $\Delta S_{assoc} = \Delta S_{lat} + \Delta S_{trans} + \Delta S_{rot}$, poses a −29.2 eu (8.7 kcal/mol at 25°C) barrier to cGG crystallization. cAA has a similar, although smaller association entropy loss of −25.5 eu. Table 4 provides a breakdown of association entropy contributions for cGG and cAA, and Table 5 shows some of the physical parameters used in the association entropy calculations, including the internal vibrational and lattice frequencies. Using the crystal-phase number density as the translational standard state, which removes any purely expansional contribution to the crystal-to-gas transfer process, the loss of three translational degrees of freedom (DOF) poses an entropy barrier of about 28 kcal/mol for cGG, as does the loss of three rotational degrees of freedom. Using Discover normal mode analysis, it is found that roughly half of this entropy is reclaimed via the addition of six lattice vibrations in the crystal. Because there is considerably more enthalpy (about 2–3 kcal/mol) associated with the lattice vibrations than with either the translational or rotational degrees of freedom, only 4 kcal/mol of free energy is regained via lattice vibrations. Consequently, our calculations indicate that the association terms pose a large 10 kcal/mol free energy barrier to crystallization. The corresponding calculated association entropy is −29.2 eu. The translational and rotational entropy terms are exact, and it is expected that NMA calculations of the intramolecular vibrational thermo-

TABLE 4 Gas-to-crystal association entropies*

Contribution	cGG	cAA
Translation ($C_{ref} = C_X$)#	−28.5	−30.0
Translation (C_{ref} = 1M)	−23.3	−25.6
Rotation	−26.6	−28.2
Intramolecular vibrations	−5.4	−5.1
Lattice vibrations (Hooke)§	17.9	19.6
Lattice vibrations (NMA)¶	25.9	32.7
Lattice vibrations (experimental)‖	41.0	NA
Association entropy (Hooke)**	−37.2	−38.6
Association entropy (NMA)**	−29.2	−25.5
Association entropy (experimental)**	−14.1	NA

*All entropies in units of eu = cal/mol/K.
#Crystal concentrations: cGG = 14.0 M; cAA = 9.2 M.
§Entropies computed using frequencies determined via explicit Hooke's law model (see Methods and Table 5).
¶Entropies computed using the six lowest crystal-phase normal mode frequencies from Discover (Table 5).
‖cGG value computed from the experimental entropy of sublimation minus rotational, translational, and internal vibrational entropy, assuming that the NMA value for the latter (−5.4 eu for cGG) is exact.
**Sum of translational, rotational, and lattice vibrational contributions, using the crystal concentration translational reference state and NMA, Hookean or "experimental" lattice vibrational contributions, respectively. The cAA value was not available because absence of sublimation data.

TABLE 5 Parameters used for entropy calculations*

Contribution	Quantity	cGG		cAA	
Translation	Thermal wavelength (Å)	0.0947		0.0849	
Rotation	Principal moments of inertia# (amu Å^2)	428		596	
		328		423	
		107		276	
Intramolecular vibration	Lowest frequencies§ (cm^{-1})	Gas	Xtal	Gas	Xtal
		662	640	667	658
		598	605	629	614
		545	523	538	559
		491	520	493	535
		482	514	430	501
		391	492	417	458
		390	442	384	428
		368	317	359	406
		191	315	357	363
		185	169	289	359
		59		287	322
				267	310
				256	282
				137	184
				128	166
				38	117
Lattice vibration	NMA frequencies¶ (cm^{-1})	163		107	
	NMA frequencies¶ (cm^{-1})	150		91	
	NMA frequencies¶ (cm^{-1})	113		75	
	NMA frequencies¶ (cm^{-1})	40		21	
	NMA frequencies¶ (cm^{-1})	30		17	
	NMA frequencies¶ (cm^{-1})	23		9	
Lattice vibration	Hooke's law rotational‖	183		186	
	Hooke's law rotational‖	124		78	
	Hooke's law rotational‖	164		158	
	Hooke's law translational**	105		106	
	Hooke's law translational**	106		69	
	Hooke's law translational**	105		108	

*All structures were minimized with Discover in the CVFF force field. Crystal structures were also optimized with respect to unit cell lengths. Data are for T = 298K.

