3. J. D. Ferry, *Viscoelastic Properties of Polymers* (Wiley, New York, ed. 3, 1980); W. W. Graessley, in *Molecular Conformation and Dynamics of Macromolecules in Condensed Systems*, M. Nagasawa, Ed. (Elsevier, Amsterdam, 1988), p. 163.
4. P. J. Flory, *Principles of Polymer Chemistry* (Cornell Univ. Press, Ithaca, NY, 1953).
5. P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell Univ. Press, Ithaca, NY, 1979).
6. S. F. Edwards, *Proc. Phys. Soc. (London)* **92**, 9 (1967); *Polymer* **9**, 140 (1977).
7. P. G. de Gennes, *J. Chem. Phys.* **55**, 572 (1971).
8. J. Klein, *Nature* **271**, 143 (1978).
9. ———, *Contemp. Phys.* **20**, 611 (1979); W. W. Graessley, *Adv. Polym. Sci.* **47**, 67 (1982); M. Tirrell, *Rubber Chem. Technol.* **57**, 523 (1984); J. Klein, in *Encyclopedia of Polymer Science and Engineering* (Wiley, New York, ed. 2, 1987), p. 205; D. Pearson, *Rubber Chem Technol.* **60**, 439 (1987); F. Brochard-Wyart, in *Fundamentals of Adhesion*, T. D. Lee, Ed. (Plenum, New York, 1988), p. 243; K. Binder and H. Sillescu, in *Encyclopedia of Polymer Science and Engineering* (Wiley, New York, ed. 2, 1989), p. 297; H. H. Kausch and M. Tirrell, *Annu Rev. Mater. Sci.* **19**, 341 (1989).
10. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford Univ. Press, Oxford, 1986).
11. L. S. Lerman and H. L. Frisch, *Biopolymers* **21**, 995 (1982); P. G. de Gennes, *C. R. Acad. Sci. Ser. II* **294**, 827 (1982); see also O. J. Lumpkin and B. H. Zimm, *Biopolymers* **21**, 2315 (1982).
12. J. Klein, *Macromolecules* **11**, 852 (1978); M. Daoud and P. G. de Gennes, *J. Polym. Sci. Polym. Phys.* **17**, 1971 (1979); P. F. Green, P. J. Mills, C. J. Palmstrom, J. W. Mayer, E. J. Kramer, *Phys. Rev. Lett.* **53**, 2145 (1984).
13. D. R. Paul and S. Newman, Eds., *Polymer Blends* (Academic Press, New York, 1978); O. Olabisi, L. M. Robeson, M. T. Shaw, Eds., *Polymer-Polymer Miscibility* (Academic Press, New York, 1979).
14. F. Bueche, W. M. Cashin, P. Debye, *J. Chem. Phys.* **20**, 1956 (1952).
15. M. Doi and S. F. Edwards, *J. Chem. Soc. Faraday Trans. II* **74**, 1789 (1978); *ibid.*, p. 1802; *ibid.*, p. 1818.
16. J. Crank, *Mathematics of Diffusion* (Oxford Univ. Press, Oxford, ed. 2, 1975).
17. J. Klein and B. J. Briscoe, *Proc. R. Soc. London Ser. A* **365**, 53 (1979).
18. F. P. Price, P. T. Gillmore, E. L. Thomas, R. L. Laurence, *J. Polym. Sci. Polym Symp.* **63**, 33 (1978); *Macromolecules* **13**, 880 (1980). [See also A. Y. Chalykli *et al*, *Polym. Sci. U.S.S.R.* **21**, 1835 (1980); *ibid.*, p. 2579.]
19. R. A. L. Jones, J. Klein, A. M. Donald, *Nature* **321**, 161 (1986).
20. G. Coulon *et al.*, *Macromolecules* **22**, 2581 (1989).
21. T. P. Russel, A. Karim, A. Mansour, G. P. Felcher, *ibid.* **21**, 1890 (1988); M. L. Fernandez *et al.*, *Polymer* **29**, 1923 (1988).
22. G. Reiter, S. Hüttenbach, M. Foster, M. Stamm, *Macromolecules*, in press.
23. E. J. Kramer, P. F. Green, C. J. Palmstrom, *Polymer* **25**, 473 (1984).
24. U. K. Chaturvedi *et al.*, *Appl. Phys. Lett.* **56**, 1228 (1990).
