Articles

Solvation Energies of Amino Acid Side Chains and Backbone in a Family of Host−Guest Pentapeptides†

William C. Wimley,‡ Trevor P. Creamer,§ and Stephen H. White*,‡

Department of Physiology and Biophysics, University of California Irvine, California 92717-4560, and Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received January 3, 1996; Revised Manuscript Received February 14, 1996⊗

ABSTRACT: Octanol-to-water solvation free energies of acetyl amino acid amides (Ac-X-amides) [Fauchère, J. L., & Pliška, V. (1983) *Eur. J. Med. Chem.−Chim. Ther. 18*, 369] form the basis for computational comparisons of protein stabilities by means of the atomic solvation parameter formalism of Eisenberg and McLachlan [(1986) *Nature 319*, 199]. In order to explore this approach for more complex systems, we have determined by octanol-to-water partitioning the solvation energies of (1) the guest (X) side chains in the host−guest pentapeptides AcWL-X-LL, (2) the carboxy terminus of the pentapeptides, and (3) the peptide bonds of the homologous series of peptides $AcWL_m$ ($m = 1-6$). Solvation parameters were derived from the solvation energies using estimates of the solvent-accessible surface areas (ASA) obtained from hard-sphere Monte Carlo simulations. The measurements lead to a side chain solvation-energy scale for the pentapeptides and suggest the need for modifying the Asp, Glu, and Cys values of the "Fauchère−Pliška" solvation-energy scale for the Ac-X-amides. We find that the unfavorable solvation energy of nonpolar residues can be calculated accurately by a solvation parameter of 22.8 ± 0.8 cal/mol/Å^2, which agrees satisfactorily with the Ac-X-amide data and thereby validates the Monte Carlo ASA results. Unlike the Ac-X-amide data, the apparent solvation energies of the uncharged polar residues are also largely unfavorable. This unexpected finding probably results, primarily, from differences in conformation and hydrogen bonding in octanol and buffer but may also be due to the additional flanking peptide bonds of the pentapeptides. The atomic solvation parameter (ASP) for the peptide bond is comparable to the ASP of the charged carboxy terminus which is an order of magnitude larger than the ASP of the uncharged polar side chains of the Ac-X-amides. The very large peptide bond ASP, −96 ± 6 cal/mol/Å^2, profoundly affects the results of computational comparisons of protein stability which use ASPs derived from octanol−water partitioning data.

The free energy ΔG_{sc} of transferring amino side chains from an organic phase to water, the solvation energy, is generally derived from studies of the partitioning of model compounds that approximate single residues (Radzicka & Wolfenden, 1988; Fauchère & Pliška, 1983; Nozaki & Tanford, 1971), the free energies of transfer of acetyl amino acid amides (Ac-X-amide) from octanol into water being most commonly used (Eisenberg et al., 1989; Eisenberg & McLachlan, 1986; Shirley et al., 1992; Fauchère & Pliška, 1983). Such solvation energies form the basis for computing so-called atomic solvation parameters (ASP),[1] which involves parameterizing the octanol-to-water transfer free energies such that $\Delta G = \sum \Delta\sigma_i A_i$ where the A_i are the atomic solvent-accessible surface areas and the $\Delta\sigma_i$ are the ASP for atomic group i (Eisenberg et al., 1989; Eisenberg & McLachlan, 1986). Even though this formalism excludes important thermodynamic details of protein stability (*e.g.*, entropy and heat capacity) and the use of octanol as a model for the interior of proteins has not been fully validated, the solvation parameter approach has nevertheless proven to be useful in computational analyses of protein stability (Juffer et al., 1995; Wang et al., 1995; Wesson & Eisenberg, 1992; Yeates et al., 1987; Eisenberg et al., 1989; Eisenberg & McLachlan, 1986). The general idea is to calculate the free energy difference between an unfolded state and the folded state from the differences in accessible surface areas (ASA)[1] of the constituent atoms in the two states. Because of the lack of knowledge of the conformation(s) of the unfolded state, one generally assumes that a residue X in the unfolded chain has the same exposure as in a Gly-X-Gly or Ala-X-Ala peptide or in a native sequence with a fully extended conformation. Further, because of a shortage of experimental data, the solvation energy of the peptide backbone is calculated using solvation parameters computed from the side chain solvation energies (Juffer et al., 1995; Holm & Sander,

† This work was supported by grants from the National Institutes of Health GM46823 to S.H.W. and GM29458 to Dr. George Rose.

* Address correspondence to S.H.W. Phone: (714) 824-7122. FAX: (714) 824-8540. E-mail: SHWhite@uci.edu.

‡ University of California, Irvine.

§ Johns Hopkins University School of Medicine.

⊗ Abstract published in *Advance ACS Abstracts,* April 15, 1996.

[1] Abbreviations: ASA, atomic solvent-accessible surface area; ASP, atomic solvation parameter; FP, Fauchère and Pliška (1983); CD, circular dichroism spectroscopy; FMOC, 9-fluorenylmethyloxycarbonyl-; HPLC, high-performance liquid chromatography; UV, ultraviolet radiation; HEPES, 4-(2-hydroxyethyl)piperazineethanesulfonic acid; EDTA, ethylenediaminetetraacetic acid; SEM, standard error of the mean; Ac, acetyl; AcA-X-A*t*Bu, Ac-Ala-Xaa-Ala-NH-*tert*-butyl.

0006-2960/96/0435-5109$12.00/0

1992; Eisenberg et al., 1989; Chiche et al., 1990). Three issues thus arise with this approach which we consider in this paper. First, the chemical environment of each side chain in an unfolded protein depends on its neighbors and possibly its covalent linkage to a multipeptide backbone (Roseman, 1988). Second, the side chains are never fully exposed to the water in the unfolded protein because of the conformational flexibility of the polypeptide chain and the presence of neighboring side chains (Creamer et al., 1995; Rose et al., 1985). Third, the solvation energy of the peptide bond is uncertain.

To examine the first two issues, we have determined the solvation free energies of amino acid side chains in pentapeptide models which have some of the features expected of unfolded proteins. Specifically, we have measured the octanol-to-water transfer free energies of the twenty natural amino acids (X) in the guest position of the host pentapeptide AcWL-X-LL that provides neighboring nonpolar side chains of moderate size. Kim and Szoka (1992) performed similar measurements using the tripeptide host Ac-Ala-Xaa-Ala-NH-*tert*-butyl but examined only eight amino acids in the guest position. In addition to the partitioning measurements, we have performed hard-sphere Monte Carlo simulations of the pentapeptides in order to estimate ASAs which are necessary for evaluating the effects of occlusion by neighboring side chains. Such measurements permit a comparison of the side chain solvation energies of the pentapeptides with those of the Ac-X-amides before and after occlusion effects are accounted for. The comparison reveals that the uncharged polar residues are apparently less polar in the pentapeptides and suggests that the Ac-X-amide transfer free energies of the Glu, Asp, and Cys side chains should be revised. Finally, the pentapeptide measurements yield a direct measurement of the ionization free energy of the carboxy terminus.

The third issue, *i.e.*, backbone solvation energy, was examined through measurements of octanol-to-water partitioning of the homologous series of peptides $AcWL_m$ ($m = 1-6$) from which we obtain the solvation free energies of the peptide bond and the glycyl unit. Early ethanol-to-water partitioning studies of simple model compounds suggested that the glycyl unit $-CH_2-CONH-$ has a solvation energy of about -1.14 kcal/mol (Cohn & Edsall, 1943). This estimate leads to a peptide bond (CONH) ASP of approximately -50 cal/mol/Å^2 (1 Å = 0.1 nm), whereas a value of -9 is expected based upon the solvation of side chain carbonyl and amide groups (Eisenberg et al., 1989). This smaller ASP value predicts that the solvation energy of the glycyl unit should be $+0.5$ (Eisenberg et al., 1989). Our measurement reveals that the Cohn and Edsall (1943) estimate is remarkably accurate. We show that, in computational comparisons of protein stability based upon octanol−water ASPs, the opposing solvation free energies of the peptide backbone and nonpolar surface are approximately equal in magnitude. This is consistent with the recent findings of Liu and Bolen (1995) that the peptide backbone plays a crucial role in determining the stability of proteins in organic solvents used for denaturing or stabilizing proteins.

MATERIALS AND METHODS

Chemistry and Spectroscopy

Peptide Design Criteria. The design criteria for both families of peptides were that the partition coefficients be measurable for all of the peptides and that the peptides not aggregate or suffer obvious changes in secondary structure. We found that the host peptide AcWL-X-LL and the family $AcWL_m$ constituted such systems. A structure taken from a Monte Carlo computer simulation of the AcWL-F-LL peptide is shown in Figure 1. As we discuss in detail below, all of the peptides appear to be monomeric in solution and to form random coils in both water and in octanol under the conditions of the partitioning experiments.

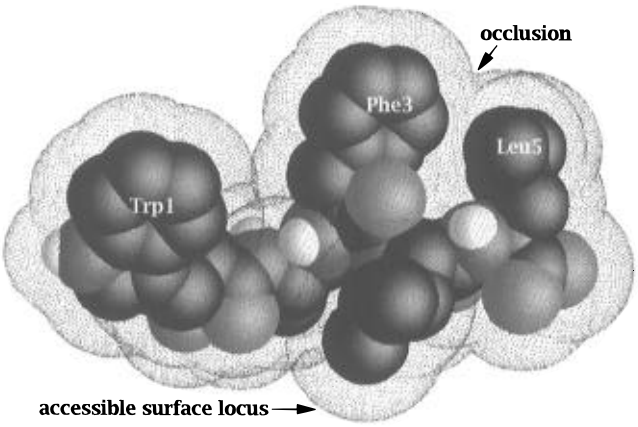

FIGURE 1: Computed conformation of the Phe member of the AcWL-X-LL peptides used in this study to estimate the energetics of solvating the 20 natural amino acids in unfolded protein chains. To derive the atomic solvation parameters from octanol-to-water partition coefficients, the solvent-accessible surface areas (ASA) of the peptides must be estimated. The locus of points derived from rolling a 1.4 Å radius sphere over the van der Waals surface is shown as the dotted surface. The average ASAs were obtained by sampling large numbers of pentapeptide conformers produced by hard-sphere Monte Carlo simulations. As shown in this example, the conformations are generally quite extended but nevertheless provide excellent estimates of ASA based upon the atomic solvation parameter results (Figure 3). An important feature of both unfolded protein chains and our peptides is occlusion of nonpolar ASA by neighboring residues. In this example, Phe3 and Leu5 mutually occlude each other's ASA. The image was created using the molecular graphics software package GRASP (Nicholls et al., 1991).

Peptide Synthesis. All of the peptides were synthesized on Wang resin using standard FMOC methodology (Atherton & Sheppard, 1989) and were cleaved for 2 h under argon in 90% trifluoroacetic acid, 5% thioanisole, 3% ethanedithiol, 2% anisole. After cleavage, peptides were extracted into 1% ammonium hydroxide from dichloromethane, lyophilized, and then purified using C18 reverse-phase HPLC and water/acetonitrile gradients with either 0.1% trifluoroacetic acid or 0.1% ammonium acetate. All peptides were better than 99% pure in both solvent systems and had the correct molecular weight by fast-atom-bombardment mass spectrometry. Concentrations of stock solutions in methanol were determined by UV absorbance.

Spectroscopy. Fluorescence spectroscopy was performed on an upgraded SPEX Fluorolog spectrometer interfaced to a computer by OLIS (Jefferson, GA) and was used to examine peptide aggregation in solution. Tryptophan emission spectra of the peptides were measured with $\lambda_{ex} = 288$, slits 10 nm.

Circular dichroism spectra were measured on a Jasco J720 CD spectrometer on samples of 50−100 μM peptide in either 50 mM phosphate buffer (pH 7.0) or in HEPES buffer-saturated octanol. All spectra were very similar and had a positive ellipticity at 225 nm that probably arises from the B-band absorption of the tryptophan residue (Woody, 1994)

and a minimum at 197 nm in buffer (∼−22,000 deg/dmol/cm^2) and in octanol (∼−17,000 deg/dmol/cm^2) that arises from peptide bond absorption. The variation between peptides in the ellipticity at 197 nm is about 3000 deg/dmol/cm^2. The shape of the spectra in both solvents and the observation that all the peptides had similar spectra suggest that the peptides have a random coil conformation in both solvents.

Peptide Aggregation and Chemical Stability

Aggregation can seriously complicate experiments with hydrophobic peptides. Although we designed these peptides to minimize solubility and aggregation problems, we nevertheless examined the aggregation state of the peptides in the aqueous and octanol phases as discussed below. With only one exception, we find no evidence for peptide aggregation in either phase.

Solubility. All of the host−guest peptides and all of the $AcWL_m$ peptides with $m < 5$ are soluble up to approximately 1 mM in buffer. For example, AcWL-W-LL, the most hydrophobic of the pentapeptides, has a solubility of 0.95 mM in buffer at pH 9 while AcWL-G-LL has a solubility of 3.7 mM. The aqueous-phase solubilities are at least 100-fold higher than the concentrations used in the partitioning experiments. For many of the peptides, the difference was even greater. Peptide solubility in octanol can be surmised from the aqueous solubility and the partition coefficients. It is at least as high, for the host−guest peptides with charged side chains, as the solubility of the hydrophobic members of the family in water. At low pH, the octanol solubility increases dramatically as expected. The relatively high solubility at high pH combined with the increased charge repulsion in the low dielectric environment leads us to believe that aggregation in the octanol phase does not occur. The circular dichroism measurements discussed below support this belief.