#Calculated from gas phase structure.

§NMA, Normal mode analysis. Only internal vibrational frequencies that contribute 0.1 kcal/mol or more to the transfer entropy are tabulated.

¶Six lowest frequencies from Discover crystal-phase vibrational analysis.

‖Frequencies corresponding to angular perturbations about the x, y, and z axes, respectively.

**Frequencies corresponding to translational perturbations in the x, y, and z directions, respectively.

dynamics are accurate, given the rigid ring structure and absence of flexible side chains of cGG. The total experimental $T\Delta S$ is −19.5 kcal/mol and, subtracting the rotational, translational, and internal vibrational terms, gives an "experimental estimate" of 41 eu for the lattice entropy, compared to 26 eu from the NMA calculations. The cGG association entropy calculations thus overestimate the rigidity of the crystal, or underestimate its entropy, by about 15 eu.

A Hooke's law analysis (HLA) of the lattice vibrations was undertaken to see whether the harmonic assumption inherent in the NMA analysis (which only uses the curvature of the potential energy at the minimum) was reasonable, and to provide another estimate of the lattice entropy. As Fig. 4 shows, the energy is in fact a quadratic function of the translational and rotational perturbation for fairly large displacements, supporting the validity of the NMA analysis. However, the lattice vibrational entropy computed using via HLA is less than that from the NMA by 8 and 13 eu for cGG and cAA, respectively. HLA thus provides a worse estimate of the cGG lattice entropy (17.9 eu) and, hence, the association entropy (−37.2 eu).

Because sublimation data for cAA are unavailable, the association entropy of cAA is experimentally unknown. For comparison with cGG, Table 4 nonetheless includes the results of association entropy calculations on cAA. Converting cGG to cAA via the addition of methyl side chains increases both the total mass and the principal moments of inertia (Table 5), and the translational and rotational entropies rise accordingly. Both the Discover NMA and the

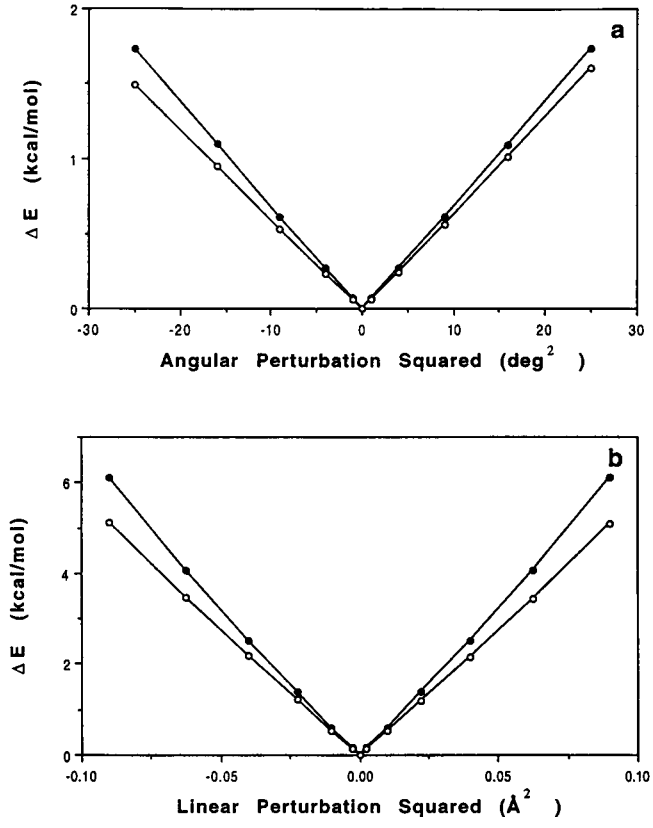

FIGURE 4 Hooke's law analysis of rigid-body perturbation of dipeptides in the crystal lattice. Change in energy from the minimized value as a function of the square of the rotational (*a*) or translational (*b*) perturbation. ○, cGG (cyclic diglycine); ●, cAA (cyclic dialanine).