25. P. J. Mills *et al.*, *ibid.* **45**, 957 (1984); J. Sokolov *et al.*, *ibid* **54**, 590 (1989).
26. E. D. Kirkendall, *Trans Am. Inst. Min. Metall Eng.* **147**, 104 (1942); L. S. Darken, *ibid.* **175**, 184 (1948).
27. F. Brochard-Wyart, J. Jouffray, P. Levinson, *Macromolecules* **16**, 1638 (1983).
28. H. Sillescu, *Makromol. Chem. Rapid Commun.* **5**, 519 (1984).
29. F. Brochard-Wyart and P. G. de Gennes, *Europhys. Lett.* **1**, 221 (1986).
30. R. J. Composto, E. J. Kramer, D. M. White, *Nature* **328**, 234 (1987).
31. E. A. Jordan *et al.*, *Macromolecules* **21**, 235 (1988).
32. F. Brochard-Wyart, in *Molecular Conformation and Dynamics of Macromolecules in Condensed Systems*, M. Nagasawa, Ed. (Elsevier, Amsterdam, 1988), p. 249.
33. A. D. Buckingham and H. G. Hentschel, *J. Polym Sci. Polym Phys.* **18**, 853 (1980); F. S. Bates, G. D. Wignall, W. C. Koehler, *Phys Rev. Lett.* **55**, 2425 (1985); *ibid.* **57**, 1429 (1986).
34. For such polymer pairs ($\chi < 0$), the phase diagram is frequently characterized by a lower critical solution temperature (*13*), although in this review I make reference only to the usual case of an upper critical temperature.
35. P. G. de Gennes, *J Chem. Phys.* **72**, 4756 (1980).
36. P. F. Green and B. L. Doyle, *Phys. Rev. Lett.* **57**, 2407 (1986); *Macromolecules* **20**, 2471 (1987).
37. G. Gee, *Contemp. Phys.* **11**, 313 (1970).
38. R. A. L. Jones, thesis, Cambridge University, Cambridge (1987).
39. R. J. Composto *et al*, *Phys. Rev Lett.* **57**, 1312 (1986).
40. J. Kanetakis and G. Fytas, *J. Chem. Phys.* **87**, 5048 (1987).
41. P. Pincus, *ibid.* **75**, 1996 (1981); K. Binder, *ibid.* **79**, 6387 (1983).
42. A homogeneous polymer mixture taken from the one-phase to the two-phase region will demix spontaneously (spinodally) into coexisting phases (*4, 5, 41*); this type of diffusive process is related to the present section but will not be considered separately.
43. E. Helfand and A. M. Sapse, *J Chem. Phys.* **62**, 1327 (1975); L. Leibler, *Macromolecules* **15**, 1283 (1982).
44. U. K. Chaturvedi *et al.*, *Phys. Rev. Lett.* **63**, 616 (1989).
45. T. Hashimoto, M. Shibayama, H. Kawai, *Macromolecules* **13**, 1237 (1980).
46. U. Steiner, G. Krausch, G. Schatz, J. Klein, *Phys Rev. Lett.* **64**, 1119 (1990).
47. P. G. de Gennes, *C. R. Acad. Sci. Ser. II* **308**, 13 (1989); J. L. Harden, *J. Phys France* **51**, 1777 (1990).
48. R. Gelman, D. C. Poppke, K. A. Piez, *J. Biol. Chem.* **255**, 8098 (1980).
49. R. A. L. Jones *et al*, in preparation; S. Rocca and J. Klein, unpublished data.
50. I thank Z. Alexandrowicz, E. Katzir, Y. Rabin, S. Reich, G. Schatz, A. Silberberg, and U. Steiner for useful comments on the manuscript, and M. Eisenstein and U. Steiner for help with the figures. This work was supported by the German-Israel Foundation, the U.S.–Israel Binational Science Foundation, and the Minerva Foundation. The author is the Herman Mark Professor of Polymer Physics at the Weizmann Institute.

