

# Energetics of the interaction between water and the helical peptide group and its role in determining helix propensities

Franc Avbelj<sup>†</sup>, Peizhi Luo<sup>\*§</sup>, and Robert L. Baldwin<sup>§¶</sup>

<sup>§</sup>Department of Biochemistry, Beckman Center, Stanford University Medical Center, Stanford, CA 94305-5307; and <sup>†</sup>National Institute of Chemistry, Hajdrihova 19, Ljubljana SI 1115, Slovenia

Contributed by Robert L. Baldwin, July 21, 2000

The alanine helix provides a model system for studying the energetics of interaction between water and the helical peptide group, a possible major factor in the energetics of protein folding. Helix formation is enthalpy-driven ( $-1.0$  kcal/mol per residue). Experimental transfer data (vapor phase to aqueous) for amides give the enthalpy of interaction with water of the amide group as  $\approx -11.5$  kcal/mol. The enthalpy of the helical peptide hydrogen bond, computed for the gas phase by quantum mechanics, is  $-4.9$  kcal/mol. These numbers give an enthalpy deficit for helix formation of  $-7.6$  kcal/mol. To study this problem, we calculate the electrostatic solvation free energy (ESF) of the peptide groups in the helical and  $\beta$ -strand conformations, by using the DELPHI program and PARSE parameter set. Experimental data show that the ESF values of amides are almost entirely enthalpic. Two key results are: in the  $\beta$ -strand conformation, the ESF value of an interior alanine peptide group is  $-7.9$  kcal/mol, substantially less than that of *N*-methylacetamide ( $-12.2$  kcal/mol), and the helical peptide group is solvated with an ESF of  $-2.5$  kcal/mol. These results reduce the enthalpy deficit to  $-1.5$  kcal/mol, and desolvation of peptide groups through partial burial in the random coil may account for the remainder. Mutant peptides in the helical conformation show ESF differences among nonpolar amino acids that are comparable to observed helix propensity differences, but the ESF differences in the random coil conformation still must be subtracted.

Baker and Hubbard (1) found that water molecules cluster around peptide C=O groups in protein x-ray structures. Ben-Naim (2) and Honig and coworkers (3, 4) used model compound data to argue that interaction of water with peptide C=O groups in protein secondary structures should be an important factor in the energetics of protein folding. If the free energy of the interaction is as small as  $-0.5$  kcal/mol, and if half the peptide groups in a 100-residue protein are stripped of water molecules when folding is complete, then breaking the interactions between water and peptide C=O groups should cost 25 kcal. If some of these water-peptide interactions are not broken in forming a molten globule folding intermediate, they would represent a very important source of stabilizing free energy.

The alanine peptide helix provides a suitable model system for determining the strength of the interaction between water and the helical peptide group and for examining the effects of side chains on this interaction. Modeling studies indicate that side chains can block the access of water to backbone C=O groups in a helix (5, 6) in a side chain-specific and rotamer-specific (6) manner. Formation of the alanine helix is known to be enthalpy-driven with an enthalpy change of  $-1.0$  kcal/mol per residue (7, 8). A high-resolution x-ray structure of an alanine peptide helix is available (9). We show below that if amides such as *N*-methylacetamide are used as models for the free peptide group in an unfolded alanine peptide, then the enthalpy of the peptide hydrogen bond is not sufficient to drive helix formation, and some additional source of enthalpy is needed. The enthalpy of interaction between amides and water is large ( $\approx -11.5$

kcal/mol, Table 1), considerably larger than the enthalpy computed by quantum mechanics for the helical peptide hydrogen bond in the gas phase ( $-4.9$  kcal/mol, ref. 10). Thus, the missing enthalpy needed for helix formation is  $-11.5 - 1.0 + 4.9 = -7.6$  kcal/mol. An important component of the missing enthalpy is probably the enthalpy of interaction between water and the peptide groups in the helix.

To study this problem, we calculate values of the electrostatic solvation free energy (ESF), by using the DELPHI program and the PARSE parameter set of partial charges and Born radii (11), of the peptide groups in alanine peptides in helical,  $\beta$ -strand, and random coil conformations, and also in mutant peptides. The DELPHI program treats the solvent as a continuum of high dielectric constant and the protein as a low dielectric cavity with fixed real and partial charges and uses a finite difference method to solve the Poisson-Boltzmann equation. This approach to determining peptide solvation has the attractive feature that there are no adjustable parameters except for peptide geometry, once the decision is made to use the PARSE parameters. Comparing DELPHI and a second, independent method based on Langevin dipoles (12) gives ESF values for tripeptides that agree within 1–2 kcal/mol when the same partial charges are used in both methods (13). The partial charges and radii in the PARSE parameter set reproduce closely the solvation free energies of the set of 67 model compounds on which they are based (11); they may be regarded as empirical parameters, useful for predicting solvation free energies for new compounds of the same type (here, peptide groups) but not to be used for other purposes. The set of model compounds used to formulate the PARSE charge set does not yet include peptides, although it does include amides. The procedure used to subtract the cavity and van der Waals contributions from the experimental solvation free energy affects the values assigned to the PARSE charges. Before presenting electrostatic calculations of solvation for peptide groups in alanine peptides, we first consider simple amides for which both experimental and theoretical data for solvation free energies are available.