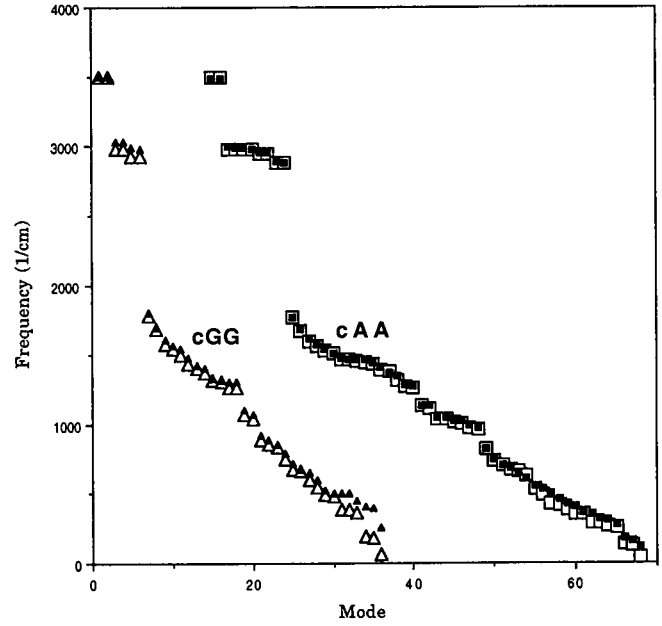

FIGURE 5 Comparison of vibrational frequencies in the gas phase (□, △) and crystal phase (■, ▲) for cGG (△, ▲) and cAA (□, ■).

explicit HLA indicate that there was more lattice vibrational entropy in cAA than in cGG, which may also be attributed to its larger mass and moments of inertia. The NMA method gives a lattice entropy difference of 6.8 eu between cGG and cAA, whereas HLA gives only 1.7 eu. The corresponding association entropy differences (cAA − cGG) are 3.7 eu and −1.4 eu for NMA and HLA, respectively.

Although not defined as part of the association entropy, intramolecular vibrational entropy (also shown in Table 4) is related to it because the intramolecular frequencies change upon association. Fig. 5 shows the intramolecular normal frequencies of cGG and cAA in crystal and gas. The highest frequency modes represent primarily hydrogen stretch vibrations about heavy atoms, the two highest being N-H stretch modes, whereas the middle grouping contains all C-H modes. The rest of the spectrum represents CO stretches, etc., and other more complicated modes. The crystal- and gas-phase frequency spectra match nicely for both compounds, except at low frequencies (<500 cm^{-1}), where the gas-phase frequencies are noticeably lower. Table 5 lists all cGG and cAA intramolecular frequencies that contributed 0.1 kcal/mol or more to $T\Delta S$ for gas-to-crystal transfer. The low-frequency modes, which contribute more to thermodynamic changes, are primarily responsible for the small (~5 eu) internal entropic barrier to crystallization for both compounds. Changes in intramolecular vibration opposed crystallization in both cGG and cAA because steric constraints imposed by the crystal environment narrow the vibrational potential wells and thus lower entropy. The change in intramolecular vibration entropy upon cAA crystallization differs negligibly from that of cGG, despite the additional methyl groups. The absolute entropy of cAA exceeds that of cGG by over 10 eu in both phases, however. The reason is that cAA has 18 more intramolecular modes than cGG and, as the larger molecule, has lower intramolecular frequencies relative to those of cGG (see Table 5).