Research Articles

A Thermodynamic Scale for the Helix-Forming Tendencies of the Commonly Occurring Amino Acids

KARYN T. O'NEIL AND WILLIAM F. DEGRADO

Amino acids have distinct conformational preferences that influence the stabilities of protein secondary and tertiary structures. The relative thermodynamic stabilities of each of the 20 commonly occurring amino acids in the α-helical versus random coil states have been determined through the design of a peptide that forms a noncovalent α-helical dimer, which is in equilibrium with a randomly coiled monomeric state. The α helices in the dimer contain a single solvent-exposed site that is surrounded by small, neutral amino acid side chains. Each of the commonly occurring amino acids was substituted into this guest site, and the resulting equilibrium constants for the monomer-dimer equilibrium were determined to provide a list of free energy difference (ΔΔG°) values.

HOW AN AMINO ACID SEQUENCE DICTATES THE THREE-dimensional structure of a protein is an intriguing, but largely unsolved question. While a general solution to the protein folding problem is not yet available, considerable progress has been made in understanding the factors stabilizing α helices. Each amino acid has distinct conformational preferences that lead to stabilization or destabilization of an α helix (*1*). Electrostatic interactions between charged side chains and either the helical

dipole or another charged residue located one turn of helix away are also important for helix stabilization (*2*). In addition, aromatic (*3*), hydrogen-bonded (*4*), and hydrophobic interactions (*5*) appear to stabilize α helices. However, our understanding of these features is largely qualitative, and the relative importance of each of these interactions is not yet clear. For instance, if one wished to design a stable helical protein, would it be better to maximize the number of electrostatic interactions or the helical potential (for example, as assessed by statistical methods) (*1*)? A classic way to resolve this question would be to individually determine the free energies associated with each of these interactions. One could then design a sequence with consideration to the multiple forces involved in stabilization of a specific conformation.

Experimental studies aimed at deciphering the factors that stabilize helices have focused largely on monomeric helices in aqueous solution. Scheraga and co-workers have studied copolymers of polyhydroxyalkylglutamine containing a low mole fraction of randomly incorporated amino acids (*6*). The temperature dependences of helicity for the copolymers were analyzed according to the Zimm-Bragg formalism (*7*) to provide estimates of the thermodynamic parameters, σ and *s*, which are believed to relate to the ease of forming and propagating a helix, respectively. However, the applicability of these parameters to natural proteins has recently been questioned by Baldwin and co-workers, who have shown that peptides as short as 15 residues can show substantial helical content in water (*2, 3, 8*). Thus, numerous recent studies (*2, 3, 8–11*) have focused on short, synthetic peptides as models for monomeric helix formation and have contributed to the realization that intrinsic conformational preferences, electrostatic interactions, and aromatic interactions are important determinants of helix stability.

However, potential limitations are associated with the study of monomeric helices as models for helix formation in proteins. Helix formation is not a simple two-state process; thus, analysis of the data requires models with multiple parameters that are not necessarily independent of one another (*6*). Also, the results obtained from the study of monomeric helices may not be representative of helices in folded proteins; an α helix in a protein lies in a distinctly anisotropic environment with non-uniform solvent exposure and dielectric constant. Finally, there is little direct evidence that monomeric helices serve as intermediates in protein folding. For instance, stabilization of the NH_2-terminal helix of ribonuclease A occurs near the end of the folding process after the β-sheet portions of the protein have folded (*12*). In cytochrome c, helices form early in the folding process, but they appear to be stabilized through a covalent cross-link and interhelical hydrophobic interactions (*13*). Thus, it became necessary to confirm and extend the conclusions derived from studies of monomeric helices with a system more closely resembling the folded state of a protein.

Site-directed mutagenesis of proteins is an alternative approach for determining helix-stabilizing interactions. Mutagenesis studies have shown that barnase (*14*) and T4 lysozyme (*15*) are stabilized by electrostatic and hydrogen-bonded interactions between side chains and the exposed amides on the ends of their α helices. In principle, this method might also be applied to obtain a thermodynamic scale for the relative helical preferences of the individual amino acids by substituting all possible amino acids at a given position in an α helix of a natural protein. In practice, however, the data would be difficult to interpret because helices in proteins are typically involved in multiple hydrogen-bonded, van der Waals, and electrostatic interactions, and mutation of a given residue would change not only the conformational preference but also other interactions. To circumvent some of the limitations associated with the study of natural proteins, we adopted a "minimalist" approach (*16*) to design a model for helical proteins that contained a single site into which various amino acids could be substituted. To minimize all but the most locally determined interactions, small, neutral side chains were included at the positions surrounding this site. We have used this model system to obtain a complete thermodynamic scale for the helix-forming tendencies of the 20 naturally occurring amino acids.