Titration Calorimetry. In order to determine if AcWL-W-LL is aggregated at high concentration in buffer, we filtered a visibly turbid solution of AcWL-W-LL with a nominal 2 mM peptide concentration to obtain a saturated solution of 0.95 mM and then used a MicroCal (North Hampton, MA) isothermal titration calorimeter to examine the heat of dilution of the peptide. Multiple titrations of 10 μL of the saturated solution of AcWL-W-LL were made at 25 °C into a cell containing 1.35 mL of buffer. These titrations gave rise to no detectable heat effects (<0.1 kcal/mol). We concluded that there is probably no significant aggregation of the peptide even in a saturated solution.

Fluorescence Spectroscopy. Fluorescence emission spectra for all of the host−guest pentapeptides were determined at 25 and 1 μM peptide concentration in buffer. After correction for the concentration difference, the position and intensity maxima of the two measurements were found to be identical to each other for all twenty peptides and very similar to free tryptophan in solution. We concluded, therefore, that the peptides probably did not aggregate in these concentration ranges which are approximately the same concentrations as were used in the partitioning experiments.

We also measured the fluorescence spectra of 5 μM solutions of AcWL-X-LL pentapeptides at pH 2 with the side chains X = Gly, Ala, Ser, Thr, Phe, Trp, and Glu and compared them with the spectra of acetyltryptophan in solution. For all of the peptides, the fluorescence spectra at pH 2 were nearly identical to acetyltryptophan in shape and in the wavelength of the intensity maximum. The possibility of aggregation was also assessed by potassium iodide (KI) quenching of 10 μM peptide solutions by titrating with KI up to 0.1 M KI. The Stern−Volmer quenching constant K_{sv} for KI quenching of the pentapeptides was found to be 6.9 ± 0.2 M^{-1} at pH 2, which is very similar to the value for acetyltryptophan (6.4 M^{-1}, pH 2). At pH 8, K_{sv} was found to be 8.1 ± 0.5 M^{-1}, which is similar to the value for acetyltryptophan at the same pH (8.2 M^{-1}). These results strongly suggest that the peptides do not aggregate in solution, even at pH 2 where they are uncharged.

The possibility of aggregation of the $AcWL_m$ peptides (m = 1−6) was also examined using fluorescence. First, the position and normalized intensity of the tryptophan peak of the peptides was compared to those of acetyltryptophan in solution. At peptide concentrations of ∼10 μM, the intensity and emission maximum of the tryptophan fluorescence were nearly identical to acetyltryptophan for all peptides except $AcWL_6$. The possibility of peptide aggregation was also assessed by potassium iodide (KI) quenching as described above. The Stern−Volmer quenching constant (Lakowicz, 1983) for KI quenching of the peptides $AcWL_m$ (m = 1−5) is K_{sv} = 8.4 ± 0.3, which is very similar that of free tryptophan (K_{sv} = 9.3) and acetyltryptophan (K_{sv} = 8.2). This result, and the spectroscopy above, suggests strongly that $AcWL_m$ peptides do not aggregate in solution for $m \leq 5$. This is consistent with our observation that partitioning is independent of concentration for all these peptides (see below). For $AcWL_6$, the fluorescence experiments indicate that the peptide aggregates in aqueous solution at concentrations above 0.5 μM. First, the normalized fluorescence intensity at 5 μM peptide is 1.4-fold higher and the emission maximum is blue shifted by 5 nm to 359 nm compared to all of the shorter peptides ($m \leq 5$). Second, the Stern−Volmer quenching constant is much smaller than for $m \leq 5$ peptides and is dependent on peptide concentration such that K_{sv} = ∼5 at 0.5 μM and ∼1 at 1.5 μM. We do not know if aggregation of $AcWL_6$ occurs at the equilibrium aqueous-phase concentration of 0.02−0.05 μM used in the partitioning experiments. Because of the possibility of aggregation, the partitioning data for $AcWL_6$ are not used in any of the analyses presented here. Our ability to detect very easily the aggregation of $AcWL_6$ and other more hydrophobic peptides (unpublished observations) with these fluorescence methods gives us confidence in our conclusion that the host−guest and $AcWL_m$ ($m < 6$) peptides do not aggregate under the conditions of the partitioning experiments.

Circular Dichroism. The CD spectra of all peptides in octanol and in buffer are consistent with a random coil structure of the peptides in both solvents (see above). Spectra were measured at concentrations approximately 5−10-fold higher than the concentrations in the partitioning experiments and were independent of concentration between 10 and 100 μM. The CD spectra for aggregates of $AcWL_5$ and $AcWL_6$ were measured in visibly turbid solutions prepared above the solubility limit of the peptides. The CD spectra of both peptides in such solutions were found to have a large negative peak at approximately 225 nm and a large maximum at 200 nm, whereas nonaggregated peptides have a small maximum at 225 nm and a minimum at 200 nm. The CD spectrum of the $AcWL_5$ filtrate, however, is indistinguishable from spectra

taken at much lower concentrations. Spectra could not be obtained from $AcWL_6$ filtrate because the concentration of the peptide was too low. We assumed from these results that peptide aggregates should generally have spectra that differ significantly from monomeric peptide. Except possibly for $AcWL_6$, we found no evidence of aggregation from the CD spectra of any of the peptides even in saturated solutions.

Concentration Dependence of Partitioning. With one exception, we did not examine systematically the partitioning of all of the peptides as functions of concentration. However, in the course of the studies the starting peptide concentrations and volumes of the phases were varied so that the equilibrium aqueous phase concentrations varied between approximately 1 and 20 μM. All partition coefficients were found to be independent of peptide concentration over this range. The concentration dependence for one of the peptides, AcWL-K-LL, was systematically examined for concentration effects as part of our recent study of the salt bridge between the side chain and carboxyl terminus (Wimley et al., 1996), and we found no concentration dependence of partitioning between 2 and 100 μM.

Chemical Stability of the Peptides. Peptide solutions in buffer were prepared from concentrated methanol solutions by evaporating the methanol under a stream of N_2 and then adding buffer. All peptides dissolved readily under these conditions, and were used in partitioning experiments within 2−3 days. Peptide stability was judged by reverse-phase HPLC. All peptides, except AcWL-C-LL, were stable in buffer for extended periods at 5 °C and at room temperature for the duration of the partitioning experiments. Specifically, buffer solutions containing Asn and Gln were examined for impurities which might result from side chain deamidation. We also examined the potential of Met to oxidize and found that oxidation did not occur in buffer. This is despite the fact that we could readily observe the oxidation products which formed rapidly from Met in the presence of 10% H_2O_2.

A very hydrophobic compound forms over a period of hours to days at the expense of the pentapeptide in solutions of AcWL-C-LL. Formation of this compound was prevented by replacing molecular oxygen with an inert gas or by adding a reducing agent such as dithiothreitol. Therefore, the compound is probably a disulfide-linked dimer of AcWL-C-LL. In partitioning experiments, the stock solutions of AcWL-C-LL were freshly made just before each experiment, so the maximum amount of dimer that formed during the partitioning experiment was about 3%−5%. Dimer formation does not affect the partition coefficient of AcWL-C-LL, however, because the monomer and dimer are distinguished by the HPLC method we used to determine peptide concentration in these experiments.

Partition Coefficients and Free Energies of Transfer

Partition Coefficients. Volume fraction octanol-to-water partition coefficients were measured using HEPES buffer (10 mM HEPES, 50 mM KCl, 1 mM EDTA, 3 mM NaN_3) at pH 9.0 and pH 1.0 by incubating an aliquot of a 2−50 μM peptide stock solution in octanol-saturated buffer with buffer-saturated octanol overnight at 25.0 °C while rotating the sample vials at 20 rpm. Typically, the concentration in the buffer phase after equilibration was 1−25 μM. After equilibration, the concentration of peptide in the buffer phase, in the octanol phase, and in the stock buffer solution were assayed by quantitative reverse-phase HPLC (Wimley & White, 1993a). Volume fraction partition coefficients are defined by $K_v = P_b/P_o$, where P_b and P_o are the peptide concentrations in the buffer and octanol phases, respectively. In many experiments, the peptide concentration in the octanol phase was determined indirectly by comparing the concentration in the buffer phase with that of the peptide stock solution so that $K_v = (V_o/V_b)(P_b)/(P_s - P_b)$ where V_o and V_b are the volumes of the octanol and buffer phases, respectively, and P_s is the relative concentration of the peptide in the stock buffer solution. The direct and indirect determinations of K_v always gave identical results. The latter method was used more frequently because octanol interferes with the HPLC analysis if volumes greater than 10 μL are injected. The volume of buffer phase that can be used in the HPLC analysis is unlimited. Water-to-octanol volume ratios, ranging between 20 and 0.05, were chosen such that the concentration of peptide in the buffer phase was about half that of the stock solution. Total volumes ranged from 1 to 20 mL. Partition coefficients were measured 5−10 times for each peptide and were always found to be independent of peptide concentration (1−25 μM), water-to-octanol volume ratio, and area of the interface between the bulk phases.

Free Energies of Transfer. Mole-fraction partition coefficients (K_x) were calculated from the volume fraction (K_v) values by $K_x = K_v(\upsilon_{wat}/\upsilon_{oct})$ where $\upsilon_{wat}/\upsilon_{oct} = 0.114$ is the ratio of the molar volumes of water and octanol. All the free energies in Table 1 are in mole-fraction units and are given by $\Delta G = -RT \ln K_x$. While mole-fraction free energies are generally used in Results and Discussion, we also provide Flory−Huggins-corrected volume-fraction units (Sharp et al., 1991) in some cases in the Appendix.

Computations

Monte Carlo Simulations. Ideally, one would like to simulate the conformations and molecular interactions of the peptides in both water and octanol. Such simulations are presently beyond the capabilities of existing force fields and computers. We therefore adopted the more achievable goal of estimating by Monte Carlo methods the amount of nonpolar accessible surface area that can be removed from contact with water upon transfer of the peptides into the octanol phase. If the nonpolar surface is completely removed from contact with water, then one needs only to know the amount of contact in the aqueous phase and can avoid simulating peptide conformations in the octanol phase. By means of X-ray diffraction measurements, Franks et al. (1993) have shown that wet octanol consists of micelles containing clusters of octanol molecules surrounding a core of water molecules that hydrate the hydroxyl groups. Franks et al. concluded that hydrated octanol "possesses a range of localized environments" capable of accommodating a wide range of solutes. One can therefore reasonably assume that the peptides can arrange themselves in a localized nonpolar environment of the octanol in such a way that nonpolar surface is completely removed from contact with water. This assumption can be tested by comparing the nonpolar solvation parameters obtained from partitioning of nonpolar solutes into octanol with those obtained from partitioning into apolar phases such as *n*-alkanes. We show in Results and Discussion that the agreement in the two cases is excellent.

Table 1: Experimental Values for the Solvation of AcWL-X-LL Peptides

partitioning of AcWL-X-LL[a]			total ASA of AcWL-X-LL (Å^2)[b]			ASA of X in AcWL-X-LL (Å^2)[c]		ASA of X in AcGG-X-GG (Å^2)[d]	
X residue	ΔG_{WLXLL}[e]	pH	A_{Tnp}	A_{Tp}	A_{Tbb}	A_{Xnp}	A_{Xp}	A_{Xnp}	A_{Xp}
Ala	0.87 ± 0.02	9	794.1	28.7	193.0	67.2	0	80.0	0
	5.81 ± 0.03	1							
Arg	2.99 ± 0.01	9	780.6	150.5	182.3	71.1	122.3	85.3	130.7
	3.87 ± 0.04	1							
Asn	0.30 ± 0.03	9	748.8	110.0	186.1	31.0	81.5	37.8	93.2
Asp	−2.46 ± 0.15	9	751.9	101.7	186.7	32.6	73.2	39.8	84.1
	5.47 ± 0.04	1							
Cys	1.23 ± 0.04	9	813.5	28.5	188.4	92.9	0	108.4	0
Gln	0.30 ± 0.03	9	764.9	116.8	184.4	50.6	88.5	61.9	97.5
Glu	−2.53 ± 0.13	9	768.1	107.7	184.6	52.4	79.3	64.0	88.0
	5.71 ± 0.03	1							
Gly	1.01 ± 0.02	9	772.2	28.6	201.3	37.4	0	43.3	0
	5.73 ± 0.03	1							
His	0.92 ± 0.02	9	812.8	71.0	183.2	100.6	42.8	116.5	47.8
	3.41 ± 0.02	1							
Ile	2.16 ± 0.01	9	846.2	28.3	180.0	133.0	0	158.1	0
Leu	2.29 ± 0.01	9	849.7	28.3	182.3	137.0	0	159.3	0
Lys	2.49 ± 0.02	9	809.5	95.6	182.8	98.6	67.4	116.0	72.3
	2.91 ± 0.03	1							
Met	1.71 ± 0.02	9	858.2	28.2	183.7	145.2	0	166.9	0
Phe	2.68 ± 0.02	9	874.1	28.3	181.6	164.1	0	187.0	0
Pro	0.90 ± 0.02	9	811.6	28.9	185.8	98.7	0	119.8	0
Ser	0.85 ± 0.02	9	768.2	64.0	190.5	43.4	35.4	51.7	41.5
	5.52 ± 0.05	1							
Thr	0.95 ± 0.02	9	791.4	56.8	185.1	71.3	28.3	85.2	34.7
	5.74 ± 0.03	1							
Trp	2.96 ± 0.01	9	882.4	54.2	178.6	177.0	26.0	199.6	28.9
Tyr	1.67 ± 0.01	9	840.2	77.7	181.1	130.6	49.4	151.8	51.3
Val	1.61 ± 0.01	9	828.3	28.4	182.9	110.2	0	133.1	0

[a] A complete set of pentapeptides of the form AcWL-X-LL was synthesized for these experiments using standard FMOC solid-phase peptide synthesis methodology (Atherton & Sheppard, 1989) and were purified with reverse-phase HPLC (see Methods). [b] Total solvent-accessible surface areas (ASA) A_{Ti} for nonpolar (i = np), polar (i = p), and backbone (i = bb) were determined using a 1.4 Å radius sphere by sampling peptide conformations generated in Monte Carlo simulations using hard-sphere potentials (see Methods). Included in the backbone are the peptide bonds and the carboxy terminus. The ASA of the carboxy terminus shows little variation with X and is approximately 69 Å^2. [c] The nonpolar and polar ASAs A_{Xi} of each residue in an AcWL-X-LL peptide (i = np or p; see note *b*). The ASA values for AcWL-X-LL are smaller than the AcGG-X-GG ASA by 18% ± 3%. [d] The polar and nonpolar ASA contribution of each side chain in an AcGG-X-GG peptide. These values are taken as the maximum ASA of a side chain with full solvent exposure in flexible peptide chain and are very similar to the stochastic GXG areas of Rose et al. (Lesser & Rose, 1990; Rose et al., 1985). [e] The solvation free energy of each homologue was determined by measuring its partitioning between *n*-octanol and buffer at pH 9.0 and in some cases pH 1. Free energies (kcal/mol) are for transfer from octanol to water and are calculated using mole-fraction partition coefficient units (see Methods). Uncertainties are estimated from the scatter in replicate experiments.