Table 6 shows the computed heat capacity changes upon crystallization for both cGG and cAA. In the high temperature limit, the vapor-phase heat capacity was $4k_B$, of which $5/2k_B$ is from translation and $3/2k_B$ from rotation. The six modes of lattice vibration in the crystal contribute $6k_b$ to the heat capacity change upon crystallization, whereas the intramolecular vibrations make no contribution in the high temperature limit. Thus, the net gas-to-crystal heat capacity increment in this limit is $2k_B$. The calculated value of the

TABLE 6 Gas-to-crystal heat capacity changes*

Contribution	cGG	cAA
Translation	−2.5	−2.5
Rotation	−1.5	−1.5
Intramolecular vibration	−1.0	−0.6
Lattice vibration#	5.9	6.0
Net	0.9	1.4

*All entries in k_B.
#Six lowest CVFF Discover normal mode frequencies used.

heat capacity is $0.9k_B$, which differs from the high temperature value principally in the $-1.0k_B$ heat capacity increment contributed by changes in intramolecular vibrations. This calculated heat capacity change and the experimental enthalpy were then used to reanalyze the experimental vapor pressure data to obtain a better value for the extrapolated vapor pressure at 25°C (see methods). The results, however, are insensitive to the exact value of ΔC_p and differ negligibly from those obtained by using the high temperature limit, shown in Table 2.

DISCUSSION

Study of the thermodynamics of dissolution of crystalline cyclic dipeptides has provided a useful thermodynamic model for protein-protein dissociation and protein unfolding (Murphy and Gill, 1990a,b, 1991; Murphy et al., 1990). Because of this, detailed thermodynamic calculations on the gas-to-water and gas-to-crystal transfer of cyclic diglycine and dialanine were performed to determine the balance of interactions involved and to gain insights into the corresponding interactions in protein-protein complexes and the protein interior.

Considering first the enthalpy of dissolution, one striking feature of protein unfolding and cyclic dipeptide dissolution is that, as the area of hydrophobic groups exposed to water increases for such processes (and, hence, as the heat capacity per residue increases), the enthalpy of unfolding/dissolution per residue becomes more favorable (Murphy et al., 1990). We found that the crystal packing contribution changes little when the -H side chain of cGG is replaced by the methyl side chain of cAA (Table 3). This result is in agreement with a previous study of cAA and c(Pro-Leu) (Lazaridis et al., 1995), in which little difference in packing effects was found, and which contradicts a previous explanation for the correlation between peptide hydrophobicity and the favorability of crystal-to-water transfer enthalpy, in which the packing in the crystal becomes poorer as the hydrophobic side chain is made larger (Creighton, 1991; Murphy and Freire, 1992). Another explanation put forward for the protein enthalpy/heat capacity behavior is that the more hydrophobic a protein is, the more desolvated the polar groups are in the folded state, primarily through burial of hydrogen-bonded groups in the formation of tertiary interactions (Yang et al., 1992). Our analysis provides an additional possible effect. The results in Table 3 show little change in the hydrogen-bonding contribution upon addition of the methyl side chains. However, the long-range electrostatic interaction is favorable in cGG and becomes less favorable in cAA, where on average the dipolar groups are further apart (Fig. 3). This "dipole dilution" effect, if present in proteins, could also contribute to the enthalpy/heat capacity effect, in addition to the desolvation effect. This conclusion is necessarily somewhat tentative, based as it is on the study of the only two cyclic dipeptides studied by Murphy et al. (1990) for which crystal structures are currently available. The crystal structures of other aliphatic side-chain dipeptides are being determined (K. P. Murphy, personal communication), and, when available, their energetics will be examined in detail. In addition, detailed examination of the long-range dipolar interactions in a series of proteins that have different specific enthalpies of unfolding, such as the set studied experimentally by Privalov, Murphy, and Gill (Murphy et al., 1990; Privalov and Gill, 1988), can be used to test this explanation. This analysis is currently under way.