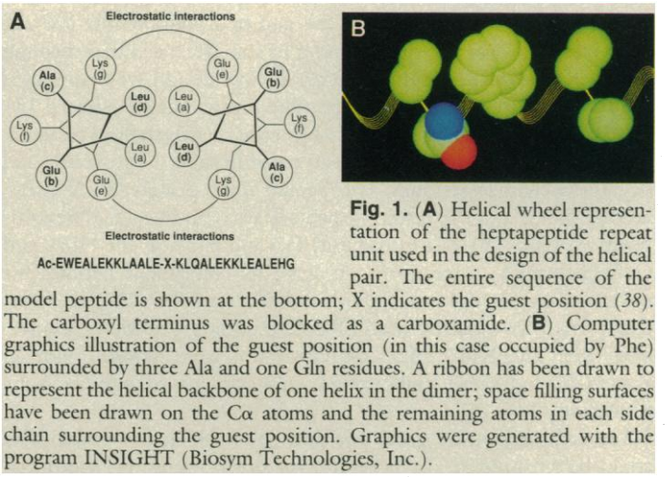

Fig. 1. (**A**) Helical wheel representation of the heptapeptide repeat unit used in the design of the helical pair. The entire sequence of the model peptide is shown at the bottom; X indicates the guest position (*38*). The carboxyl terminus was blocked as a carboxamide. (**B**) Computer graphics illustration of the guest position (in this case occupied by Phe) surrounded by three Ala and one Gln residues. A ribbon has been drawn to represent the helical backbone of one helix in the dimer; space filling surfaces have been drawn on the Cα atoms and the remaining atoms in each side chain surrounding the guest position. Graphics were generated with the program INSIGHT (Biosym Technologies, Inc.).

Design of an α-helical dimer. We have chosen a noncovalent homodimer of α helices as a simple model for a helical protein. Previously, Kim and co-workers showed that "leucine zipper" peptides form parallel, α-helical dimers that are presumably short α-helical coiled coils such as those found in fibrous proteins (*17*). Such a structure would have a number of advantages as a model system for studying helix stability: (i) CD measurements of "leucine zipper" peptides indicate that they exist in an equilibrium between nonhelical monomers and α-helical dimers (*17*). Thus, helix formation is linked to a two-state thermodynamic process (monomer-dimer equilibrium) with an association constant K_a (related to the free energy for dimerization $\Delta G°$ by $\Delta G° = -RT \ln K_a$, where R is the gas constant and T is the absolute temperature) that can be easily measured by CD spectroscopy. (ii) The midpoint for the denaturation curves of homodimers depends on the peptide concentration, which provides a distinct advantage when comparing variants of differing stabilities. The stability of monomeric proteins is typically determined by analyzing temperature- or denaturant-induced unfolding transitions (*18*). In cases where variants unfold at very different temperatures or concentrations of denaturant, data must be extrapolated to a common condition to allow a comparison of their relative stabilities. Such extrapolations often add considerable uncertainty to the results (*18*). In contrast, the midpoint of a monomer-dimer equilibrium depends on the concentration of the peptide. Thus, by varying the peptide concentration, variants of differing stabilities can be examined at the same concentration of denaturant and temperature, thereby eliminating the need to extrapolate to a common solution condition. (iii) The small size and homodimeric structure of the peptides allow small differences in stability (<0.1 kcal/mol) to be measured with a high degree of accuracy.