The pentapeptides were modeled using Monte Carlo computer simulations with the Metropolis sampling algorithm (Metropolis et al., 1953) as previously described (Creamer & Rose, 1994). The Metropolis algorithm is a widely used method for generating molecular configurations from the Boltzmann distribution. Briefly, a molecular system is perturbed randomly from some initial configuration to produce a new (trial) configuration and the difference ΔE between initial and new configuration is calculated. If $\Delta E < 0$ or if $\Delta E > 0$ and a normal random deviate between 0 and 1 is less than the Boltzmann energy $\exp(-\Delta E/RT)$, then the trial configuration is accepted and thereby becomes the new initial configuration. If neither condition is satisfied, then the trial configuration is rejected (*i.e.*, the initial configuration is retained). The distribution of configurations obtained in this way will satisfy the Boltzmann distribution. The interactions between atoms within the peptides were described by a hard-sphere potential in which atoms have only excluded volume with no attractive components. United atoms were employed: CH, CH_2, and CH_3 groups were treated as single atoms with inflated radii. The atomic radii used were scaled to 90% of their van der Waals values (Bondi, 1968). The bond lengths and angles were fixed at their standard values, and the peptide units were kept rigid and planar ($\omega = 180°$). Accessible surface area (ASA) was calculated using the method of Richmond (1984) with a probe radius set to that of a water molecule (1.4 Å). Conformations of the peptides were generated by making rotations about randomly selected torsions by random amounts (from 0 to ±180°) and displacing atoms by small random distances (~0.005 Å). All of the simulations employed an equilibration period of 10^5 Monte Carlo iterations, with data being collected every 3000 steps from the subsequent 3×10^7 steps. Thus, 10^4 conformations were sampled.

ASAs were determined for all atoms in the AcWL-X-LL and $AcWL_m$ peptides for each of the 10^4 conformations sampled. The means and standard deviations (SD) of the ASAs were accumulated in order to establish for each peptide the average ASAs and their fluctuations. Total ASAs ranged from 1002 Å^2 (SD = 39 Å^2 or 3.89%) for X = Gly to 1115 Å^2 (SD = 42 Å^2, 3.77%) for X = Trp. The standard deviations of about 4% do not, of course, measure uncertainties in the mean values. Rather, they measure the range of values sampled; *i.e.*, the width of the density of ASA values. The typical per cent standard error of the mean (SEM) for $n = 10^4$ is $\%SD/n^{1/2}$ or 4×10^{-4}. Standard statistical methods can be used to compare the Gly and Trp ASA densities. The expected range of the Trp %SD for our n that includes 95% of the observations is 3.77 ± 0.05. The width of the Gly

distribution is therefore, with statistical significance, somewhat broader, as might be expected from its greater conformational flexibility.

ASAs were also determined for AcGG-X-GG in order to estimate the exposure of side chains in peptides without interfering neighboring side chains. The ASAs of Ac-X-amides could not be determined for technical reasons. However, an examination of the ASAs of several Ac-X-Gly peptides yielded ASA values for X which were only 3% larger than the values obtained from AcGG-X-GG. We therefore used the AcGG-X-GG values in calculations involving the Ac-X-amides.

Corrections for Host Side Chain Occlusion. The residues of the peptides occlude one another's ASA so that their total nonpolar ASA is not equal to the sum of the fully-exposed side chain ASA values estimated from AcGG-X-GG. The simulations revealed that occlusion of the host residues depended upon the residue in the X position. The guest-dependent occlusion of the nonpolar (np) ASA of the host residues therefore causes the hydrophobic-effect contribution of the host to vary in the partitioning experiments. We therefore corrected for host occlusion in the following way.

For reasons described in Results and Discussion, the reference pentapeptide is taken as an AcWL-X-LL pentapeptide containing a so-called virtual glycine (GLY* or G*) in the X position. The host nonpolar ASA of AcWL-X-LL is defined as $A_{host}(X) = A_{Tnp}(WLXLL) - A_{Xnp}(WLXLL)$, where Tnp refers to the total nonpolar ASA of the pentapeptide and Xnp refers to the nonpolar ASA of X. The change in the host nonpolar ASA for a G*-to-X substitution will be $\Delta A_{host} = A_{host}(X) - A_{host}(G^*)$. The corrected relative free energy contribution of the X side chain in the context of bulky neighbors is

$$\Delta G_X^{cor} = \Delta G_{WLXLL} - \Delta G_{WLG^*LL} + \Delta\sigma_{np}\Delta A_{host} \quad (1)$$

where $\Delta\sigma_{np}$ is the nonpolar solvation parameter determined from pentapeptides partitioning (see Results and Discussion). ΔG_X^{cor} accounts for the occlusion of the host by the guest residue X but does not account for the occlusion of the guest by the host. It therefore represents the free energy contribution of the X side chain when occluded by the neighboring host residues.

Corrections for Guest Side Chain Occlusion. In order to compare the X side chain free energy ΔG_{sc} of AcWL-X-LL with that of Ac-X-amide, one must further adjust ΔG_X^{cor} for occlusion by the host residues. The fully adjusted AcWL-X-LL value for ΔG_{sc} is defined as

$$\Delta G_X^{GXG} = \Delta G_X^{cor} + \Delta\sigma_{np}\Delta A_X \quad (2)$$

where $\Delta A_X = A_{Xnp}(WLXLL) - A_{Xnp}(GGXGG)$.

Computational Comparisons of Protein Stability. In order to examine the effect of the choice of the backbone atomic solvation parameter on computational comparisons of protein stability, we examined a set of proteins drawn from the sets used in the thermodynamic/computational study of Khechinashvili et al. (1995) and the computational study of Juffer et al. (1995). The high-resolution crystallographic coordinates of all of the proteins were available in the Brookhaven Protein Data Bank (Bernstein et al., 1977). The proteins used, including their size (number of residues) and PDB identifiers, were the following: crambin (46, 1CRN), pancreatic trypsin inhibitor (58, 4PTI), Ca-binding protein (75, 3CIB), plastocyanin (99, 2PCY), parvalbumin (108, 5CPV), ribonuclease A (124, 3RN3), hen lysozyme (129, 2LZT), myoglobin (153, 4MBN), phage T4 lysozyme (164, 3LZM), papain (212, 9PAP), trypsin (223, 2PTN), elastase (240, 3EST), bovine β-trypsin (245, 1TLD), α-chymotrypsin (245, 2CHA), carbonic anhydrase (256, 2CAB), carboxypeptidase (307, 5CPA), penicillopepsin (323, 3APP), and phosphoglycerate kinase (416, 3PGK).

The ASAs of the side chains and backbones of the residues in the "unfolded" state were taken as the ASAs determined for the X-residue and backbone of the AcWL-X-LL peptides because exposures in the pentapeptides appear to give more realistic estimates for proteins than the exposures in G-X-G or A-X-A models (see Results and Discussion). The ASAs of residues in the folded state were determined from the crystallographic coordinates using the program ACCESS (Lee & Richards, 1971; Richards, 1977). The primary difference between our computation and the computations of others was that we separated the peptide bond backbone atoms from all others in order to evaluate their specific contributions to the change in the solvation energy upon "folding". The solvation free energy changes were computed using side chain solvation parameters determined from our modified Fauchère−Pliška solvation-energy scale and the solvation parameter we determined for the peptide bond (see Results and Discussion). Two computations were performed for each protein based upon a low and a high value of peptide bond solvation energy: (1) the change in solvation free energy, ΔG_{loBB}, assuming the peptide bond had the same solvation energy as the uncharged polar groups determined for the side chains and (2) the change, ΔG_{hiBB}, using our measured peptide bond solvation parameter. We did not distinguish between cysteine and cystine because the collective contribution of sulfur atoms to stability is negligible in the calculations compared to the contributions of other atoms.

RESULTS AND DISCUSSION

Octanol-to-Water Pentapeptide Free Energies: Comparison with Ac-X-Amides. The octanol-to-water transfer free energies[2] ΔG for all of the host−guest pentapeptides were measured at pH 9 where the carboxyl terminus is fully deprotonated. Some of the peptides were also examined at pH 1 where the carboxyl group is protonated in order to evaluate the energetics of deprotonation. The experimental values of ΔG are summarized in Table 1. The differences in ΔG between the pH 9 and pH 1 data for the pentapeptides with the uncharged residues X = Gly, Ala, Ser, and Thr are virtually equal with mean value of -4.78 ± 0.06 (SEM) kcal/mol. We take this value to be the free energy cost of deprotonating a carboxy terminus charge.

The AcWL-X-LL data of Table 1 are compared to the Ac-X-amide data of Fauchère and Pliška (1983) (referred to hereafter as FP)[1] in Figure 2. The side chain free energies for both peptides are computed relative to the X = Ala peptides so that $\Delta G_{sc} = \Delta G_X - \Delta G_A$. Although the correlation between the two data sets is high, several

[2] Free energies are calculated using mole-fraction partition coefficients (see Methods). The use of Flory−Huggins-corrected volume-fraction partition coefficients for computing free energies is discussed in the Appendix.

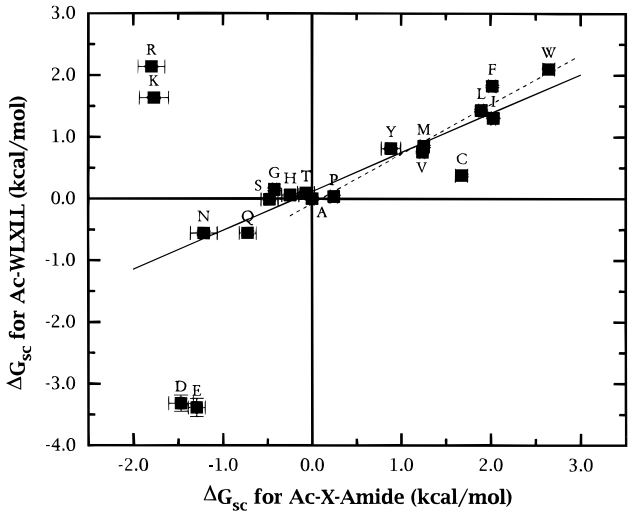

FIGURE 2: Comparison of the side chain free energies of transfer for the AcWL-X-LL peptides reported in this paper with those for the Ac-X-amides reported by Fauchère and Pliška (1983). The solid line is a linear regression through all points except K, R, D, and E. Its slope is 0.63. The dashed line has a slope of 0.8 as determined by a linear regression through the A, V, I, L, F, and W points. The slopes are less than 1 primarily because of side chain occlusion in the pentapeptides (see text). The values of K, R, D, and E for the pentapeptides are grossly different from the Ac-X-amide values. The K and R values differ because of a salt bridge between the side chain and the carboxy terminus (Wimley et al., 1996) (see text). The D and E values probably differ because of the lack of buffer in the aqueous phase of the experiments of Fauchère and Pliška (1983).

important differences are apparent. The two most striking differences are seen for the charged side chains. First, Arg and Lys in the pentapeptides appear to be about as hydrophobic as tryptophan. At pH 1, however, both side chains are very hydrophilic (Table 1). We have used ^{13}C NMR to examine this effect and find that it is due to an ionic interaction (salt bridge) in the octanol phase between the basic side chain and the acidic carboxy terminus that occurs when both groups are charged (Wimley et al., 1996). The NMR data indicate that this interaction is absent when the carboxyl group is protonated at pH 1. We therefore use the pH 1 data for Arg and Lys in our calculations of the residue solvation energies (see below). Second, Glu and Asp are much more hydrophilic in the pentapeptides than the Ac-X-amides. This may be because the aqueous phase in the FP experiments was unbuffered. Although FP reported that they adjusted the pH of the aqueous phase to pH 7.1, our experience is that this is difficult to accomplish reliably without the use of buffers. Furthermore, partition coefficients determined by radiolabeling can be strongly affected by trace impurities for compounds such as Ac-Asp-amide and Ac-Glu-amide, which have very high octanol-to-water partition coefficients. In any case, unlike our data, the FP ΔG_{sc} values for Asp and Glu are surprisingly close to values for Asn and Gln. The differences in the pentapeptide values of ΔG for X = Glu and Asp between pH 9 and pH 1 (Table 1) are approximately twice the value determined for the carboxy terminus because two carboxyl groups are titrated. Furthermore, Kim and Szoka (1992) found in their study of Ac-Ala-Xaa-Ala-NH-*tert*-butyl (henceforth referred to as AcA-X-A*t*Bu)[1] that the Asp and Glu side chains were about 2 kcal/mol more hydrophilic than the FP values (Table 2). This gives us confidence in our values of ΔG_{sc} for Asp and Glu.