## Amides as Models for the Interaction with Water of the Free Peptide Group

Transfer data (vapor phase to aqueous) are available for both the enthalpy and free energy of interaction of the amide group with water. The experimentally observed transfer quantities are called solvation free energy and enthalpy after they have been

Abbreviation: ESF, electrostatic solvation free energy.

<sup>\*</sup>Present address: Xencor, 2585 Nina Street, Pasadena, CA 91107.

<sup>¶</sup>E-mail: bbaldwin@cmgm.stanford.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.200343197. Article and publication date are at [www.pnas.org/cgi/doi/10.1073/pnas.200343197](http://www.pnas.org/cgi/doi/10.1073/pnas.200343197)

**Table 1. Gas → water transfer energetics for amides: Nonpolar and electrostatic contributions to the enthalpy and standard free energy**

Solvation energies <sup>†</sup> , kcal/mol	Acetamide	<i>N</i> -methyl acetamide	<i>N,N</i> -dimethyl acetamide	Propionamide
$\Delta G_{\text{obs}}^{\ddagger}$	−9.7	−10.07	−8.54	−9.38
$\Delta G_{\text{vdw+cavity}}^{\ddagger}$	1.95	2.16	2.32	2.12
$\Delta G_{\text{pol}}^{\ddagger}$	−11.65	−12.23	−10.86	−11.50
$\Delta G_{\text{es}}^{*\ddagger}$	−11.75	−12.20	ND	−11.57
$\Delta H_{\text{obs}}^{*\ddagger}$	−16.32	−17.07	−16.56	−17.45
$\Delta H_{\text{vdw+cavity}}^{\ddagger}$	−4.67	−5.58	−6.25	−5.41
$\Delta H_{\text{pol}}^{\ddagger}$	−11.65	−11.49	−10.31	−12.04
Polar ASA (Å <sup>2</sup> ) <sup>§</sup>	95	51	32	88
Nonpolar ASA (Å <sup>2</sup> )	97	173	217	130

ND, not determined. ASA, accessible surface area.

<sup>†</sup>The units are kcal/mol, and the temperature is 25°C.

<sup>‡</sup> $\Delta G$  is the standard Gibbs free energy of transfer from the gas phase to the dilute aqueous solution (or the solvation free energy), and  $\Delta H$  is the corresponding enthalpy of transfer, or solvation enthalpy; both are referred to the standard state of Ben-Naim (17), (see also refs. 15 and 16).  $\Delta G_{\text{pol}}$  and  $\Delta H_{\text{pol}}$  are the polar contributions to the solvation free energy and enthalpy, respectively. They are found by subtracting the sum of the cavity and van der Waals terms from the solvation free energy of enthalpy.  $\Delta G_{\text{es}}^*$  is the calculated value, based on DELPHI and the PARSE parameter set, of the ESF. The transfer free energy data from amides are from Wolfenden (17) and Wolfenden *et al.* (18). The enthalpy data for amides are from Della Gatta *et al.* (19). The alkane transfer data, which are used to give the sum of the cavity and van der Waals terms, are from Cabani *et al.* (42) and Ben-Naim and Marcus (43). The equations used to fit the alkane data are  $\Delta H = 0.101 - 0.0278x$  and  $\Delta G = 0.727 + 0.0064x$ , where  $x$  is ASA in Å<sup>2</sup>; the data point for methane was not used in fitting in both cases.

<sup>§</sup>Values of polar and nonpolar ASA were calculated with NACCESS 2.1 (S. Hubbard and J. Thornton, University College, London) by using a probe radius of 1.4 Å and van der Waals radii from Chothia (44).

corrected to the standard state introduced by Ben-Naim (14), which takes account of the different volumes available to a solute molecule in the gas and liquid phases. The standard state correction is large in the case of free energy but modest in the case of enthalpy (see the discussion by Lee, refs. 15 and 16). Amides have long been used as models for the interaction of water with the free peptide group (17), and earlier studies reveal that amides interact very strongly with water. Wolfenden and coworkers (17, 18) succeeded in measuring directly the free energy of transfer of amides from the vapor phase to aqueous solution by a procedure based on isotopic labeling and detection. The experimental solvation enthalpies of amides (19) are determined by combining results from separate experiments, yielding the heat of solution and the heat of vaporization.

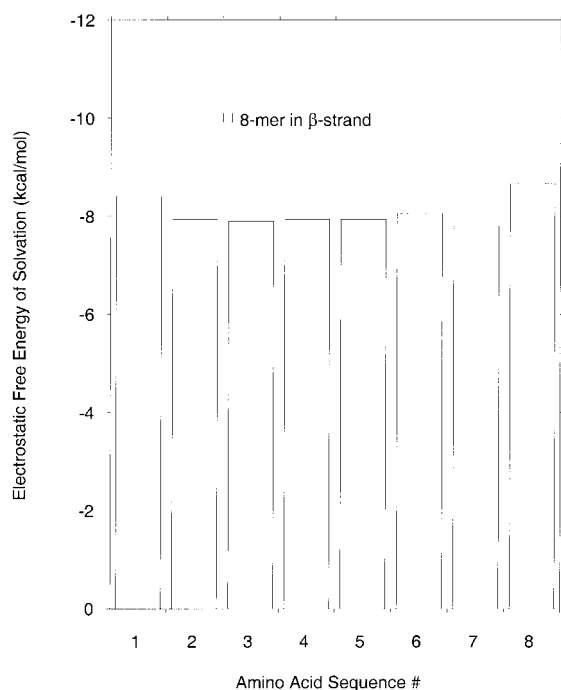
The enthalpy of interaction between the amide polar groups and water is the largest single factor when either the solvation free energy or enthalpy is measured in an amide transfer experiment (vapor phase to aqueous solution), but other terms also enter in. For amides, the enthalpy of interaction with water is approximately equal to the ESF, which may be calculated theoretically by a model with specified partial charges and atomic radii. To determine the enthalpy of interaction with water from an amide transfer experiment (which gives the solvation enthalpy), it is necessary to estimate the sum of two other terms and subtract them. The cavity term, which arises from forming a cavity in water for the solute, is chiefly entropic and makes a large unfavorable contribution to the solvation free energy. Solvent reorganization at the cavity surface introduces an enthalpic contribution into the cavity term (see discussion for nonpolar solutes, refs. 15 and 16). The van der Waals term arises from interactions between solute and solvent after insertion of the solute in the cavity.