Our design of a parallel α-helical dimer contains a number of elements important for the stability of coiled coils in fibrous proteins. Hodges and co-workers have designed a series of repeating polyheptapeptide models for tropomyosin (*19*). Their most successful heptapeptide repeat ($\underset{a}{Leu}$-$\underset{b}{Glu}$-$\underset{c}{Ala}$-$\underset{d}{Leu}$-$\underset{e}{Glu}$-$\underset{f}{Gly}$-$\underset{g}{Lys}$) contains hydrophobic Leu residues at the a and d positions (Fig. 1A) to stabilize

The authors are members of the scientific staff in the Central Research and Development Department, E. I. du Pont de Nemours and Company, Wilmington, DE 19880–0328.

the structure through van der Waals and hydrophobic interactions, and Glu and Lys at positions e and g to stabilize the dimer through interhelical electrostatic interactions. In our design we maintained the basic features of their heptapeptide repeat but substituted Lys for Gly (at the f position). This change enhances the helical potential (*1*) and provides a structure with an equal number of positively and negatively charged residues. This heptad was repeated approximately four times, and the resulting peptide was modified at numerous positions (Fig. 1B). We introduced Trp and His near the NH_2- and COOH-termini, respectively, to serve as convenient handles for nuclear magnetic resonance (NMR) and ultraviolet (UV) spectroscopy; Trp is particularly useful for determining peptide concentration (*20*). The NH_2- and COOH-termini were blocked with an acetyl-Glu and a Gly-carboxamide, respectively, to stabilize helix formation (*2, 4, 21*).

We next created a "guest position" into which any amino acid could be substituted. To be useful for determining the conformational preferences of the amino acids, this site should be distant from the dimerization surface. In this way, differences in the overall stability of the dimer should be attributable to changes in the helix coil transition, and not the interhelical packing interactions. Therefore, the guest position was situated at the f position of the middle heptad (Fig. 1). Also, the side chain at the guest position should undergo a minimal change in solvent accessibility upon forming a helix, and therefore residues close to this position in the helix (residues at $i - 4$, $i - 3$, $i + 3$, and $i + 4$) should be small, neutral, and nonperturbing. Thus Ala was chosen for these positions with the exception of position $i + 3$, where the neutral, polar residue Gln was included to improve peptide solubility.

A set of 21 peptides was synthesized in which each of the commonly occurring amino acids or α-aminoisobutyric acid (Aib, a commonly used helix-promoting amino acid) (*22*) was incorporated at the guest position. The peptides were synthesized by the rapid multiple peptide synthesis method (*23*) modified to use the fast-Boc protocol of Kent (*24*), which is rapid, minimizes the use of solvent, and provides reasonably homogeneous peptides. This method allowed the synthesis of ten different peptides per week. Each peptide was synthesized on 200 mg of 4-methylbenzhydrylamine resin (substitution level = 0.7 mmol/g), providing 20 to 50 mg of peptide after purification to homogeneity by reversed-phase high-performance liquid chromatography (HPLC). The composition and purity of the peptides were confirmed by HPLC, fast atom bombardment mass spectrometry, and amino acid analysis.

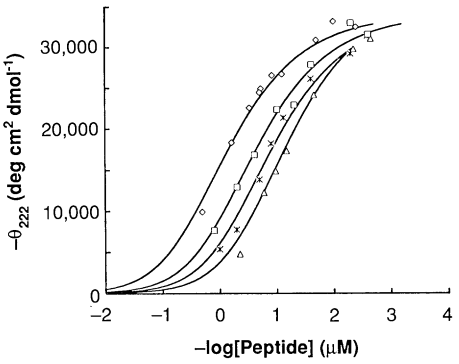

Fig. 3. Concentration dependence of helix formation for the peptides with the following amino acids substituted in the guest position: A la (△); Phe (✳); Ile (□); and Asp (◇). Solutions of peptides were prepared in 10 mM MOPS, 5 M urea, pH 7.5, and θ_{222} was measured in cells of various pathlengths (0.01 cm, 0.1 cm, 0.5 cm, and 1 cm) depending on peptide concentration. Duplicate measurements were reproducible to within ±2%. Data were fit to a monomerdimer equilibrium [θ_{222}(monomer) = 0, θ_{222}(dimer) = −34,000]. Theoretical curves are drawn in solid lines by using the equilibrium constant obtained from the fit for each peptide.