Four additional, but more subtle, differences between the pentapeptide and Ac-X-amide ΔG_{sc} values are also apparent in Figure 2. First, our cysteine value is much more hydrophilic than that of FP. We note, however, that our value is consistent with the results of other workers. Radzicka and Wolfenden (1988) found for cyclohexane/water partitioning that hydrophobicity increased in the order Cys < Ala < Val < Leu. On the basis of the burial of side chains in proteins, Rose et al. (1985) found Ala < Cys < Val < Leu, which agrees with our measurements. The FP data, on the other hand, follow the series Ala < Val < Cys < Leu. Furthermore, Saunders et al. (1993) have found cystine to more hydrophobic than cysteine whereas FP obtained the opposite result. We thus suggest, as Roseman did earlier (1988), that the FP value for Cys is problematic. Second, our Gly value is more hydrophobic than the Ac-G-amide value. This difference is probably a result of the effect of Gly on the conformation of the pentapeptide (see below). Third, a linear regression of the pentapeptide data against the Ac-X-amide data (excluding Asp, Glu, Lys, and Arg) yields a slope of 0.63 ± 0.06 (solid line, Figure 2). This slope differs from a value of 1 because of mutual occlusion of the ASA of the guest and host residues and because of differences in the *apparent* solvation energies of the uncharged polar side chains (see below). The occlusion effect becomes apparent by performing a linear regression for the hydrophobic residues Ala, Val, Leu, Ile, Phe, and Trp. The resulting straight line (dashed line, Figure 2) has a slope of 0.80 ± 0.10, which is consistent with occlusion of the nonpolar surface of X by its neighbors.

The fourth difference between the pentapeptide and Ac-X-amide data revealed by Figure 2 is the discordance of the solvation energies of the uncharged polar side chains: The magnitudes of Asn, Gln, Gly, Ser, His, and Thr are smaller for AcWL-X-LL than for Ac-X-amide, and Gly, His, and Thr have opposite signs. Furthermore, the pentapeptide ΔG_{sc} values for Ser, Gly, His, and Thr are all approximately equal to Ala. Why is the solvation energy of these polar side chains apparently more unfavorable than the Ac-X-amide values? The simplest explanation is that the flanking peptide bonds reduce side chain polarity, as suggested by Roseman (1988). Interestingly, Kim and Szoka (1992) found for AcA-X-A*t*Bu that ΔG_{sc} (Table 2, relative to Ala) equaled −0.13 and +0.03 kcal/mol for Gly and His, respectively, compared to the respective FP values of −0.42 and −0.24. That is, the Kim and Szoka values are also very close to Ala and are less polar than the FP values. However, octanol is a complex interfacial phase (Franks et al., 1993) that is likely to affect the conformation and hydrogen bond formation of oligopeptides. Subtle differences in conformation between water and octanol, which are unlikely to be revealed by CD spectroscopy or accounted for by the hard-sphere simulations, could have significant effects on apparent solvation energies. The anomalous behavior of the Gly peptide, discussed below, is consistent with strong effects of conformation on apparent solvation energies. NMR measurements such as those of Kemmink et al. (1993) should reveal such differences. One scenario is that the proximity of neighboring nonpolar residues combined with peptide conformation effects could interfere with water−side chain hydrogen bond formation for the polar side chains. Alternatively, the polar side chains might cause greater exposure of the host nonpolar ASA. Another possibility is very favorable side chain−backbone

Table 2: Solvation Free Energies of the Side Chains (X) of the 20 Natural Amino Acids in AcWL-X-LL and Ac-X-Amide

		mole fraction[b]				Flory−Huggins[c]	
residue[a]	charge	ΔG_X^{cor} [d]	ΔG_X^{GXG} [e]	ΔG_X^{FP} [f]	ΔG_X^{KS} [g]	ΔG_X^{cor} [d]	ΔG_X^{GXG} [e]
Ala		+0.65	+0.81	+0.42	+0.13	+0.69	+0.99
Arg	+1	−0.66	−0.47	−1.37		+1.44	+1.81
Asn		+0.30	+0.32	−0.79		+1.06	+1.10
Asp	0	+0.72	+0.75			+1.33	+1.39
Asp	−1	−2.49	−2.46	−2.46 (−1.05)	−3.50	−1.88	−1.83
Cys		+1.17	+1.39	+1.39 (+2.10)		+1.72	+2.14
Gln		+0.38	+0.50	−0.30		+1.66	+1.90
Glu	0	+1.04	+1.17			+2.19	+2.44
Glu	−1	−2.48	−2.35	−2.35 (−0.87)	−3.12	−1.33	−1.08
GLY*		0	0	0[h]	0[h]	0[h]	0[h]
His	+1	−1.18	−0.96			+0.24	+0.68
His	0	+1.04	+1.27	+0.18	+0.16	+2.46	+2.90
Ile		+2.27	+2.70	+2.46		+3.72	+4.56
Leu		+2.40	+2.77	+2.30		+4.20	+4.92
Lys	+1	−1.65	−1.39	−1.35		+0.17	+0.67
Met		+1.82	+2.18	+1.68		+3.45	+4.14
Phe		+2.86	+3.24	+2.44	+2.19	+4.96	+5.71
Pro		+1.01	+1.35	+0.67	+0.29	+1.59	+2.25
Ser		+0.69	+0.74	−0.05		+0.78	+0.89
Thr		+0.90	+1.08	+0.35		+1.58	+1.93
Trp		+3.24	+3.62	+3.07	+2.52	+6.15	+6.88
Tyr		+1.86	+2.21	+1.31		+4.08	+4.75
Val		+1.61	+1.99	+1.66		+2.86	+3.61

[a] Residue solvation free energies of the 20 natural amino acids relative to glycine calculated from the data in Table 1. Free energies were corrected for the occlusion of neighboring residue areas (see text) and for the anomalous properties of glycine (see text). [b] Residue solvation free energies calculated with mole-fraction units. [c] Residue solvation free energy calculated with the Flory−Huggins correction (Sharp et al., 1991; De Young & Dill, 1990) (see Appendix). Constituent molar volumes were taken from Makhatadze et. al. (1990). [d] Residue solvation free energies for the X residue in the context of a AcWL-X-LL peptide calculated from the free energies in Table 1 using the virtual glycine (GLY*) as the reference (see text). $\Delta G_X^{cor} = \Delta G_{WLXLL} - \Delta G_{WLG*LL} + \Delta\sigma_{np}\Delta A_{host}$ where $A_{host}(X) = A_{Tnp}(WLXLL) - A_{Xnp}(WLXLL)$. These "corrected" values account for X-dependent changes in the nonpolar ASA of the host peptide. Values for Arg and Lys were calculated from experimental free energies measured at pH 1 where the ionic interaction between the side chain and carboxyl group does not occur. ΔG_X^{cor} is the best estimate of the solvation energy of residues occluded by neighboring residues of moderate size. [e] Residue solvation free energies for the X residue in the context of a AcGG-X-GG peptide calculated from ΔG_X^{cor} and the data in Table 1. $\Delta G_X^{GXG} = \Delta G_X^{cor} + 22.8\Delta A_X$ where $\Delta A_X = A_{Xnp}(WLXLL) - A_{Xnp}(GGXGG)$. This additional correction accounts for occlusion of the guest residue by the host (see text). ΔG_X^{GXG} is the best estimate of the solvation energy of the fully exposed residue. [f] Modified Fauchère and Pliška (1983) solvation energies, relative to Gly, for the transfer of acetyl amino acid amides from n-octanol to unbuffered aqueous phase. In this modified scale, the original values of FP for Asp, Glu, and Cys have been replaced by the ΔG_X^{GXG} in the left-hand adjacent column (see text). The original values of FP for Asp, Glu, and Cys are shown in parentheses. [g] Residue solvation free energies relative, relative to Gly, for the transfer of AcA-X-AtBu tripeptides from *n*-octanol to buffer, pH 7.2. Data are those of Kim and Szoka (1992). [h] Reference state is the experimentally determined Gly value rather than GLY*.

hydrogen bonding in the octanol phase. Our data do not permit us to distinguish among these possibilities. Indeed, all of them may be occurring. All we can say at present is that the *apparent* solvation energies of uncharged polar side chains in the pentapeptide can be unfavorable.

Solvent-Accessible Surface Areas (ASA) and the Solvation Parameter Formalism. Eisenberg (1989, 1986) has generalized the model compound approach to solvation energies by parameterizing the octanol-to-water ΔG values of the Ac-X-amides (Fauchère & Pliška, 1983) in terms of atomic solvation parameters ($\Delta\sigma_i$) and atomic solvent-accessible surface areas (A_i) for atomic group i so that $\Delta G = \sum\Delta\sigma_i A_i$. In order to assess the solvation parameters in the AcWL-X-LL host−guest system, we estimated the A_i by sampling conformations generated through hard-sphere Monte Carlo simulations. An example conformation and the ASA locus for AcWL-F-LL are shown in Figure 1. In general, the conformations of all of the simulated peptides are expected to be more extended than peptides in solution because of the hard-sphere potentials used in the simulations. Nevertheless, the simulations appear to provide reasonable estimates of A_i for nonpolar residues because the resulting values of $\Delta\sigma_i$ for nonpolar atoms agree well with generally accepted values (see below). Presented in Table 1 are values for the *total* ASAs ($A_{Ti} = \sum A_i$) and the guest−side chain ASAs ($A_{Xi} = \sum A_i$) of the AcWL-X-LL peptides. The subscripts i = np, p, and bb refer to nonpolar, polar, and bb atoms, respectively. The α carbon is included in the nonpolar ASA so that the backbone ASA includes only the peptide bond and the C-terminus carboxyl. Also included in Table 1 are guest−side chain ASA values for the peptide AcGG-X-GG. Compared to these values, the X-residue values in AcWL-X-LL are 18% ± 3% smaller because of the presence of leucine and tryptophan neighbors. Similar differences in ASA are seen when protein sequences in an extended chain are compared with the sum of ideal areas based upon tripeptide G-X-G areas (Rose et al., 1985; Lesser & Rose, 1990; Livingstone et al., 1991; Khechinashvili et al., 1995).

Nonpolar Residues. From hydrocarbon solubility data (mole-fraction units) and calculations of ASA, Reynolds et al. (1974) found a linear relationship between ΔG and ASA from which they estimated that $\Delta\sigma_{np}$ has a value of 21−25 cal/mol/Å^2. Chothia (1974) estimated from the amino acid solubility measurements of Nozaki and Tanford (1971) for Ala, Val, Leu, and Phe a value of 22 cal/mol/Å^2, assuming maximum exposure for each side chain. Rose et al. (1985) included the Nozaki−Tanford Gly and Trp data and arrived at a value of 18.9 ± 0.7 using so-called stochastic ASA measurements of X in G-X-G. Although estimates of $\Delta\sigma_{np}$ such as these have been discussed exhaustively in the

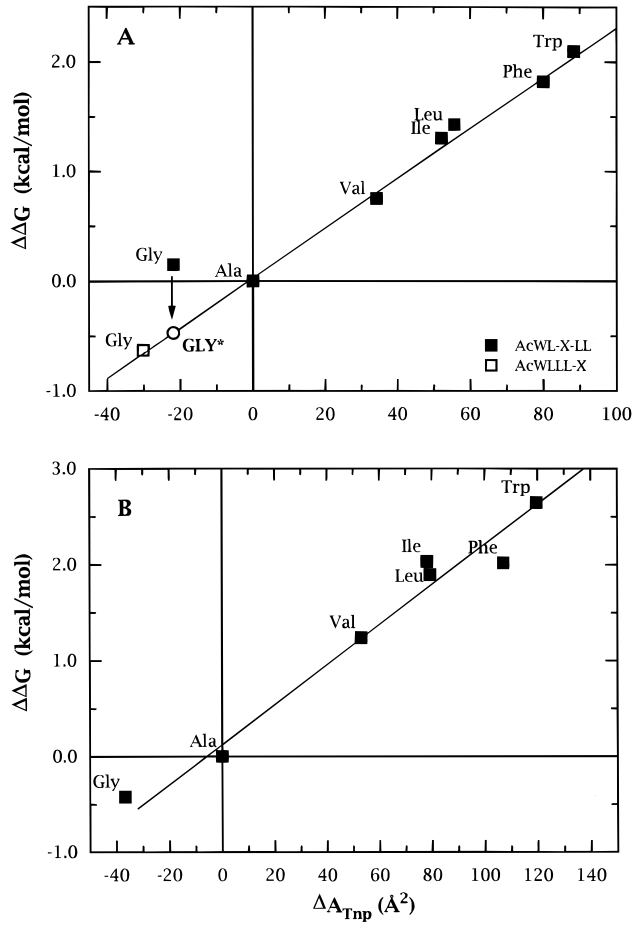

FIGURE 3: Free energy of solvation ($\Delta\Delta G$) for the hydrophobic amino acids as a function of the nonpolar relative ASA (ΔA_{Tnp}). (A) Measurements reported in this paper for AcWL-X-LL (■) and AcWLLL-X peptides (□). The diameters of the data points are about equal to the experimental uncertainties. Free energies and ASAs are relative to the Ala pentapeptides so that $\Delta\Delta G = \Delta G_{XPeptide} - \Delta G_{AlaPeptide}$ and $\Delta A_{Tnp} = A_{Tnp}(XPeptide) - A_{Tnp}(AlaPeptide)$. The straight line is a least-squares fit of the data to $\Delta\Delta G = K + \Delta\sigma_{np}\Delta A_{Tnp}$, excluding AcWL-G-LL, where $\Delta\sigma_{np}$ is the nonpolar solvation parameter. $\Delta\sigma_{np} = 22.8 \pm 0.8$ cal/mol/Å^2. Note that the value for AcWL-G-LL is anomalous in that it is considerably more hydrophobic than expected from its ASA. The origin of this effect is unknown but the observation that AcWLLL-G is not anomalous indicates that it is not a property of glycine *per se* but instead results from glycine occupying the middle position in AcWL-G-LL. Because of the glycine anomaly, we define a virtual glycine (○), GLY*, to use as a reference state (see text). (B) Measurements reported by Fauchère and Pliška (1983). In this case, $\Delta\sigma_{np} = 20.9 \pm 2.5$ cal/mol/Å^2. The excellent agreement with the AcWL-X-LL value indicates that the Monte Carlo simulations of the AcWL-X-LL peptides give reasonable estimates of ASA.

literature (Yu et al., 1995; Nozaki & Tanford, 1971; Richards, 1977; Eisenberg & McLachlan, 1986; Reynolds et al., 1974), the only consensus on the value is that it lies between 18 and 28 cal/mol/Å^2. If our Monte Carlo-based estimates of ASA are correct, then we should expect our data to yield a value $\Delta\sigma_{np}$ that is within that range. As shown in Figure 3A, our ΔG and A_{Tnp} data for X = Ala, Val, Leu, Ile, Phe, and Trp (Table 1) yield a good fit to a straight line with slope $\Delta\sigma_{np} = 22.8 \pm 0.8$ when $\Delta\Delta G = \Delta G_{AcWLXLL} - \Delta G_{AcWLALL}$ is plotted against $\Delta A_{Tnp} = A_{Tnp}(AcWLLXLL) - A_{Tnp}(AcWLALL)$. Shown in Figure 3B is a similar treatment of the FP data using the GGXGG A_{Xnp} values of Table 1. In this case, $\Delta\sigma_{np}$ has a value of 20.9 ± 2.5. Our value of $\Delta\sigma_{np}$ is thus well within the expected range. We therefore conclude that the solvation of nonpolar residues in the pentapeptides is the same as in the Ac-X-amides and that our hard-sphere Monte Carlo method for estimating the ASA of nonpolar residues is satisfactory.