We follow the procedure suggested by Sitkoff *et al.* (11) of approximating the sum of these two terms by experimental transfer data for nonpolar alkanes, evaluated at corresponding values of water-accessible surface area. There are no adjustable parameters in this procedure. When used together with the DELPHI program and PARSE parameter set, this approximation gives predicted values of solvation free energy that reproduce the

experimental values closely (11). By using alkane data to represent the sum of the cavity and van der Waals terms, the van der Waals interactions between water and the polar atoms of an amide are assumed to be the same as those between water and an alkane. The uncertainty caused by this approximation is difficult to evaluate (see discussion, refs. 11 and 20). Makhatadze and Privalov (21) used a similar procedure for evaluating the sum of the cavity and van der Waals terms, but their analysis of transfer enthalpies is based on assigning group enthalpies and assuming that they are additive; their value for the enthalpy of interaction between water and the peptide group (−14.2 kcal/mol) differs from the average value given here in Table 1 (−11.5 kcal/mol). They find that the interaction enthalpy is almost independent of temperature near 25°C.

The data in Table 1 show that the enthalpy and standard free energy of interaction of the amide polar groups with water are nearly equal for all four amides. This equality should not be surprising, because the polar interaction with water may be considered as H-bonding for amides in water, and the enthalpy of an amide-water H-bond is substantial whereas the loss in entropy from binding one water molecule in liquid water is small (compare ref. 21). The enthalpy of a water-amide H-bond lies between −5 and −6 kcal/mol, as computed by quantum mechanics for the gas phase (22), whereas the entropy loss from binding a water molecule is only 0–2 kcal/mol in terms of free energy, according to thermodynamic data for crystalline hydrate salts (23). Table 1 also shows that the calculated ESF values of these amides are equal to the standard free energies of the polar interaction with water, as expected (11). Different amides yield similar ESF values (Table 1) as well as similar values of the enthalpy of interaction with water.

There is no close correlation between polar accessible surface area and these values of polar free energy or enthalpy. The polar accessible surface area decreases 3-fold from acetamide to *N,N*-dimethylacetamide whereas the ESF value decreases by less than 10%. Because the ESF value is close to the enthalpy of interaction with water for amide polar groups, we use predicted values of ESF to estimate the enthalpy of interaction with water of peptide groups in alanine peptides.



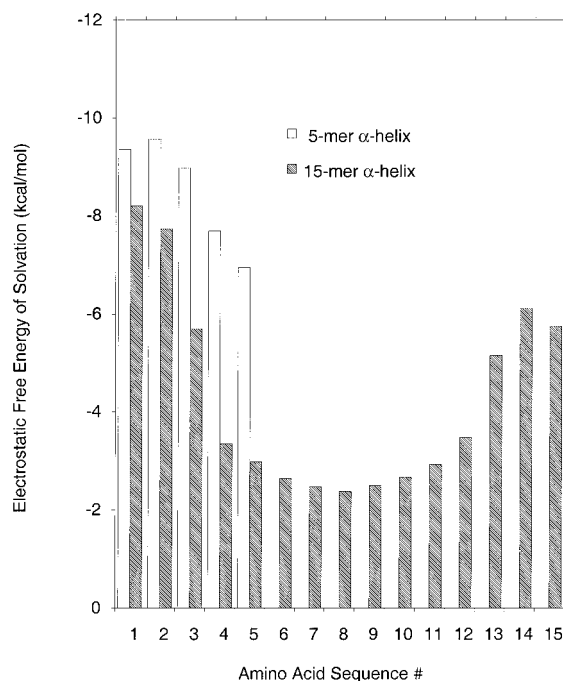
**Fig. 1.** The calculated ESF of the peptide groups in an eight-residue alanine peptide,  $\beta$ -strand conformation are plotted against residue number. With the numbering system used here, the number of peptide groups equals the number of alanine residues because, in a series of peptides with varying numbers of alanine residues, the amide bond created by joining the N-terminal acetyl and C-terminal *N*-methyl blocking groups (to make *N*-methylacetamide) is not counted as a peptide bond. The ESFs have been calculated as outlined (13) by using DELPHI and the PARSE parameter set (11); the internal dielectric constant is 2, and the external dielectric constant is 80. The peptide geometry is generated by using INSIGHT II with the  $\phi, \psi$  backbone angles set to  $-120^\circ$  and  $120^\circ$ , respectively, for the  $\beta$ -strand.