Characterization of the α-helical dimers. The circular dichroism (CD) spectra of the peptides at micromolar concentrations showed minima at 208 and 222 nm and a maximum at 190 nm, consistent with a helical conformation (*25*). The ellipticities at 222 nm (θ_{222}) of the peptides were −34,000 ± 2,000 deg cm^2 dmol^{-1}, corresponding to approximately 100% helical content (*25*). Analytical ultracentrifugation (*26*) of the Phe peptide (the peptide with Phe in the guest position) at 800 μM concentration gave a molecular weight of 6,200 ± 250, in good agreement with that expected for a dimer. Dimerization of each of the 21 peptides was confirmed by analyzing the concentration dependence of their CD spectra in the presence of urea (see below).

The stabilities of the helical dimers were rapidly assessed by measuring the variation in θ_{222} as a function of urea concentration. Sigmoidal curves were observed (Fig. 2), with midpoints that depended markedly on the nature of the "guest" amino acid (Table 1). Substitution of amino acids that occur frequently in helices (*1*) into the guest position gave peptides that were most resistant to denaturation; for instance, the midpoints of the Ala peptide and the Gly peptide are separated by greater than 2 M urea. The urea denaturation curves were analyzed with the use of a two-state model (*27*), so that the dissociation constants for the monomer-dimer equilibrium could be calculated as a function of urea concentration. Within the limit of experimental error, ΔG° scaled linearly with urea concentration, and parallel lines were obtained for each of the peptides (Fig. 2). The ΔG° values were extrapolated to 0 M urea concentration for each peptide (Table 1). Although the linear extrapolation is long, dimerization is highly favorable in each case. Thus the perturbation to the stability associated with each amino acid substitution was minor relative to the overall folding free energy, and it is reasonable to assume that all of the peptides adopted similar, dimeric structures. One exception was the Pro peptide, whose extrapolated stability was less favorable than the others by about 5 to 7 kcal/mol, although θ_{222} for the dimer of this peptide (about −32,000 deg cm^2 dmol^{-1} as measured at a concentration of 2.4 mM) was not significantly different from the other peptides.

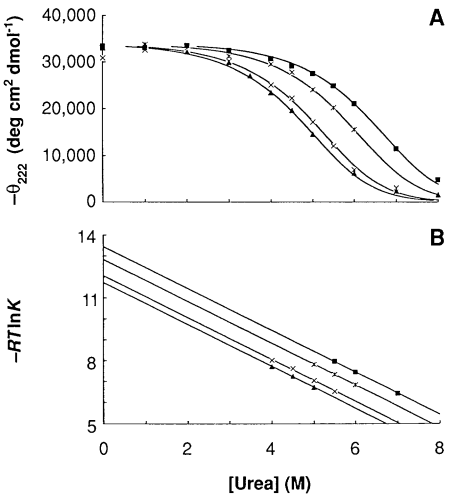

Fig. 2. (**A**) Urea denaturations for the peptides with the following residues substituted in the guest position: Ala (■); Met (✳); Phe (x); and Asp (▲). The θ_{222} values were measured at various urea concentrations in a 0.5-cm cell with peptide concentration approximately 10 μM in MOPS buffer, pH 7.5. Peptide concentration was determined spectrophotometrically (*20*) and by amino acid analysis. Measurements were made on a Jasco J-500 Spectropolarimeter at 23° ± 2°C. (**B**) ΔG° versus urea concentration; ΔG° was calculated with a two-state model assuming a random coil monomer ($\theta_{222} = 0$) and folded dimer (θ_{222} = −34,000 deg cm^2 dmol^{-1}). Extrapolation by linear regression to 0 M urea was made with data from (A) where the fraction of dimer was between 0.75 and 0.25.

Determination of ΔΔG° for helix formation. Typically, the free energy difference between two variants of a given protein (ΔΔG°) is determined from an analysis of their denaturation curves. However, the error associated with this analysis (~1 kcal/mol) (*18*) is large when compared to the effects we expected to observe in this study. Therefore the dimerization constant for each peptide was deter-