An interesting feature of the data included in Figure 3A is that $\Delta\Delta G$ for the Gly peptide is anomalous: It is more hydrophobic than the Ala peptide, which is contrary to our expectations from the Ac-X-amide data of FP (Figure 3B). We hypothesized that this anomaly resulted from conformational effects arising from the presence of Gly in the middle position of the pentapeptide. To examine that possibility, we synthesized the peptides AcWLLL-A and AcWLLL-G in order to minimize the effect of Gly on conformation. The values of $\Delta\Delta G$ and ΔA_{Tnp} for those peptides are included in Figure 3A. As can be seen, the peptides have the "normal" behavior expected and thereby support our hypothesis. The conformational effect of Gly may be due to a favorable interaction in the aqueous phase between Trp and the Gly amide. NMR measurements by Kemmink et al. (1993) revealed such an interaction in tetrapeptides when an aromatic residue is separated from a Gly residue by one intervening residue, as in AcWL-G-LL. Because computations of solvation free energies are logically divided into side chain and backbone contributions in which glycine is taken as the fundamental backbone unit, we define a virtual glycine, GLY*, as shown in Figure 3A, that has the ΔA_{Tnp} of AcWL-G-LL but a value of $\Delta\Delta G$ that is 0.62 kcal mol^{-1} smaller.

Polar Residues. The solvation energies of the polar side chains are generally assumed to arise from additive contributions from the polar and nonpolar parts of the side chain. If that assumption is correct, then in a plot of $\Delta\Delta G$ against ΔA_{Tnp} for all twenty of the peptides, one would expect to find the nonpolar side chains to lie on a straight line with slope $\Delta\sigma_{np}$ as defined by Figure 3A,B. The uncharged polar residues (including protonated Asp and Glu) should be offset from the line by amounts equal to the contributions to partitioning of their polar moieties (*e.g.*, −OH, −CONH$_2$, or −COOH). If the contribution of the polar moiety (p) is described by a solvation parameter $\Delta\sigma_p$, then one should be able to account for the offsets of the polar residue data points by means of the equation (Eisenberg et al., 1989)

$$\Delta\Delta G = \Delta\sigma_{np}\Delta A_{np} + \Delta\sigma_p\Delta A_p \qquad (3)$$

This approach is tested in Figures 4 and 5 for the Ac-X-amide and the pentapeptide data sets, respectively. For the Ac-X-amides (Figure 4), all of the polar residues lie below the nonpolar reference line as expected. Interestingly, however, the offsets from the reference line of the hydroxylated side chains appear to be length dependent while the amidated side chains do not (inset, Figure 4). The fact that all of the polar side chains lie below the reference line makes it possible to compute consistent values of $\Delta\sigma_p$ using eq 3 (Eisenberg et al., 1989; Eisenberg & McLachlan, 1986). However, the apparent length dependence of the hydroxylated side chains will introduce uncertainties into the values.

As might be expected from their relatively unfavorable solvation energies, the behavior of the pentapeptide polar side chains (Figure 5) is quite different: Many of the polar residues, specifically Asn, Asp, Gln, Glu, Ser, and Thr, lie *above* the nonpolar reference line and therefore appear to have unfavorable free energies of solvation. Note that the sulfur-containing residue Met, which is often considered as hydrophobic, lies below the line and thus appears to have

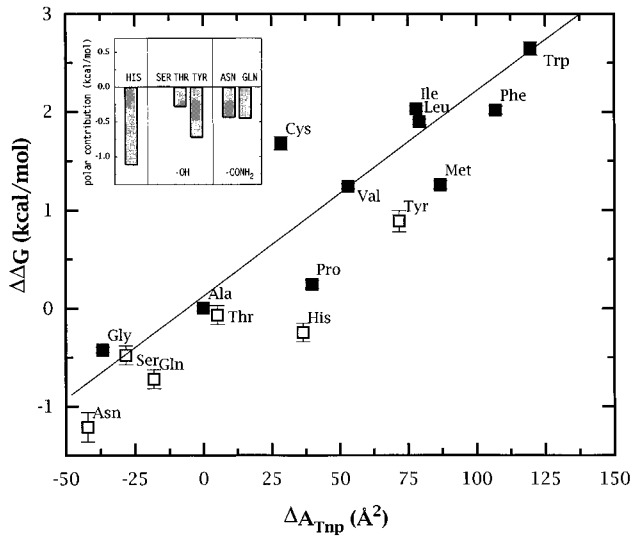

FIGURE 4: Free energy of solvation ($\Delta\Delta G$) for uncharged amino acids as a function of the nonpolar relative ASA (ΔA_{Tnp}) for the Ac-X-amides. Free energies and ASAs are relative to the Ala peptide so that $\Delta\Delta G = \Delta G_X - \Delta G_A$ (Table 2) and $\Delta A_{Tnp} = A_{Tnp}$(AcGGXGG) $- A_{Tnp}$(AcGGXGG) (Table 1). Closed squares (■) are the nonpolar residues of Figure 3B and open squares (□) the polar residues. The straight line corresponds to the hydrophobic solvation parameter of 20.9 cal/mol/Å^2 from Figure 3B. Points above the line indicate unfavorable solvation free energies of the polar moieties of polar side chains while those below indicate favorable contributions of the moieties. (Inset) The bars are the distance of the $\Delta\Delta G$ values of the polar residues above or below the 20.9 line. Here the polar moiety contributions are all favorable.

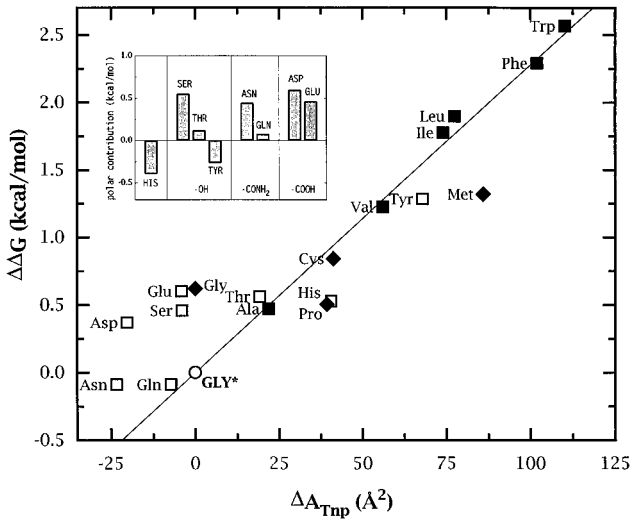

FIGURE 5: Free energy of solvation ($\Delta\Delta G$) for uncharged amino acids as a function of the nonpolar relative ASA (ΔA_{Tnp}) for the AcWL-X-LL peptides. Free energies and ASAs are relative to the virtual glycine GLY* (see text and Figure 3A) so that $\Delta\Delta G = \Delta G_{AcWLXLL} - \Delta G_{AcWLG*LL}$ and $\Delta A_{Tnp} = A_{Tnp}$(AcWLXLL) $- A_{Tnp}$(AcWLG*LL). Closed squares (■) are the nonpolar residues of Figure 3A, open squares (□) are the polar residues, and closed diamonds (◆) are those that are sometimes considered to be nonpolar or moderately polar. The straight line corresponds to the hydrophobic solvation parameter of 22.8 cal/mol/Å^2 from Figure 3A. The values of $\Delta\Delta G$ on this plot for Asp and Glu are from ΔG values determined at pH 1 and thus correspond to the uncharged side chains. Plotting the data in this fashion compensates for the solvation free energy of the nonpolar atoms of the polar residues. Points above the line indicate unfavorable solvation free energies of the polar moieties of polar side chains while those below indicate favorable contributions of the moieties. (Inset) The bars are the distance of the $\Delta\Delta G$ values of the polar residues above or below the 22.8 line. The polar moiety contributions are highly variable and, remarkably, the aqueous solvation of all of the nonaromatic polar groups is *unfavorable* and depends strongly on the size of the side chain.

polar character consistent with the ability of the Met sulfur to form hydrogen bonds in some proteins (Schulz & Schirmer, 1979). This is true for Ac-M-amide as well (Figure 4). We considered the possibility that the hydrophilicity was due to oxidation of the Met sulfur (Creighton, 1984), but control experiments showed that significant oxidation did not occur in our experiments (see Methods). Tyr, His, and Pro also lie below the line and therefore have relatively favorable free energies of solvation. The fact that polar residues are found both above and below the nonpolar line indicates that the pentapeptide data cannot be adequately reconciled through the $\Delta\sigma_p\Delta A_p$ term in eq 3. This is confirmed by the fact that the nonpolar line alone provides as good a fit, judged by χ^2 values, as the best fit of the data to eq 3. Another difficulty with the two-parameter model is shown in the inset of Figure 5, where the apparent polar moiety contributions, defined as the offset from the nonpolar reference line, are compared on the basis of polar moiety type. The contributions of all of the nonaromatic polar groups are unfavorable and depend strongly on the size of the side chain. Interestingly, the smallest polar side chains have the most unfavorable apparent polar moiety solvation energies.

The partitioning of the tryptophan side chain requires comment. We found earlier (Wimley & White, 1992, 1993b) from studies of the partitioning of indole compounds from cyclohexane to water that the solvation energy of the Trp imide group was −1.2 kcal/mol. Nevertheless, there is no significant deviation of Trp from the nonpolar reference lines in Figures 3−5. A logical explanation is that the imide group and/or the indole ring can form equally good hydrogen bonds in the octanol and buffer phases. This is consistent with the notion of Radzicka and Wolfenden (1988) that octanol exerts a "specific attraction" on Trp.

The apparent unfavorable polar moiety solvation energies of the pentapeptides, considered apart from the whole-residue energies, may indicate a failure of the hard-sphere Monte Carlo simulations to model the conformational behavior of the polar residues accurately. Notice in Figure 5 that an error in ΔA_{Tnp} of −25 to −50 Å^2 could account for the appearance of the polar residues above the reference line. Such errors could be accounted for by systematic errors of 3%−6% in the estimation of A_{Tnp}, which is about 800 Å^2 for the pentapeptides (Table 1). Systematic errors for polar side chains could occur for the obvious reason that the hard-sphere potential does not include electrostatic and hydrogen bonding effects. Such difficulties emphasize the problems inherent to computational approaches to comparisons of protein stability and thus the value of experimental studies of polypeptide models.

Side Chain Occlusion and the Energetics of Peptide Solvation. The contribution of the free energy of solvation to protein stability is generally calculated under the assumption that the side chains are fully exposed to solvent in the unfolded protein. If the assumption were correct, then the solvation free energies obtained from the Ac-X-amide partitioning data would suffice for the calculation. The nonpolar ASA data shown in Table 1 show that that assumption is incorrect because of mutual occlusion of ASA between the X residue and its neighbors. Consider, for example, the Phe peptide. A_{Xnp} for Phe in AcGG-F-GG is

144 Å^2 whereas in AcWL-F-LL it is 127 Å^2 because the neighbors occlude some of the nonpolar surface of Phe (see Figure 1). But, Phe also occludes nonpolar surface of its neighbors, directly through immediate proximity and indirectly through effects on peptide conformation. Because of these mutual occlusion effects, the net change in the nonpolar ASA accompanying the replacement of Gly by Phe is A_{Tnp}(AcWLFLL) − A_{Tnp}(AcWLGLL) = 102 Å^2 (Table 1). Thus, in a protein sequence, the true hydrophobic free energy of a Gly-to-Phe substitution could easily be reduced by (144 − 102)22.8 ≈ 1 kcal/mol compared to the fully-exposed assumption. This analysis explains why we calculated $\Delta\sigma_{np}$ using ΔA_{Tnp} rather than ΔA_{Xnp} data (Figure 3A).

Occlusion-Corrected Side Chain Solvation Energies. A desirable goal in the calculation of the contributions of solvation free energy to protein stability is to be able to perform a simple sum of the individual contributions of each residue without regard to a residue's neighbors. One can arrive at a useful single-residue estimate for such calculations using the data of Table 1 through a simple correction procedure that involves determining the effect of a G-to-X substitution on the nonpolar ASA of the host pentapeptide as described in Methods (eq 1). The corrected side chain contributions ΔG_X^{cor} are given in Table 2.