### Calculated ESF Values for Peptide Groups in Alanine Peptides

Because the peptide C=O and NH groups have large dipoles, main-chain dipole-dipole interactions have a major influence on the electrostatic potential of the peptide backbone and on conformational transitions such as the  $\alpha \rightarrow \beta$  transition (24, 25). Brant and Flory (24) found that they were unable to account for the measured dimensions of polypeptide chains without taking main-chain dipole-dipole interactions into account. Neighboring C=O and NH dipoles are antiparallel in the extended ( $\beta$ ) conformation but parallel in the  $\alpha$  conformation, so that main-chain dipole-dipole interactions strongly favor  $\beta$  over  $\alpha$ . Peptide dipoles in the main chain also interact strongly with water dipoles, and side-chain perturbation of the electrostatic interactions between water and main chain dipoles is predicted to be a main factor determining helix propensities (5, 13, 25, 26).

The predictions of ESF values below follow the same procedure as those used to fit solvation free energies of amides. The geometry of the peptide is first assigned by using INSIGHT II, whereas the DELPHI program and the PARSE parameter set are used to calculate the ESF values ( $\Delta G_{es}^*$  in Table 1) of peptide groups at various positions in the peptide. The results depend on the assigned geometry as well as on the use of the PARSE parameter set, and it should be kept in mind that experimental data for the solvation free energies of peptides are not yet available. The partial charges and radii assigned to the peptide NH and C=O groups may be changed when such information becomes available.

Fig. 1 plots ESF values of peptide groups in a blocked alanine peptide containing eight alanine residues in the  $\beta$ -strand con-



**Fig. 2.** The calculated ESFs, in kcal/mol, of the peptide groups in two helical alanine peptides are plotted against residue number. The peptides contain either five or 15 alanine residues, plus N-terminal acetyl and C-terminal *N*-methylamide blocking groups. The backbone  $\phi, \psi$  angles are  $-65^\circ$  and  $-40^\circ$  respectively. For other information, see legend to Fig. 1.

formation (no peptide H-bonds). In the numbering system used here, the amide group formed by joining the acetyl and *N*-methyl blocking groups (to make *N*-methylacetamide) is not counted as a peptide group and the number of peptide groups equals the number of alanine residues. In  $\beta$ -strand peptides with from two to five alanine residues, the exterior (N terminal and C terminal) peptide groups have uniformly high ESF values ( $-8.5$  to  $-8.3$  and  $-8.8$  to  $-8.7$  kcal/mol, respectively) and interior peptide groups have lower and nearly constant ESF values near  $-7.9$  kcal/mol, as in Fig. 1 with eight residues.

These results indicate that amides are not close models for the interaction of the peptide group with water. The ESF value for *N*-methylacetamide differs by  $-4$  kcal/mol from the value for an interior alanine peptide group. An analysis by quantum mechanics of the charge distribution of the peptide group compared with amides (27) predicts that at least a blocked dipeptide is needed to provide a satisfactory model for the peptide bond. Regarding solvation of peptide groups in  $\beta$ -strand peptides, it is necessary to distinguish exterior from interior peptide groups, and the solvation of the central peptide group of a tripeptide differs substantially from the dipeptide value. These results agree with the conclusion (13) that solvent accessibility is a major factor determining the ESF.

An earlier study by Yang and Honig (28) used DELPHI calculations, among other approaches, to examine the determinants of stability in alanine peptide helices. They found that the electrostatic solvation free energy of peptide groups in the unfolded peptide is comparable in size to the enthalpy of the peptide hydrogen bond, and they concluded that hydrogen bonds do not drive helix formation, but rather the burial of nonpolar surface area is the main determinant.

### ESF Values for Alanine Peptides in the Helical Conformation

Results are shown in Fig. 2 for two blocked alanine peptides in the helical conformation: one peptide with five Ala residues



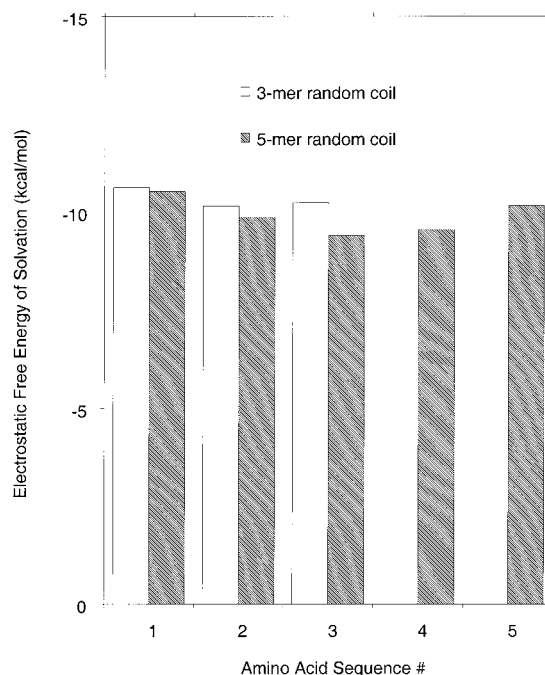
(5-mer) and one with 15 residues (15-mer). The results are more complex than the ones for the  $\beta$ -strand conformation, both because the accessibility of water to peptide groups changes in successive residues at each end of the helix and because hydrogen bonding has a major effect on the electrostatic potential (24, 25). The C=O of Ala-1 is H-bonded to the NH of Ala-5 and the NH groups of Ala residues 1–4 do not make peptide H-bonds, although in the helical geometry specified here the C=O of the N-terminal acetyl group H-bonds to the NH of Ala-4. At the C-terminal end of the 15-mer helix, the NH of Ala-15 is H-bonded to the C=O of Ala-11 and the C=O groups of Ala-12–15 do not make peptide H-bonds, although the C-terminal NHCH<sub>3</sub> group H-bonds to the C=O of Ala-12.