The solvation calculations of the FDPB/γ method are based on free energy parameterizations, and so do not explicitly provide an enthalpy/entropy breakdown. However, if one assumes that the electrostatics provides the enthalpic term, whereas the surface area part provides the entropic part, the enthalpy of crystal-to-water dissolution is $-18.5 - (-24.6) = 6.1$ kcal/mol, compared to the experimental value of 6.3 kcal/mol (Murphy and Gill, 1990a) (Table 7). Privalov and Makhatadze previously observed that the absolute values of gas-to-water solvation enthalpies are about 20% larger than the corresponding free energies (Privalov and Makhatadze, 1993). Using this approach instead, the enthalpy of solvation of cGG would be -19.7 kcal/mol, leading to a crystal-to-water enthalpy estimate of 4.9 kcal/mol, which is somewhat lower than the experimental value, but is still within the accuracy of these types of calculations.

Turning now to the entropy changes, it was found that in cGG the decrease in experimental gas-to-crystal transfer free energy was about 77% of the corresponding enthalpy decrease, because of an opposing entropy change of about -19.5 eu. Of this entropy loss, a small but significant

TABLE 7 Cyclic dipeptide crystal-to-solution transfer thermodynamics*

	Contribution	cGG	cAA
Entropy ($T\Delta S$)	Association	8.7 (4.2)	7.6 (4.4)
	Intramolecular	1.6	1.5
	Solvation#	−2.1	−2.4
	Net calculated	8.2 (3.7)	6.7 (3.5)
	Experimental§	3.6	1.0
Enthalpy/energy	Inter- and intramolecular¶	24.6	21.9
	Solvation	−18.5	−17.8
	Net calculated	6.1	4.1
	Experimental§	6.3	3.3
Free energy	Net calculated	−2.1 (2.4)	−2.6 (0.6)
	Experimental§	2.7	2.3

*All energies in kcal/mol, entropies expressed as $T\Delta S$ at T = 298K. Figures in brackets are obtained assuming the lattice entropy is 74.5% of the translational plus rotational entropy (i.e., assuming the experimental specific residual entropy defined by Eq. 26 of cGG applies to both solutes). Other data are taken from Tables 1 and 2.

#Assuming that hydrophobic contribution is purely entropic, and the electrostatic term is purely enthalpic (see Discussion).

§Data obtained from Murphy and Gill (see Methods).

¶Includes −2.5 kcal/mol association/vibrational contribution.

contribution was due to restriction of internal vibrations, the NMA treatment of which was assumed to be accurate for these small rigid molecules. This assumption allows one to accurately extract an association entropy contribution of −14.1 eu from the experimental value for total entropy. The association entropy accounted for most of the net gas-to-crystal entropy loss. Because the rotational and translational contributions are essentially exact, they can be used to extract an accurate estimate of the lattice entropy from the experimental data. The sum of translational and rotational entropies is −55.1 eu (using the crystal concentration as the standard state), which yields an "experimental" estimate of the lattice entropy of 41.0 eu. The presumed accuracy of this value enables us to assess the accuracy of the lattice entropies obtained with the CVFF force field and the NMA method.