Examination of the ΔG_X^{cor} values in the first column of numbers in Table 2 reveals that all uncharged polar side chains now have unfavorable apparent solvation free energies relative to GLY*. This result, as noted earlier, contradicts our expectations based upon the partitioning of the Ac-X-amides which have a single fully-exposed side chain. To assess the extent of the contradiction, however, one must adjust the ΔG_X^{cor} values to account for the occlusion of X by neighboring residues as described in Methods (eq 2). These adjusted values, ΔG_X^{GXG}, are shown as the second column of numbers in Table 2. They indicate that correction for the occlusion of X increases the apparent unfavorable energy of solvation of the uncharged polar residues and, of course, the nonpolar residues.

Modified Fauchère−Pliška Side Chain Solvation Energies. The accuracy of these corrections for occlusion is uncertain, especially for the polar side chains, because of the limitations of the hard-sphere Monte Carlo simulations. Nevertheless, the correction process is useful because it reveals the complexities of the unfolded state that must be accounted for in computational comparisons of protein stability. The efficacy of the corrections can be assessed by comparing the ΔG_X^{GXG} data with the FP ΔG_X^{FP} data (Table 2, see footnotes). This has been done in Figure 6 where we again, as in Figure 2, plot ΔG_{sc} values using X = Ala as the reference.

The solid line passing through the origin of Figure 6 has a slope of 1 and broadly describes the relationship between the two data sets and especially so for the Lys, GLY*, Ala, Val, Met, Leu, Ile, Phe, and Trp residues. However, as in Figure 2, in which the uncorrected pentapeptide data were compared with the FP data, there are several striking differences. First, the uncharged polar residues are shifted upward by about 0.5 kcal/mol (dashed line, Figure 6), consistent with the polar side chains in the pentapeptides being apparently less polar than in the Ac-X-amides. As discussed earlier, this shift may be due to subtle conformational differences between the pentapeptides in octanol and

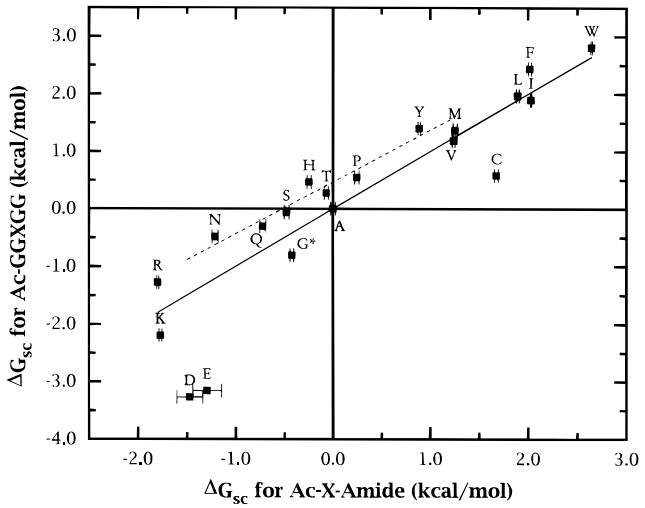

FIGURE 6: Comparison of the side chain free energies of transfer for AcGG-X-GG with those for the Ac-X-amides. The AcGG-X-GG values are derived from the AcWL-X-LL data by correcting for side chain occlusion (see text). The X = Gly used for AcGG-X-GG is the "virtual" glycine (G*, see text). The solid line passing through the origin has a slope of 1. The dashed line has a slope of 0.9 as determined by a linear regression through the N, Q, S, H, T, P, and Y points which are offset from the solid line by about +0.5 kcal/mol. This offset reveals clearly that uncharged polar residues in the pentapeptides appear to be less hydrophilic than in the Ac-X-amides. The fact that most points lie along lines with a slope of about 1 indicates that the procedure for correcting for occlusion is effective. The Asp (D), Glu (E), and Cys (C) points are outliers probably because of experimental conditions in the Ac-X-amide experiments (see text).

water that are not accounted for in the Monte Carlo simulations. However, the differences could also be due to the effect of flanking peptide bonds which Roseman (1988) suggests should reduce the polarity of polar side chains. The true origin of the effect will remain uncertain until NMR studies of the peptide conformations are completed. For the present, we attribute the effect to conformational differences and use the FP values until additional information is available.

The second major difference is that our Cys value is much more polar than the corresponding FP value. The preponderance of evidence, as discussed earlier, suggests that the pentapeptide value is more reasonable. We therefore recommend that our value be used instead of the FP value.

The third significant difference is seen for Asp and Glu, which are much more polar than the values of FP. We believe that the pentapeptide values are more reasonable for two reasons, discussed earlier. First, they are more consistent with the free energy cost of deprotonating the terminal carboxyl group. Second, the Kim and Szoka (1992) tripeptide values for Asp and Glu are similar to those of the pentapeptides. We are thus forced to conclude that the FP values for Asp and Glu are certainly too low and should be replaced by either our values or those of Kim and Szoka (1992) (Table 2). Because the Kim and Szoka values do not account for occlusion effects, we choose to use our values for Asp and Glu. The revised FP scale using our values for Asp, Glu, and Cys is presented in Table 2 as ΔG_X^{FP} (the original FP values for Asp, Glu, and Cys are shown in parentheses).

Backbone Solvation Energy. We determined the octanol-to-water partition coefficients for the homologous series of

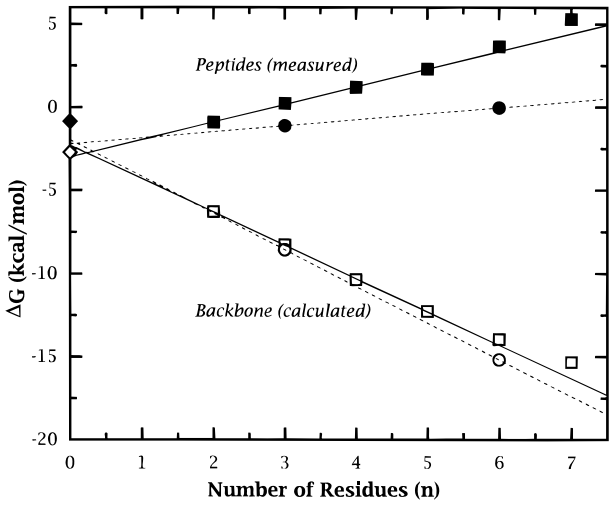

FIGURE 7: Length (n) dependence of the measured free energies of transfer ΔG (solid symbols) from octanol to water of two families of homologous peptides and the calculated free energies of transfer (open symbols) of their peptide bonds. Solid squares (■), $AcWL_m$ ($m = n - 1 = 1-6$) peptides; solid circles (●), $AcWV_m$ ($m = 2$ and 6). The peptide bond contributions of the peptides were calculated by subtracting from the measured free energies the free energy of deprotonation of the carboxy termini (−4.78 kcal/mol) and the hydrophobic-effect contribution of the Leu side chains and α carbons calculated from $22.8A_{Tnp}$. The result is the free energy of transfer of $(CONH)_nCOOH$. Open squares (□) are values calculated from the Leu peptides, and the open circles (○) are the values calculated from the Val peptides. Similarly, the free energy of transfer of a COOH group (−2.7 ± 0.05 kcal mol^{-1}, open diamond, ◊), was calculated from the free energy of transfer of acetate (solid diamond, ♦) by accounting for the hydrophobic contribution of the methyl group. The best straight line through the $n = 2-5$ backbone points yielded a slope of −2.00 ± 0.11 kcal/mol per CONH and an intercept of −2.31 ± 0.41 kcal/mol. The CONH values for Val agree reasonably well with the Leu values and, combined with the acetate results, validate the Leu measurements.

peptides $AcWL_mCOO^-$ for $m = 1-6$ (pH 9) for the purpose of determining the solvation energy of the peptide backbone. The resulting free energies of transfer are plotted as a function of peptide length ($n = m + 1$) in Figure 7 (solid squares, ■). The data are accurately linear for $n = 2-5$ but break upward for $n > 5$. Although CD spectroscopy measurements reveal no secondary structure for the longer peptides, Monte Carlo simulations using the OPLS force field suggest that the longer peptides can collapse by forming hydrogen bonds between backbone atoms (data not shown). The apparent increase in hydrophobicity for $n > 5$ is consistent with that possibility. Our computation of the solvation free energy of the backbone is therefore based upon the $n = 2-5$ ($m = 1-4$) data points. The computation is done by subtracting from the ΔG of each peptide the ionization free energy the carboxy terminus (−4.78) and the hydrophobic-effect contribution of the total nonpolar ASA of the peptide ($22.8A_{Tnp}$). This results in the calculated free energy of transfer of $(CONH)_nCOOH$ because the α carbons and the acetyl CH_3 are included in A_{Tnp}. The result of this computation is shown in Figure 7 by the open squares (□). A nonlinear least-squares fit of the points to $\Delta G = \Delta G_{COOH} + n\Delta G_{CONH}$ yielded $\Delta G_{COOH} = -2.3 \pm 0.4$ kcal/mol and $\Delta G_{CONH} = -2.00 \pm 0.11$ kcal/mol. The solvation energy of the glycyl unit $-CH_2-CONH-$ is obtained by adding the hydrophobic effect contribution of a Gly methylene. The result is $\Delta G_{glycyl} = -1.15 \pm 0.11$ kcal/mol which agrees with the ethanol-to-water value of −1.14 obtained by Cohn and Edsall (1943) from the partitioning of model compounds. It also agrees well with the values of −0.98 to −1.33 kcal/mol estimated by Tanford (1970) from peptide ethanol/water solubility data (Cohn & Edsall, 1943). Liu and Bolen (1995) have discussed extensively the importance of the backbone's preference for water compared to mixed organic solvents (such as wet octanol) for understanding protein stabilization or destabilization in various solvents.

How reliable is our estimate of ΔG_{CONH}? To answer that question, we performed two additional experiments. First, we measured the partitioning of acetate (solid diamond, ♦, Figure 7) from which we calculated the solvation energy of COOH by subtracting the hydrophobic-effect contribution of the CH_3 group. The agreement with the value of ΔG_{COOH} determined from the partitioning of the Leu peptides is satisfactory (open diamond, ◊, Figure 7). Second, we determined the solvation energies of $AcWV_2$ and $AcWV_5$ (solid circles, ●, Figure 7) from which we calculated the free energies of transfer of $(CONH)_3COOH$ and $(CONH)_6$-COOH (open circles, ○, Figure 7). The agreement is reasonably satisfactory, although the $(CONH)_6COOH$ point is statistically different from the equivalent Leu-determined point. This could indicate a dependence of peptide bond solvation energy on the side chain, but we caution that the valine peptide measurements are difficult to accomplish and are near the limits of experimental feasibility.

We considered the possibility that length-dependent changes in the ionization of the carboxy terminus, arising, *e.g.*, from steric effects (Creighton, 1984), might have affected our measurements and their interpretation. The aqueous insolubility of the longer Leu peptides precluded the possibility of measuring directly the carboxyl solvation energies of all of them. However, two experiments indicate that the carboxyl solvation energy is independent of length. First, measurements at pH 1 and pH 9 of AcWL and AcWLL showed that their carboxyl solvation energies agreed within experimental error with the value of −4.78 kcal/mol used in our computation. Second, the differences in the free energies of transfer at pH 7 and pH 9 were virtually identical for all of the Leu peptides. This indicates that there was no apparent change in solvation energy or pK_a with length. We thus believe that there were no length-dependent changes in the ionization of the carboxy terminus in our measurements.

The −2 kcal/mol solvation energy of the peptide bond is considerably larger than the polar moiety solvation energies of Ser, Thr, Tyr, Asn, Gln, and His (Figure 4) but is roughly equivalent to the polar moiety solvation energies of Asp, Glu, Arg, Lys, and His, which we estimate to be −2.36, −2.80, −1.63, −3.25, and −1.25 kcal/mol, respectively (Wimley et al., 1996). It is somewhat larger than the value of −1.74 kcal/mol estimated by Roseman (1988) from the FP data for Ac-X-amide peptide bonds but smaller than the value of −2.71 kcal/mol that he calculated by means of Hansch and Leo fragmental constants [see Hansch (1993)]. Roseman (1988) suggested that in polypeptides the flanking peptide bonds could further reduce the peptide bond free energy to −0.76 kcal/mol, which is considerably smaller than our value.

From our value of $\Delta G_{glycyl} = -1.15 \pm 0.11$ kcal/mol and the side chain solvation energy values of Table 2, one can estimate the solvation energy of a complete amino acid residue. For the pentapeptide side chain values of Table 2,

all residues favor the aqueous phase except for Val, Cys, Ile, Leu, Met, His, Tyr, Phe, and Trp. The FP side chain values behave similarly except for His, which is predicted to favor the aqueous phase.

The very favorable octanol-to-water transfer free energy of the peptide bond indicates that wet octanol provides a rather poor hydrogen bonding environment for the peptide backbone. This suggests that octanol may not be a good model for the interior of proteins in terms of backbone hydrogen bonding. The often debated issue of the contribution of backbone hydrogen bonding to protein stability remains unresolved. One must thus be cautious about incorporating octanol−water-derived peptide bond solvation energies into computational comparisons of protein stability [see below and extensive discussion of Liu and Bolen (1995)].