A large drop in the absolute value of the peptide ESF occurs in going from Ala-1 to Ala-5 in the 15-mer helix, accompanying the formation of the first peptide group in which both the C=O and NH groups make peptide H-bonds. A striking feature of the 15-mer helix results is that the ESF values of peptide groups keep on changing after Ala-5, in contrast to the step function behavior predicted by standard H-bond models of helix formation; the changing values result from the long-range nature of the dipole-dipole interactions within the helix (5, 13, 24). A much smaller drop in absolute ESF occurs in the 5-mer between Ala-1 and Ala-5; the 5-mer forms only a single peptide H-bond plus two H-bonds provided by the acetyl and amide blocking groups. Likewise, a large increase in absolute ESF value occurs at the C-terminal end of the 15-mer helix in going from Ala-11 to Ala-15, as the C=O groups cease to be H-bonded. In the interior of the helix, where both the NH and C=O moieties of the peptide groups are H-bonded, the ESF value of a peptide group lies between  $-2.4$  and  $-2.7$  kcal/mol for alanine residues 6–10. Approximately  $-0.5$  kcal/mol partitions with the peptide NH and  $-2.0$  kcal/mol with the peptide C=O (data not shown).

The results in Fig. 2 indicate that interior peptide groups in a helix are significantly solvated, and the free energy is large enough to be highly important in the energetics of protein folding. The ESF of an interior helical peptide group is only about one-half the gas phase energy of the H-bond formed between water and an amide C=O group (22), but there are many peptide groups in a protein, and their solvation introduces a major factor in the energetics of folding. The existence of significant solvation in helical peptide groups is consistent with a molecular dynamics study of the solvation of the hydrogen-bonded formamide dimer (29), which found that the dimer has more affinity for water than for CCl<sub>4</sub> by  $-2.2$  kcal/mol. A high-resolution x-ray structure of a 37-residue peptide containing chiefly alanine residues (9) shows water molecules clustering around peptide C=O groups but no fixed H-bonds between water and C=O groups. Clustering of water molecules near C=O groups is consistent with a molecular dynamics simulation (30) of water molecules surrounding a peptide helix and also with an NMR study (31) of the solvation of a peptide helix. An infrared study of the amide I' and amide II' bands shown by a dimeric coiled-coil helix indicates that solvent-exposed peptide groups are solvated (45).

### ESF Values of Alanine Peptides in the Random Coil Conformation

Results for two blocked alanine peptides in the random coil conformation, containing three or five residues, are shown in Fig. 3. The random coil has been defined previously (13). Structures in the Protein Data Bank provide a library of sequences with specified torsional angles, and residue conformations are chosen at random from this library. There are several differences between these random coil results and the ones for  $\beta$ -strand peptides (Fig. 1). In the  $\beta$ -strand peptides, there is a major difference in ESF between exterior and interior peptide groups, but all interior groups have the same ESF value



**Fig. 3.** The calculated ESF values, in kcal/mol, are plotted against residue number for two blocked alanine peptides in the random coil conformation. The peptides contain either three or five alanine residues. The definition of the random coil conformation is given in the text and in ref. 13, and the method of calculating the results follows ref. 13.

and this remains true as the number of alanine residues is varied. In contrast, the ESF values of the peptide groups in the random coil change steadily with residue position and are different in the three- and five-residue peptides. The decrease in solvation with increasing peptide length is explained by reduced solvent accessibility of the peptide groups in longer peptides (13).

Because the ESF value changes both with residue position and the number of residues in the peptide, it is not possible to give a single value for the ESF of a peptide group in the random coil. The ESF values for the three- and five-residue peptides in the random coil (Fig. 3) are larger in absolute value than that of interior peptide groups in the  $\beta$ -strand conformation ( $-7.9$  kcal/mol) because the random coil contains a mixture of  $\alpha$ ,  $\beta$ , and “other conformations,” and  $\alpha$  has a larger absolute value ESF than  $\beta$  (5, 13, 26, 28, 32, 33), as illustrated here in Figs. 1 and 2. Data for the proportions of  $\alpha$ ,  $\beta$ , and other conformations are given by Smith *et al.* (34) for the various amino acid residues and for two models of the random coil based on structures in the Protein Data Bank. [One model (ALL) includes residues contained within secondary structure, whereas the second model (COIL) excludes them.] In the COIL model, the proportions of  $\beta$ ,  $\alpha$ , and other conformations for the alanine residue are 0.51, 0.32, and 0.17, respectively (34). Thus, the fact that the peptide absolute ESF value is higher in the  $\alpha$  than in the  $\beta$  conformation explains why alanine peptide absolute ESF values are higher in the short random coil peptides shown here than in the  $\beta$ -strand conformation, but the dependence of ESF on peptide length in the random coil, versus the invariance observed in the  $\beta$ -strand, adds an additional twist. The importance of understanding the ESF behavior of the random coil was pointed out by Tobias and Brooks (36) in an early simulation of alanine and valine peptide helix formation.