Two kinds of lattice entropy calculations were performed, both within the Einstein model of lattice vibrations. The first was a Hooke's law analysis, which provided a lattice entropy estimate of only 17.9 eu. While confirming the harmonic nature of the whole-body motions of the dipeptide in the crystal that were assumed for NMA, this analysis only provides a lower limit on the entropy, because there is no coupling of the whole-body motions to internal motions or to motions in neighboring molecules in the lattice. This coupling would have the effect of increasing the overall entropy of the system. Normal Mode Analysis with periodic boundary (minimum image) conditions allows the molecule to undergo internal and whole-body vibrational modes in a fixed array of neighboring molecules, and thereby incorporates the coupling of internal and whole-body vibrations. Because of this coupling NMA analysis gave a considerably larger lattice entropy of 25.9 eu for cGG, an increase of 8 eu over the Hooke's law model. Coupling between unit cells (i.e., lattice phonons), however, remains unaccounted for, and consequently the lattice model is still Einsteinian. This would account for the underestimate of the experimental lattice entropy in these calculations by 15 eu. The periodic boundary condition facility in Discover unfortunately does not include coupling between the motions of different molecules in the lattice, i.e., a full lattice dynamics calculation that involves adding lattice phonon wave vector terms to the dynamical matrix (Ashcroft and Mermin, 1976). This involves a substantial amount of calculation that lies outside the scope of the present study, and we are currently developing the software to perform an accurate calculation of this term using the Discover force field. Nevertheless, this analysis of thermodynamic data on cGG provides one of the more accurate estimates of the association entropy currently available because the total entropy and internal vibrational entropies are both well determined.

Considering the gas/crystal/water cycle, if the experimental entropy of −19.3 eu for gas-to-crystal transfer is combined with Privalov and Makhatadze's observation (Privalov and Makhatadze, 1993) that free energy of solvation is about 80% of the solvation enthalpy needed to extract the solvation entropy, then the crystal-to-water entropy change is $1000(0.2x - 16.4)/298 - (-19.5) = 8.5$ eu. This value is close to, but somewhat lower than, the experimental value of 12.0 e.u. extracted from the data of Murphy and Gill (1990a). Of course, if we use our calculated gas-to-crystal entropy then, because this is too high in magnitude, the resulting crystal-to-water entropy is also too high (26 eu).

A detailed analysis of the entropic contributions to dissolutions of crystalline cGG and cAA is summarized in Table 7. The resulting breakdown of the entropy terms for cGG is somewhat different from a previous analysis by Murphy et al. (1994). They concluded that there was no solvent contribution to entropy, leading to a translational/mixing entropy contribution equal to the total entropy of dissolution, 17 eu. Of this they attribute 8 eu to mixing (obtained using a mole fraction of 1/55 for the solute in water at the 1 M standard state concentration) and 9 eu to the gain of rotational and translational entropy. Our analysis provides a different split of the crystal-to-water dissolution entropy, primarily because of an estimated −7 to −9 eu from solvation, with contributions of +14 eu from association (28 eu of which is translational), and 5.4 eu from internal vibrational entropy. It should be noted that another recent analysis of solvation also attributes a significant solvation entropy decrease to polar as well as nonpolar groups, which is consistent with the net negative solvation entropy found here (Makhatadze and Privalov, 1996).

Considering the heat capacity changes associated with the gas-to-crystal transfer, the association (translational and rotational minus lattice vibrational) contribution is $1.9k_b$ (Table 6), very close to the $2k_b$ value expected from high temperature limit analysis. There is a small but significant $-1k_b$ contribution from upward shifts in intramolecular vibrational frequencies upon crystallization. This gives a net gas-to-crystal heat capacity increment of $0.9k_b$, still not very different from that expected by the high temperature limit analysis. It is interesting in this regard that the results of Tidor and Karplus (1994) on insulin dimerization show the opposite behavior, i.e., that the net vibrational heat capacity change (internal plus whole body) is $7k_b$, which, with a maximum of $6k_b$ from the six new whole-body vibrational modes, implies that the $3n - 6$ internal vibrational contributions increase the heat capacity. However, the vibrational behavior in such a large, anharmonic system is extremely difficult to determine accurately, so this qualitative difference in heat capacity behavior may not be significant. Considering the solvation contribution, only the free energy is given directly by the FDPB/γ method, not the heat capacity change. Using previous empirical solvation heat capacity models based on surface area, however, the heat capacity of solvation of cGG can be estimated as