Computation of the Octanol-to-Water Free Energies of Transfer of Hydrophobic Peptides. The above considerations lead to the formula

$$\Delta G = -4.78 + \Delta G_{\mathrm{COOH}} + n\Delta G_{\mathrm{CONH}} + \Delta\sigma_{\mathrm{np}}A_{\mathrm{np}} \quad (4)$$

for computing the free energies of transfer (kcal/mol) of short *N*-acetyl hydrophobic peptides (*n* residues) from octanol to water where A_{np} is determined from Monte Carlo simulations of accessible surface areas and −4.78 kcal/mol is the ionization free energy of COOH. The accuracy of the formula for $n = 5$ is easily verified using the data of Table 1. It should be accurate for $n < 5$ but may be inaccurate for $n > 5$ if there is significant collapse of peptides into more compact structures.

Solvation Parameters

Side Chains. We determined the solvation parameters for both the pentapeptides and the acetyl amino acid amides (modified FP solvation energies, Table 2) in the manner of Eisenberg and McLachlan (1986) with three modifications. First, we determined the solvation parameter for the "aliphatic" groups as in Figure 3 by fitting the solvation energies for Ala, Val, Ile, Leu, Phe, and Trp to the equation

$$\Delta G_{\mathrm{np}} = K + \Delta\sigma_{\mathrm{np}}\Delta A_{\mathrm{np}} \quad (5)$$

where *K* is a constant. Second, we classified the uncharged and charged polar moieties of the side chains as single atomic groups rather than attempting to subdivide the moieties by atom type (see footnotes of Table 3). The groups are designated as polar (pol, uncharged polar groups), positive (pos, positively charged moieties), and negative (neg, negatively charged moieties). Sulfur (S) was treated as a distinct group. Third, we determined the solvation parameters using standard nonlinear least-square methods to minimize the differences between the experimentally determined solvation free energies $\Delta G_{\mathrm{exp}}(\mathrm{X})$ and those computed from the solvation parameters using

$$\Delta G_{\mathrm{calc}}(\mathrm{X}) = K + \Delta\sigma_{\mathrm{np}}A_{\mathrm{np}} + \Delta\sigma_{\mathrm{S}}A_{\mathrm{S}} + \Delta\sigma_{\mathrm{pol}}A_{\mathrm{pol}} + \Delta\sigma_{\mathrm{pos}}A_{\mathrm{pos}} + \Delta\sigma_{\mathrm{neg}}A_{\mathrm{neg}} \quad (6)$$

K and $\Delta\sigma_{\mathrm{np}}$ are held constant during the fit at the values determined using eq 5. For want of a better term, we refer to *K* as the residual. Our purpose in retaining it in eq 6 was to force the nonpolar residues to fall along a line of slope 1 when the values of $\Delta G_{\mathrm{calc}}(\mathrm{X})$ were plotted against ΔG_{exp}-(X). This helps reveal systematic variations of the calculated values from the experimental ones in the manner of Figure 3. The nonpolar ASP was held constant because Figure 3 suggests that it is accurately estimated from eq 5. The value of $\Delta\sigma_{\mathrm{np}}$ obtained from a completely unconstrained fit is invariably smaller than the one obtained using eq 5 which causes the $\Delta G_{\mathrm{calc}}(\mathrm{X})$ values for the hydrophobes to be systematically smaller than $\Delta G_{\mathrm{exp}}(\mathrm{X})$. For the pentapeptides, we used the direct experimental data of Table 1 taken relative to the Ala peptide to avoid the glycine anomaly. For the Ac-X-amides, we used the modified FP values of Table 2 but referenced them to the Ala peptide in order to be consistent with the pentapeptide calculation.

Table 3: Computed Solvation Parameters from Partitioning of AcWL-X-LL and Ac-X-Amide from Octanol to Water

atomic group	AcWL-X-LL[a]	Ac-X-amide[b]	EWY89[c]
nonpolar[d]	22.8 ± 0.8	20.9 ± 2.5	18 ± 1
sulfur[e]	16.7 ± 8.4	19.5 ± 6.6	−5 ± 6
polar[f]	4.0 ± 3.0	−6.55 ± 3.0	−9 ± 3
positive[g]	−25.0 ± 4.4	−22.3 ± 3.6	−38 ± 4
negative[h]	−34.1 ± 5.1	−27.4 ± 4.0	−37 ± 7
residual[i]	27 ± 47	113 ± 204	
peptide bond[j]	−96 ± 6		∼−50[k]
carboxy terminus[l]	−102 ± 6		

[a] Solvation parameters in units of cal/mol/Å^2 determined using eq 6 and the pentapeptide free energies ΔG of Table 1. Side chain free energies ΔG_{exp} relative to the Ala peptide were used in the determinations: $\Delta G_{\mathrm{exp}} = \Delta G_{\mathrm{WLXLL}} - \Delta G_{\mathrm{WLALL}}$. [b] Solvation parameters in units of cal/mol/Å^2 determined using eq 6 and the modified FP side chain solvation energies $\Delta G_{\mathrm{exp}} = \Delta G_{\mathrm{X}}^{\mathrm{FP}}$ of Table 2. [c] Equivalent solvation parameters determined by Eisenberg, Wesson, and Yamashita (EWY) (1989). [d] Nonpolar includes aliphatic (CH, CH_2, and CH_3) and aromatic carbons of Trp and Phe in the determination of the solvation parameter, $\Delta\sigma_{\mathrm{np}}$. Nonpolar also included the aromatic and aliphatic carbons of Tyr when finding the polar, charged, and sulfur solvation parameters. [e] Sulfur includes −SH in Cys and −S− in Met. [f] Polar atomic groups: −OH of Ser, Thr, and Tyr; $-CONH_2$ of Asn and Gln; −COOH of protonated Asp and Glu. [g] Positively charged groups: entire $-NH-C(NH_2)_2^+$ guanido group of Arg and the $-NH_3^+$ of Lys. [h] Negatively charged groups: entire $-COO^-$ of Asp and Glu. Does not include the carboxy terminus. [i] Constant term *K* in eq 5. See text. [j] CONH. [k] Value estimated by Eisenberg and McLachlan (1986) from data of Cohn and Edsall (1943). [l] Determined from the free energy of deprotonation and free energy of transfer of COOH (see text).

The side chain solvation parameters for the pentapeptides and Ac-X-amides are summarized in Table 3 and the calculated side chain solvation energies compared to the experimental ones in Figure 8. The two sets of solvation parameters are in good agreement except for $\Delta\sigma_{\mathrm{pol}}$ which is positive for AcWL-X-LL and negative for Ac-X-amide. The sign reversal occurs, of course, because the polar side chains of the pentapeptides are apparently less polar for reasons described earlier (Figures 4 and 5). For comparison, the solvation parameters of Eisenberg et al. (1989) determined from the FP Ac-X-amide data are included in Table 3 under the heading EWY89. The primary differences are that our Ac-X-amide value for the $\Delta\sigma_{\mathrm{np}}$ is larger (see above) while the $\Delta\sigma_{\mathrm{pol}}$, $\Delta\sigma_{\mathrm{pos}}$, and $\Delta\sigma_{\mathrm{neg}}$ values are smaller. The latter differences occur because we attributed the solvation energy to the entire polar moiety rather than attempting to subdivide the moiety into specific atom types. The subdivision attributes, in effect, the free energy to a smaller ASA which increases the solvation parameter. The biggest difference occurs for $\Delta\sigma_{\mathrm{S}}$ probably because of the modification of the FP data to include our pentapeptide value for Cys. Despite these differences, however, the agreement between the

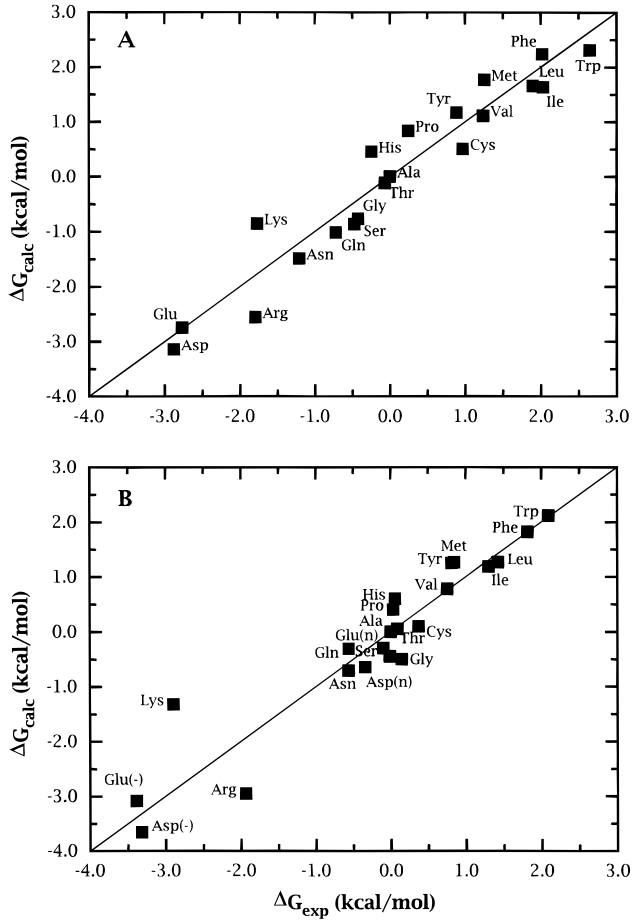

FIGURE 8: Comparisons of the side chain free energies calculated using solvation parameters [ΔG_{calc}(X); eq 6 and Table 3] with the measured values ΔG_{exp}(X). (A) Results for the modified Ac-X-amide values (Table 2). (B) Results for the AcWL-X-LL peptides (Table 1) including the data for both protonated (n) Asp and Glu and deprotonated (−) Asp and Glu. See text.

calculated and experimental values of side chain solvation energies is about equivalent to the agreement observed by Eisenberg and his colleagues (Eisenberg et al., 1989; Eisenberg & McLachlan, 1986).

Charged Terminal Carboxyl. The hard-sphere Monte Carlo calculations indicate that the accessible surface area of the carboxy terminus is 69.4 ± 0.4 Å^2. The free energy of transfer of the terminal carboxyl equals the free energy of transfer of −COOH (−2.3 kcal/mol, see above) plus the free energy cost of deprotonation (−4.78 kcal/mol) or −7.08 ± 0.40 kcal/mol. From these numbers the solvation parameter $\Delta\sigma_{COO}$ for the charged carboxy terminus is found to be −102 ± 6 cal/mol/Å^2, which is considerably larger than $\Delta\sigma_{neg}$ determined from the Asp and Glu of the pentapeptides. This is consistent with Roseman's (1988) conclusion that terminal polar groups should be much more hydrophilic than interior polar groups.

Peptide Bond. The average ASA of the Leu peptide bonds in the AcWL$_m$ peptides is 20.8 ± 1.2 Å^2, which leads to a CONH solvation parameter $\Delta\sigma_{CONH}$ of −96 ± 6 cal/mol/Å^2. This value is about twice as large as that estimated by Eisenberg and McLachlan (1986) and only slightly smaller than the carboxy terminus solvation parameter.

The large magnitude of $\Delta\sigma_{CONH}$ raises a concern about the interpretation of the ΔG_{sc} values determined for the pentapeptides. For example, a difference in exposure of the peptide bonds of 10 Å^2 between two peptides would result in a 1 kcal/mol difference in free energy that might be attributed erroneously to the side chains. Such a difference in peptide bond exposure exists (Table 1) between the Ala and Val pentapeptides, yet their difference in free energy appears to be explained entirely by the changes in nonpolar accessible surface area (Figure 3A). Considerations of the other pentapeptides in this vein does not reveal any systematic differences in ΔG_{sc} (Figure 5) that can be attributed to changes in exposure of the peptide backbone. Two possible explanations for the apparent lack of a backbone effect are errors in ASA estimates arising from deficiencies of the Monte Carlo simulations or failure of the peptide bond solvation energy to be described by $\Delta\sigma_{CONH}\Delta A_{CONH}$. The latter explanation would call into question the use of the solvation parameter formalism for computing peptide bond contributions in computational comparisons of protein stabilities were it not for the fact that about 80% of the peptide bonds are fully buried in folded proteins (see below). That is, the peptide bond contribution is given approximately by $n\Delta G_{CONH}$, where n is the number of residues.

Computational Comparisons of Protein Stability. In computational comparisons of protein stability using solvation parameters derived from octanol−water partitioning, the solvation energy of the peptide backbone is generally calculated using solvation parameters computed from the side chain solvation energies (Juffer et al., 1995; Holm & Sander, 1992; Eisenberg et al., 1989; Chiche et al., 1990). The results presented above indicate that that assumption is not valid. How large an effect will the use of the correct peptide bond solvation parameter have on such calculations? We explored that question for a number proteins by calculating the difference in solvation free energy between the folded state and a hypothetical unfolded state in which the protein chain is fully extended (see Methods). We determined the solvation free energy differences ΔG_{loBB} and ΔG_{hiBB} for the proteins using −6.55 and −96 cal/mol/Å^2, respectively, for the peptide bond solvation parameter (Table 3) by means of eq 6. The results are summarized in Figure 9. Whereas ΔG_{loBB} always favors the folded state, ΔG_{hiBB} never does. Furthermore, the use of the correct peptide bond solvation parameter leads to the conclusion that the favorable contribution of the nonpolar ASA to stability is about equal in magnitude to the unfavorable contribution of the peptide backbone. ΔG_{hiBB} is approximately given by the sum of the solvation energy differences of the polar and charged moieties of the side chains as a result (Figure 9). Because the solvation parameter formalism for computing protein stability does not include the entropy of folding, which will be unfavorable, one must suspect that the free energy cost of exchanging the backbone:water hydrogen bonds of an unfolded protein for backbone:backbone hydrogen bonds in a folded protein is likely to be smaller than the value suggested by octanol−water partitioning. Our computations indicate that about 80% of the peptide bonds are completely buried in the folded protein, consistent with the observation of Stickle et al. (1992) that 68% of the hydrogen bonds in folded proteins are of the backbone:backbone type. Thus, even modest reductions, in the free energy cost of transferring peptide bonds from water to the interior of proteins, relative to the water-to-octanol values, could dramatically alter the results of Figure 9.