### Origin of the Favorable Enthalpy Change for Alanine Helix Formation

The problem posed in the Introduction is the existence of a major deficit ( $-7.6$  kcal/mol) in accounting for the enthalpy

change (−1.0 kcal/mol) that drives alanine helix formation, when experimental data for amides (−11.5 kcal/mol) are used to model the enthalpy of interaction between water and the peptide group in the random coil and a quantum mechanics value (−4.9 kcal/mol) is used for the enthalpy of the peptide H-bond in a helix, and other sources of enthalpy are ignored. The deficit is reduced to −1.5 kcal/mol when ESF calculations are used to model the peptide-water interaction in the random coil by the value (−7.9 kcal/mol) for interior peptide groups in  $\beta$ -strand peptides, and a value (−2.5 kcal/mol) is given to the ESF of interior peptide groups in a helix, but calculations for some random coil peptides indicate that ESF values in the random coil are a complex, not yet well understood, subject.

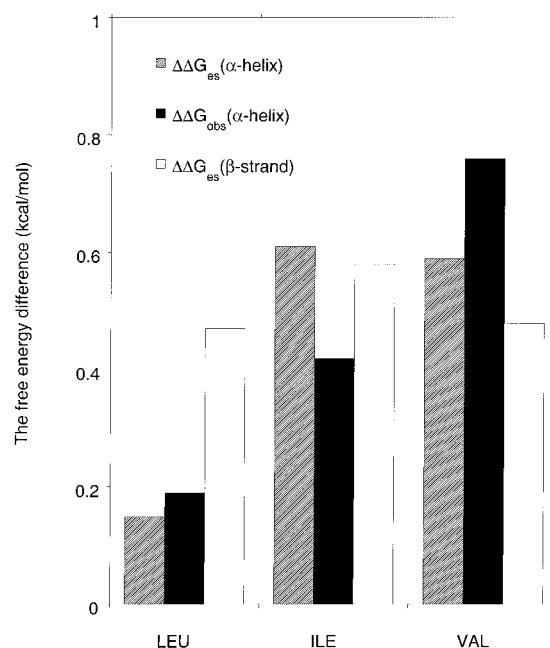
In considering the enthalpy needed to drive helix formation, one term has not yet been discussed: the electrostatic free energy difference between the  $\alpha$  and  $\beta$  conformations arising from main-chain dipole-dipole interactions (5, 13–15). Quantum mechanics calculations indicate that the  $\alpha$  conformation (without peptide H-bonds) is less favorable than  $\beta$  in the gas phase by about 2.5 kcal/mol (32, 33, 35) whereas, in solution, increased solvation of peptide groups in the  $\alpha$  conformation increases the stability of the  $\alpha$  conformation relative to  $\beta$ . The ESF value of an interior peptide group in the  $\beta$  conformation is −7.9 kcal/mol (Fig. 1) whereas in the  $\alpha$  conformation it is −9.6 kcal/mol for the first interior peptide group in a 5-mer (Fig. 2). Because the overall free energy difference between  $\alpha$  and  $\beta$  in solution (without peptide H-bonds) is unfavorable for helix formation, this factor adds to the need for additional enthalpy to drive helix formation.

#### ESF Values of Mutant Peptides in Relation to Helix Propensities

Both theoretical analysis (5, 13, 26) and experimental data (6) have been used to argue that side-chain perturbation of the interaction between water and main-chain polar groups should be a main determinant of helix propensities. Calculation of ESF values for mutant peptides provides a direct test of the size of this perturbation without introducing adjustable parameters. The geometry of the peptide is held fixed in the  $\alpha$  or  $\beta$  conformation, and the size of the perturbation is calculated from the difference between the ESF values of the mutant and an all-Ala peptide.

Fig. 4 gives difference ESF values for peptides in the  $\beta$ -strand conformation, in which a single Leu, Ile, Val, or Gly residue is substituted for Ala. Because the ESF values of interior peptide groups do not vary with peptide length in  $\beta$ -strand peptides, a five-residue peptide is chosen as the host, and the substitution is made at the central residue. The largest change in ESF occurs at the site of substitution, but significant changes also may be seen at the two flanking Ala residues and the difference values are summed over all five residues. Val, Ile, and Leu all show substantial, nearly equal, changes in ESF for the  $\beta$ -strand peptides. [Gly shows a change of opposite sign, corresponding to increased solvation; data not shown.] The changes in ESF values are correlated with decreased accessibility of water to the peptide backbone, caused by the amino acid side chain. For the Val, Ile, and Leu mutant peptides, these changes in ESF resemble the changes in peptide NH exchange rate observed by Bai *et al.* (37) in blocked dipeptides, which have been attributed to steric blocking by the side chain of the access of water to the peptide group.