$$\Delta C_p^{solv} = c_p^{np} A^{np} + c_p^{pol} A^{pol}, \qquad (26)$$

where A^{np} and A^{pol} are the nonpolar and polar accessible surface areas, respectively, and c_p^{np} and c_p^{pol} are the heat capacity changes per unit nonpolar or polar surface area. Using $c_p^{np} = 0.32$ cal/mol/K/Å and $c_p^{pol} = -0.14$ cal/mole/K/Å (Spolar et al., 1992) gives 16.4 cal/mol/K for cGG, whereas using $c_p^{np} = 0.45$ cal/mol/K/Å and $c_p^{pol} = -0.26$ cal/mol/K/Å (Xie and Freire, 1994) gives 14.0 cal/mol/K. Another approach is to use the group contribution values of Makhatadze and Privalov (1990) of −10.4 cal/mol/K for the CHCONH moiety, and 18 cal/mol/K for the glycine H side chain, giving a net value of 15 cal/mole/K for cGG. All of the methods give about the same heat capacity change. Adding this to the calculated crystal-to-gas contribution of −1.8 cal/mol/K gives a "calculated" heat capacity increment of 13.2 cal/mol/K for the crystal-to-water dissolution step. The experimental value is -3.6 ± 4 (Murphy and Gill, 1990a). We currently have no explanation for the poor agreement, except to note that both the experimental crystal-to-water value and the gas-to-water values are difficult numbers to extract from experimental data, and the results for a small molecule such as cGG in which the polar and nonpolar contributions largely cancel are likely to have rather large error bars.

The overall conclusion from our calculations, based on comparison with experimental data for cyclic diglycine, for which the most data are available, is that the solvation free energy can be calculated accurately, as can the intermolecular interaction contribution to the enthalpy. The entropy of association is negative, but is overestimated in magnitude compared to experiment, because of underestimation of the vibrational freedom in the crystal, either from over-rigid representation of the intermolecular forces or from neglect of the entropy of coupling between intermolecular vibrations and crystal lattice vibrations. Conclusions about the ability to accurately calculate heat capacity changes are ambiguous because of both experimental and theoretical uncertainties. Future directions include the extension of these calculations to other cyclic dipeptides for which thermodynamic data are known when these structures become available, and the application of these methods to the study of protein-protein and protein-ligand association.

We have enjoyed extended discussions with Kip Murphy and have benefited from access to his data on cyclic dipeptides.

Financial support from NSF grant MCB95-06900, NIH grant GM-54105, and the E. R. Johnson Research Foundation is acknowledged.

REFERENCES

Ashcroft, N. W., and N. D. Mermin. 1976. Solid State Physics. W. B. Saunders, Philadelphia.

Ben-Naim, A., and Y. Marcus. 1984. Solvation thermodynamics of nonionic solutes. *J Chem. Phys.* 81:2016–2027.

Ben-Shaul, A., N. Ben-Tal, and B. Honig. 1996. Statistical thermodynamic analysis of protein association into lipid membranes. *Biophys. J.* (in press).

Creighton, T. 1991. Stability of folded conformations. *Curr. Opin. Struct. Biol.* 1:5–16.

Degeilh, R., and R. Marsh. 1959. Three dimensional structure of diketopiperazine. *Acta Crystallogr.* 12:1007.

Gilson, M. K., J. A. McCammon, and J. A. Given. 1996. Principles of noncovalent binding. *Biophys. J.* (in press).

Hagler, A. T., E. Huler, and S. Lifson. 1974. Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. *J. Am. Chem. Soc.* 96:5319–5327.

Hill, T. L. 1985. Cooperativity Theory in Biochemistry. Springer-Verlag, New York.

Hill, T. 1986. An Introduction to Statistical Thermodynamics. Dover Books, New York.

Holtzer, A. 1995. The "Cratic correction" and related fallacies. *Biopolymers.* 35:595–602.

Horton, N., and M. Lewis. 1992. Calculation of the free energy of association for protein complexes. *Protein Sci.* 1:169–181.