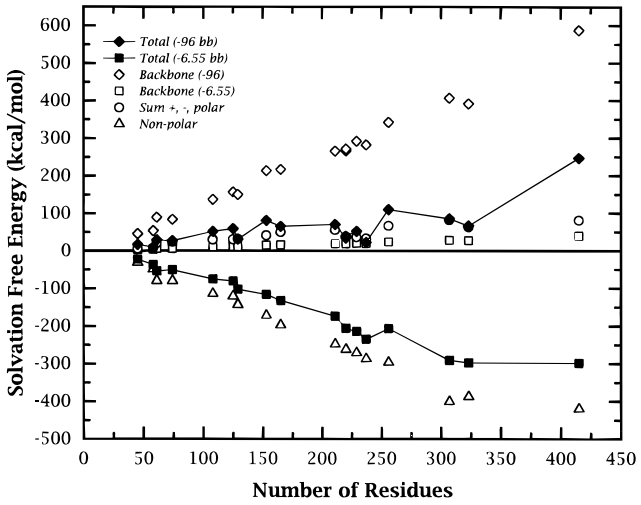

FIGURE 9: Computational comparisons of the stabilities of proteins of known structure using two different values for the peptide bond solvation parameter. The meanings of the symbols are indicated in the upper left-hand corner of the figure. The solvation free energy is calculated by means of eq 6 from the differences in the accessible surface areas between the native state determined from the crystallographic coordinates and an unfolded state defined as the fully-extended protein chain (see Methods). The contribution to the free energy of the peptide bonds assuming solvation parameters of −6.55 and −96 cal/mol/Å^2 are indicated by open squares (□) and open diamonds (◇), respectively. The total solvation free energies assuming solvation parameters of −6.55 and −96 cal/mol/Å^2 are indicated by closed squares (■) and closed diamonds (◆), respectively. Solvation energies calculated using the larger solvation parameter always disfavor the folded state. Notice that the magnitude of the nonpolar contribution (open triangles, △) is about equal in magnitude to the unfavorable contribution (open diamonds, ◇) of the peptide bonds calculated using the larger value of the solvation parameter.

Computing ΔG_{hiBB} with the set of pentapeptide solvation parameters (Table 3), which has a positive solvation parameter for uncharged polar groups, has no significant effect on the results. However, we note that the use of Flory−Huggins-corrected solvation parameters (see Appendix) leads to large negative values of ΔG_{hiBB} rather than the positive values obtained with the mole-fraction-based units used throughout this paper. This occurs because the Flory−Huggins correction doubles the nonpolar solvation parameter without significantly affecting the polar solvation parameters. Does this result support the use of the Flory−Huggins formalism? We do not believe that question can be answered for three reasons: (1) The standard formalism probably over estimates the correction, and there is no agreement as to how the correction should be made in practical cases (see Appendix); (2) octanol may not be a good model for the interior of folded proteins; and (3) the solvation parameter formalism ignores important thermodynamic details of protein folding such as entropy and heat capacity changes.

CONCLUSIONS

The overall agreement between the side chain solvation energies of the AcWL-X-LL pentapeptides and the acetyl amino acid amides is excellent except for the uncharged polar residues which are apparently much less polar (Table 2). The same conclusion is reached when the solvation parameters for the two classes of peptides are compared (Table 3). The difference in the behaviors of the uncharged polar groups may be due to the presence of a larger number of flanking peptide bonds (Roseman, 1988) or conformational effects or both. NMR measurements of peptide conformations in octanol and water may allow the role of conformational effects to be evaluated.

Analyses of the accessible surface areas of the pentapeptides determined by hard-sphere Monte Carlo simulations show how mutual occlusion of nonpolar surface area among the side chains can affect peptide solvation energies. The analyses confirm that it is inappropriate to assume that the solvent exposure of side chains in unfolded proteins is equivalent to the exposure found in Gly-X-Gly or Ala-X-Ala tripeptides. This conclusion is supported by the observation that the apparent side chain solvation energies of residue X in AcWL-X-LL are reduced by 20%−40% compared to the Ac-X-amides.

The measurement of the solvation energy of the peptide bond is consistent with the early estimate of Cohn and Edsall (1943) and reveals that the peptide bond is extremely polar. When the solvation energies of the peptide backbone and the side chains are combined, only the hydrophobic residues larger than Ala are found to favor the octanol phase unequivocally. The peptide bond solvation parameter, which normalizes the solvation energy for accessible surface area, is about equal to the solvation parameter of the carboxy terminus and is much larger than the solvation parameters of the uncharged polar side chains. In computational comparisons of protein stability which use mole-fraction-based solvation parameters derived from octanol−water partitioning, the unfavorable solvation free energy arising from the burial of peptide bonds is found to be approximately equal in magnitude to the favorable contribution arising from the burial of nonpolar surface. This result suggests that octanol−water partitioning of peptide bonds may greatly overestimate the cost of burying peptide bonds in folded proteins. Solvation parameters derived from Flory−Huggins-corrected free energies, on the other hand, lead to the conclusion that the free energy decrease arising from the burial of nonpolar surface is much larger than the increase arising from burial of peptide bonds. However, we do not believe that this finding supports the use of Flory−Huggins corrections because of uncertainties about the implementation of Flory−Huggins corrections, the appropriateness of octanol as a model for the interior of folded proteins, and the adequacy of the solvation parameter formalism as a means of computing the free energy of folding.

ACKNOWLEDGMENT

Ms. Sherin Sabet played a key role in the synthesis of the peptides, and Ms. Reneé Dillinger carried out circular dichroism measurements on all of the peptides. We are pleased to acknowledge the advice and comments of Drs. George Rose and Mark Roseman. We thank Dr. Fred Richards for providing the source code for the program ACCESS and Dr. Rodney Biltonen for the use of his titration calorimeter.

APPENDIX

Shown in the last two columns of Table 2 are values of ΔG_X^{cor} and ΔG_X^{GXG} computed using the Flory−Huggins (FH) formalism (Sharp et al., 1991; De Young & Dill, 1990) for calculating values of ΔG which includes a free energy correction for solute and solvent molar volumes. These data are given relative to the experimental Gly rather than a GLY*

because the glycine anomaly is less apparent for the FH-corrected values of $\Delta\Delta G$ for the hydrophobes than for mole-fraction unit values.

The need for the FH correction to solvent partitioning was supported by experiments reported by de Young and Dill (1990) and subsequently by a theoretical analysis of the formalism by Sharp et al. (1991). The applicability of FH partitioning free energies to protein folding is highly controversial (Sitkoff et al., 1994a,b; Lee, 1994b; Holtzer, 1994, 1995b). Two recent analyses provide convincing support (Kumar et al., 1995; Chan & Dill, 1994) but suggest that the standard formalism probably overestimates the correction. Because there is no agreement on precisely how the correction should be made in practical cases, we include the standard formalism FH side chain free energies in Table 2 and caution that the true FH solvation free energies are probably smaller. The qualitative conclusions of our work are unchanged, however, by the use of the FH formalism. Hydrophobic free energies remain proportional to ΔA_{Trp} as in Figure 3, and the apparent solvation energies for all of the polar side chains remain unfavorable as in Figure 5. The effects of FH corrections are strictly quantitative ones: The solvation parameters derived from the pentapeptide solvation-energy scale (Table 3) become, in units of cal/mol/Å^2, nonpolar, $+43.9 \pm 2.3$; sulfur, $+31.3 \pm 9.2$; polar $+17.3 \pm 3.3$; positive, -11.4 ± 4.7; negative, -19.8 ± 5.6; CONH, -85.6; COO^-, -144. The solvation parameters derived from the modified FP scale become, in units of cal/mol/Å^2, nonpolar, $+39.2 \pm 3.5$; sulfur, $+35.5 \pm 8.6$; polar -6.6 ± 3.8; positive, -10.5 ± 4.6; negative, -13.5 ± 5.1.

REFERENCES

Atherton, E., & Sheppard, R. C. (1989) *Solid Phase Peptide Synthesis*, IRL Press, Oxford.

Bernstein, F. C., Koetzle, T. F., Williams, J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Bondi, A. (1968) in *Physical Properties of Molecular Crystals, Liquids, and Glasses*, John Wiley and Sons, New York.

Chan, H. S., & Dill, K. A. (1994) *J. Chem. Phys. 101*, 7007−7026.

Chiche, L., Gregoret, L. M., Cohen, F. E., & Kollman, P. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3240−3243.

Chothia, C. (1974) *Nature (London) 248*, 338−339.

Cohn, E. J., & Edsall, J. T. (1943) in *Proteins, Amino Acids, and Peptides As Ions and Dipolar Ions*, Hafner Pub. Co., New York.

Creamer, T. P., & Rose, G. D. (1994) *Proteins: Struct., Funct., Genet. 19*, 85−97.

Creamer, T. P., Srinivasan, R., & Rose, G. D. (1995) *Biochemistry 34*, 16245−16250.

Creighton, T. E. (1984) in *Proteins. Structures and Molecular Properties*, W. H. Freeman, New York.

De Young, L. R., & Dill, K. A. (1990) *J. Phys. Chem. 94*, 801−809.

Eisenberg, D., & McLachlan, A. D. (1986) *Nature (London) 319*, 199−203.

Eisenberg, D., Wesson, M., & Yamashita, M. (1989) *Chem. Scr. 29A*, 217−221.

Fauchère, J.-L., & Pliška, V. (1983) *Eur. J. Med. Chem.−Chim. Ther. 18*, 369−375.

Franks, N. P., Abraham, M. H., & Lieb, W. R. (1993) *J. Pharm. Sci. 82*, 466−470.

Hansch, C. (1993) *Acc. Chem. Res. 26*, 147−153.

Holm, L., & Sander, C. (1992) *J. Mol. Biol. 225*, 93−105.

Holtzer, A. (1994) *Biopolymers 34*, 315−320.

Holtzer, A. (1995) *Biopolymers 35*, 595−602.

Juffer, A. H., Eisenhaber, F., Hubbard, S. J., Walther, D., & Argos, P. (1995) *Protein Sci. 4*, 2499−2509.

Kemmink, J., van Mierlo, C. P. M., Scheek, R. M., & Creighton, T. E. (1993) *J. Mol. Biol. 230*, 312−322.

Khechinashvili, N. N., Janin, J., & Rodier, F. (1995) *Protein Sci. 4*, 1315−1324.

Kim, A., & Szoka, F. C. (1992) *Pharm. Res. 9*, 504−514.

Kumar, S. K., Szleifer, I., Sharp, K., Rossky, P. J., Friedman, R., & Honig, B. (1995) *J. Phys. Chem. 99*, 8382−8391.

Lakowicz, J. R. (1983) in *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.

Lee, B. (1994) *Biophys. Chem. 51*, 263−269.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.

Lesser, G. J., & Rose, G. D. (1990) *Proteins: Struct., Funct., Genet. 8*, 6−13.

Liu, Y., & Bolen, D. W. (1995) *Biochemistry 34*, 12884−12891.

Livingstone, J. R., Spolar, R. S., & Record, M. T. (1991) *Biochemistry 30*, 4237−4244.

Makhatadze, G. I., Medvedkin, V. N., & Privalov, P. L. (1990) *Biopolymers 30*, 1001−1010.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953) *J. Chem. Phys. 21*, 1087−1092.

Nicholls, A., Sharp, K. A., & Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

Nozaki, Y., & Tanford, C. (1971) *J. Biol. Chem. 246*, 2211−2217.

Radzicka, A., & Wolfenden, R. (1988) *Biochemistry 27*, 1664−1670.

Reynolds, J. A., Gilbert, D. B., & Tanford, C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2925−2927.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151−176.

Richmond, T. J. (1984) *J. Mol. Biol. 178*, 63−89.

Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., & Zehfus, M. H. (1985) *Science 229*, 834−838.

Roseman, M. A. (1988) *J. Mol. Biol. 200*, 513−523.

Saunders, A. J., Young, G. B., & Pielak, G. J. (1993) *Protein Sci. 2*, 1183−1184.

Schulz, G. E., & Schirmer, R. H. (1979) in *Principles of Protein Structure*, Springer-Verlag, New York.

Sharp, K. A., Nicholls, A., Friedman, R., & Honig, B. (1991) *Biochemistry 30*, 9686−9697.

Shirley, B. A., Stanssens, P., Hahn, U., & Pace, C. N. (1992) *Biochemistry 31*, 725−732.

Sitkoff, D., Sharp, K. A., & Honig, B. (1994a) *Biophys. Chem. 51*, 397−409.

Sitkoff, D., Sharp, K. A., & Honig, B. (1994b) *J. Phys. Chem. 98*, 1978−1988.

Stickle, D. F., Presta, L. G., Dill, K. A., & Rose, G. D. (1992) *J. Mol. Biol. 226*, 1143−1159.

Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.

Wang, Y., Zhang, H., & Scott, R. A. (1995) *Protein Sci. 4*, 1402−1411.

Wesson, L., & Eisenberg, D. (1992) *Protein Sci. 1*, 227−235.

Wimley, W. C., & White, S. H. (1992) *Biochemistry 31*, 12813−12818.

Wimley, W. C., & White, S. H. (1993a) *Anal. Biochem. 213*, 213−217.

Wimley, W. C., & White, S. H. (1993b) *Biochemistry 32*, 9262.

Wimley, W. C., Gawrisch, K., Creamer, T. P., & White, S. H. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2985−2990.

Woody, R. W. (1994) *Eur. Biophys. J. 23*, 253−262.

Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P., & Feher, G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6438−6442.

Yu, M.-H., Weissman, J. S., & Kim, P. S. (1995) *J. Mol. Biol. 249*, 388−397.

BI9600153