The mutant peptides in helical conformation have the same sequences as 12-residue peptides studied experimentally (6), and the calculated differences in ESF may be compared directly with the measured differences in helix propensity. The experiments, which were made as a function of temperature, gave results showing that the rank order of helix propensities tends to change with temperature: in some conditions, Val actually has a higher helix propensity than Ala (6). Consequently an enthalpic factor



**Fig. 4.** Comparison between the measured helix propensity difference (6) (filled bars) and the calculated ESF difference for a mutant peptide versus an alanine host in either the  $\beta$ -strand conformation (empty bars) or the helical conformation (hatched bars). The units are kcal/mol. The comparison is made for Leu, Ile, and Val. The  $\beta$ -strand peptides contain five alanine residues plus N-terminal acetyl and C-terminal N-methylamide blocking groups, and only the central Ala residue is substituted by another amino acid. The ESF values of the five central residues of the mutant  $\beta$ -strand peptide are subtracted from those of the Ala peptide and then summed. The peptide sequence used for calculating ESF values in the helical conformation corresponds to the peptide sequence used for measuring helix propensities (6), which is Ac-K(A)<sub>4</sub>X(A)<sub>4</sub>KGY-NH<sub>2</sub>, where X is either A, L, I, V, or G. In calculating ESF differences between helical peptides, the values for the five central peptide groups (groups 4–8) in the mutant peptide are subtracted from the corresponding values for the Ala peptide and then summed. The side-chain torsion angles are: Leu,  $\chi_1$ −60°,  $\chi_2$  180°; Ile,  $\chi_1$ −60°,  $\chi_2$  180°; Val,  $\chi_1$  180°; the backbone angles are given in the legends to Figs. 1 and 2.

(inferred to be the interaction of water with the peptide group) was predicted to be an important determinant of helix propensities. The calculated ESF values for the helical peptides (Fig. 4) are similar to the calculated  $\beta$ -strand results for Ile and Val, but not for Leu. The experimental changes in helix propensity also are shown in Fig. 4. [Glycine is excluded because of a large contribution from backbone conformational entropy to its helix propensity.]

The observed helix propensity differences are comparable in size to the calculated ESF differences both in the helical and  $\beta$ -strand conformations. The helix propensity differences are known to be correlated with the loss in side-chain entropy on helix formation (found, for example, by Monte Carlo simulation for a flexible helix, ref. 38). The free energy changes caused by the side-chain entropy loss correspond to about 1/3 of the helix-propensity differences between alanine and other nonpolar amino acids, as measured in alanine-based peptides (39). The free energy changes arising from calculated ESF differences in the helical conformation are larger than the changes arising from side-chain entropy loss, but the ESF differences in the random coil conformation must be subtracted (13) to compare calculated ESF differences with observed helix propensity differences. The main conclusion from the calculated ESF differences shown here for mutant peptides is that they are comparable in size, before the random coil term is taken into account, to the observed helix

propensity differences. Note that no adjustable parameters are used in calculating the ESF differences besides the peptide geometry and use of the PARSE parameter set.

Even larger ESF differences in the random coil conformation than shown here for the helix have been found for randomly generated tripeptides (13) in which both the sequences and torsion angles are selected randomly, using a model similar to the ALL model of Smith *et al.* (34) (and also using a different set of partial charges than the PARSE set used here). These calculations have been repeated for blocked tripeptides (AXA) and pentapeptides (AAXAA) where X = A, L, I, V, or G (F.A., unpublished work), by using the COIL library of torsion angles (34) and the PARSE parameter set. The results are similar to the ones given in ref. 13, as regards both the large ESF differences between mutant and all-Ala peptides in the random coil conformation and the decrease in ESF values with increasing peptide length. Srinivasan and Rose (40) have argued recently that the random coil is not random but tends to have segments containing either the  $\alpha$  or  $\beta$  conformation.

The calculations given here for mutant peptides suggest an answer to a long-standing puzzle: why are the helix propensity differences measured in alanine-based peptides approximately twice as large as those measured in the natural sequence peptide from RNase T1 or in the RNase T1 protein (41)? The difference between the ESF value of alanine ( $-7.94$  kcal/mol), at an interior peptide group in a  $\beta$ -strand peptide, and valine is approximately twice as large in an all-valine eight-residue peptide (whose ESF is  $-7.42$  kcal/mol, data not shown) as in a single-site mutant, whose ESF is  $-7.73$  kcal/mol. The reason is that neighboring valine residues contribute to the desolvation in an all-valine peptide, whereas neighboring residues have little effect on a valine residue in an alanine-based peptide. Likewise,

in a peptide helix the side chains of neighboring residues larger than Ala are likely to contribute to the desolvation of the peptide group at the site of a substitution (6); thus, the amount of desolvation produced by a particular substitution such as Ala $\rightarrow$ Val will be larger in an alanine-based peptide than in a natural sequence peptide because the peptide group at the site of substitution in the natural sequence peptide already has reduced solvation before the substitution is made.

### Concluding Comment

A basic feature of these ESF calculations is that no adjustable parameters are used besides peptide geometry and the PARSE parameter set; the results are not scaled but instead are compared directly with experiment. The model compounds used to fix the parameters in the PARSE set do not include peptides, and some changes in the parameters are likely to be made when experimental results for peptides become available. Nevertheless, the main conclusions from this study probably will remain. The conclusions are: (i) solvation of the peptide group in the helix is large enough to be important, both in the enthalpy of helix formation and the free energy of protein folding, (ii) amides are poor models for the solvation of the peptide group in the  $\beta$ -strand conformation, and (iii) side-chain interference with solvation of the peptide group is a major factor affecting helix propensities.

We thank Charles Brooks III, Bill DeGrado, David Eisenberg, Bertrand Garcia-Moreno, Pehr Harbury, Barry Honig, Nick Pace, Carol Rohl, George Rose, John Schellman, and Stephen White for discussion and Janet Thornton for use of NACCESS, and we especially thank B. K. Lee for his suggestions concerning preparation of this paper. This work was supported by the Ministry of Science and Technology of Slovenia.