Janin, J., and C. Chothia. 1978. Role of hydrophobicity in the binding of coenzymes. *Biochemistry.* 17:2943–2948.

Lazaridis, T., G. Archontis, and M. Karplus. 1995. Enthalpic contributions to protein stability. *Adv. Protein Chem.* 47:231–306.

Lee, B. K. 1991. Isoenthalpic and isoentropic temperatures and the thermodynamics of protein denaturation. *Proc. Natl. Acad. Sci. USA.* 88: 5154–5158.

Makhatadze, G., and P. Privalov. 1990. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* 213:375–384.

Makhatadze, G., and P. Privalov. 1996. On the entropy of protein folding. *Protein Sci.* 5:507–510.

McQuarrie, D. 1976. Statistical Mechanics. Harper and Row, New York.

Murphy, K., and E. Freire. 1992. Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem.* 43: 313–361.

Murphy, K., and S. Gill. 1990a. Group additivity thermodynamics for the dissolution of solid cyclic dipeptides into water. *Therm. Acta.* 172: 11–20.

Murphy, K. P., and S. J. Gill. 1990b. Calorimetric measurement of the enthalpy of dissolution of diketopiperazine in water as a function of temperature. *Thermochim. Acta.* 139:279–290.

Murphy, K. P., and S. J. Gill. 1991. Solid model compounds and the thermodynamics of protein unfolding. *J. Mol. Biol.* 222.

Murphy, K. P., P. L. Privalov, and S. J. Gill. 1990. Common features of protein unfolding and dissolution of hydrophobic compounds. *Science.* 247:559–561.

Murphy, K. P., D. Xie, K. Thompson, M. Amzel, and E. Freire. 1994. Entropy loss in biological processes: estimate of translational entropy loss. *Proteins.* 18:63–67.

Peitzsch, R. M., and S. McLaughlin. 1993. Binding of acylated peptides and fatty acids to phospholipid vesicles: pertinence to myristoylated proteins. *Biochemistry.* 32:10436–10443.

Privalov, P. L., and S. J. Gill. 1988. Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39:191–234.

Privalov, P., and G. Makhatadze. 1993. Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs free energy of hydration. *J. Mol. Biol.* 232:660–679.

Seki, S., K. Suzuki, and T. Koide. 1956. Vapor pressure of molecular crystals. *Kogyo Kagaku Zasshi.* 77:346.

Sharp, K. A., A. Jean-Charles, and B. Honig. 1992. A local dielectric constant model for solvation free energies which accounts for solute polarizability. *J. Phys. Chem.* 96:3822–3828.

Sitkoff, D., K. Sharp, and B. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988.

Sletten, E. 1970. Conformation of cyclic dipeptides. The crystal and molecular structures of cyclo-D-alanyl-L-alanyl and cyclo-L-alanyl-L-

alanyl 3,6-dimethylpipererazine-2,5-dione. *J. Am. Chem. Soc.* 92: 172–177.

Spolar, R., J. Livingstone, and M. T. Record. 1992. Use of liquid hydrocarbon and amide transfer data to estimate contributions to protein folding. *Biochemistry.* 31:3947–3955.

Sridharan, S., A. Nicholls, and B. Honig. 1992. A new vertex algorithm to calculate solvent accessible surface areas. *Biophys. J.* 61:abstract 995.

Steinberg, I., and H. Scheraga. 1963. Entropy changes accompanying association reactions of proteins. *J. Biol. Chem.* 238:172–181.

Tidor, B., and M. Karplus. 1994. The contribution of vibrational entropy to molecular association. *J. Mol. Biol.* 238:405–414.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* 7:230–252.

Xie, D., and E. Freire. 1994. Structure based prediction of protein folding intermediates. *J. Mol. Biol.* 242:62–80.

Yang, A., K. Sharp, and B. Honig. 1992. Analysis of the heat capacity dependence of protein folding. *J. Mol. Biol.* 227:889–900.