- Baker, E. N. & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol.* **44**, 97–179.
- Ben-Naim, A. (1991) *J. Phys. Chem.* **95**, 1437–1444.
- Yang, A. S., Sharp, K. A. & Honig, B. (1992) *J. Mol. Biol.* **227**, 889–900.
- Honig, B. & Yang, A. S. (1995) *Adv. Protein Chem.* **46**, 27–58.
- Avbelj, F. & Moulton, J. (1995) *Biochemistry* **34**, 755–764.
- Luo, P. & Baldwin, R. L. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4930–4935.
- Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M. & Bolen, D. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2854–2858.
- Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M. & Baldwin, R. L. (1991) *Biopolymers* **31**, 1463–1470.
- Sicheri, F. & Yang, D. S. C. (1995) *Nature (London)* **375**, 427–431.
- Mitchell, J. B. O. & Price, S. L. (1990) *J. Comp. Chem.* **11**, 1217–1233.
- Sitkoff, D., Sharp, K. A. & Honig, B. (1994) *J. Phys. Chem.* **98**, 1978–1988.
- Florian, J. & Warshel, A. (1997) *J. Phys. Chem. B* **101**, 5583–5595.
- Avbelj, F. (2000) *J. Mol. Biol.* **300**, 1335–1359.
- Ben-Naim, A. (1978) *J. Phys. Chem.* **82**, 792–803.
- Lee, B. K. (1991) *Biopolymers* **31**, 993–1008.
- Lee, B. K. (1995) *Methods Enzymol.* **259**, 555–557.
- Wolfenden, R. (1978) *Biochemistry* **17**, 201–204.
- Wolfenden, R., Andersson, L., Cullis, P. M. & Southgate, C. C. (1981) *Biochemistry* **20**, 849–855.
- Della Gatta, G. D., Barone, G. & Elia, V. (1986) *J. Solution Chem.* **15**, 157–167.
- Tannor, D. J., Marten, B., Murphy, R., Friesner, R. A., Sitkoff, D., Nicholls, A., Ringnalda, M., Goddard, W. A. & Honig, B. (1994) *J. Am. Chem. Soc.* **116**, 11875–11882.
- Makhatadze, G. I. & Privalov, P. L. (1993) *J. Mol. Biol.* **232**, 639–659.
- Mitchell, J. B. O. & Price, S. L. (1991) *Chem. Phys. Lett.* **180**, 517–523.
- Dunitz, J. D. (1994) *Science* **264**, 670.
- Brant, D. A. & Flory, P. J. (1965) *J. Am. Chem. Soc.* **87**, 663–664.
- Avbelj, F. & Moulton, J. (1995) *Proteins Struct. Funct. Genet.* **23**, 129–141.
- Avbelj, F. & Fele, L. (1998) *J. Mol. Biol.* **279**, 665–684.
- Price, S. L., Faerman, C. H. & Murray, C. W. (1991) *J. Comp. Chem.* **12**, 1187–1191.
- Yang, A.-S. & Honig, B. (1995) *J. Mol. Biol.* **252**, 351–365.
- Sneddon, S. F., Tobias, D. J. & Brooks, C. L., III (1989) *J. Mol. Biol.* **209**, 817–820.
- Soman, K. V., Karimi, A. & Case, D. A. (1991) *Biopolymers* **31**, 1351–1361.
- Brüschweiler, R., Morikis, D. & Wright, P. E. (1995) *J. Biomol. NMR* **5**, 353–356.
- Head-Gordon, T., Head-Gordon, M., Frisch, M. J., Brooks, C. L., III & Pople, J. A. (1991) *J. Am. Chem. Soc.* **113**, 5989–5997.
- Beachy, M. D., Chasman, D., Murphy, R. B., Halgren, T. A. & Friesner, R. A. (1997) *J. Am. Chem. Soc.* **119**, 5908–5920.
- Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M. & Dobson, C. M. (1996) *J. Mol. Biol.* **255**, 494–506.
- Brooks, C. L., III & Case, D. A. (1993) *Chem. Rev.* **93**, 2487–2502.
- Tobias, D. J. & Brooks, C. L., III (1991) *Biochemistry* **30**, 6059–6070.
- Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993) *Proteins Struct. Funct. Genet.* **17**, 75–86.
- Creamer, T. P. & Rose, G. D. (1994) *Proteins Struct. Funct. Genet.* **19**, 85–97.
- Rohl, C. A., Chakrabarty, A. & Baldwin, R. L. (1996) *Protein Sci.* **5**, 2623–2637.
- Srinivasan, R. & Rose, G. D. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 14258–14263.
- Myers, Y. K., Pace, C. N. & Scholtz, J. M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 2833–2837.
- Cabani, S., Gianni, P., Mollica, V. & Lepori, L. (1981) *J. Solution Chem.* **10**, 563–595.
- Ben-Naim, A. & Marcus, Y. (1984) *J. Chem. Phys.* **81**, 2016–2027.
- Choithia, C. (1976) *J. Mol. Biol.* **55**, 379–400.
- Manas, E.S., Getahun, Z., Wright, W. W., DeGrado, W. F. & Vanderkooi, J. M. (2000) *J. Am. Chem. Soc.*, in press.