mined by measuring the dependence of θ_{222} on the peptide concentration at a fixed urea concentration of 5.0 M (Fig. 3). In each case the data were well described by a monomer-dimer equilibrium, so that $\Delta G°$ could be determined from K_a to within ±0.1 kcal/mol (corresponding to a 20% error in the K_a values) based on duplicate measurements of some of the peptides. The differences in the free energies of helix stabilization (designated $\Delta\Delta G_\alpha$, Table 2) for each amino acid relative to Gly were calculated by subtracting $\Delta G°$ for the appropriately substituted peptide from $\Delta G°$ for the Gly peptide, and have been divided by 2 to correct for the number of monomers per dimer. Although the values of $\Delta\Delta G_\alpha$ were calculated from data obtained in 5 M urea, the identity of the slopes for $\Delta G°$ plotted as a function of the urea concentration (Fig. 2) for the entire set of peptides would indicate that the values of $\Delta\Delta G_\alpha$ should be approximately independent of the urea concentration. The value for Pro could not be measured directly by this method because the Pro peptide failed to show appreciable dimerization at 5 M urea. This value was approximated by linear extrapolation of the denaturation curve for the Pro peptide to 5.0 M urea.

Correspondence to other scales. Recently, several groups have examined the helix-forming tendencies of various amino acids by making single or multiple substitutions in monomeric, helix-forming peptides (*8–11*). Kallenbach and co-workers (*11*) have examined ten neutral amino acids (Ala, Leu, Met, Gln, Ser, Ile, Val, Thr, Asn, and Gly), and determined their free energies for helix formation by the Zimm-Bragg multistate model (*7*). These values were expressed as free energies relative to Gly and hence may be directly compared with our $\Delta\Delta G_\alpha$ values. The two measures are linearly related (Fig. 4A) with a slope near unity (0.78 ± 0.09). Thus, the two scales agree on the relative abilities of these residues to stabilize helices and the magnitude of the effect is approximately equal in both systems. The largest outlier in this analysis is Ser, which may show a cooperative helix destabilizing effect when it is repeated in a sequence as in the Kallenbach study (*11*). The good overall agreement between the two studies encourages one to believe that the values obtained may indeed be measures of intrinsic conformational preferences and are largely independent of the specific features of the host peptide.

The agreement between the two systems also supports our assumption that variations in $\Delta\Delta G_\alpha$ primarily reflect differences in helix stability as opposed to helix-helix packing. By comparison, while the overall rank order between $\Delta\Delta G_\alpha$ and the s values obtained by the host guest method of Sheraga (*6*) show some similarities (Table 2), there is a less precise correspondence than that observed for the Kallenbach data.

Merutka and Stellwagen (*9*) have also examined water-soluble monomeric helices and find Ser and Met to be less helix-stabilizing than Ala by 0.5 and 0.3 kcal/mol, respectively (*28*). These values are in excellent agreement with the values found in this work (0.4 and 0.3, respectively). However, their data were analyzed by a two-state model, which is approximate for monomeric helix formation (*7, 11*). Baldwin and co-workers have published a similar set of experiments, investigating the helix-stabilizing tendencies of five residues (*8*). They find a rank order for helix-stabilizing tendencies of Ala = Leu > Ile > Phe > Val, similar but not identical to that determined in our work, that is, Ala > Leu > Phe > Ile > Val.

Table 1. Approximate $\Delta G°$ values (kilocalories per mole) for the peptides extrapolated to 0 M urea; the amino acid listed in the first column indicates the residue in the guest position (see legend to Fig. 2B).

Amino acid	$-\Delta G°$	Amino acid	$-\Delta G°$	Amino acid	$-\Delta G°$
Aib	14.3	Trp	12.4	Val	11.9
Ala	13.4	Ser	12.2	Asp	11.7
Arg	13.3	Glu	12.2	Thr	11.7
Lys	13.0	Tyr	12.1	His	11.6
Leu	13.0	Phe	12.1	Cys	11.6
Met	12.8	Ile	12.0	Gly	11.3
Gln	12.8	Asn	12.0	Pro	6.5

The $\Delta\Delta G_\alpha$ values from our work also correlate modestly well with the relative probability (P_α) that a given type of residue will occur in a helical conformation in the crystal structures of proteins. A plot of P_α(mid) (*29*) as a function of $\Delta\Delta G_\alpha$ is shown in Fig. 4B for each of the amino acids with the exception of Pro and Cys, Pro because its helix-forming probability varies considerably between different statistical scales and Cys because P_α values do not differentiate between the free sulfhydryl and disulfide-bonded forms of this residue. The data show a modest linear relation (Fig. 4B) with a correlation coefficient of 0.75. The largest outliers in this analysis are Ser, Tyr, and Gly, which may reflect the frequent inclusion of these residues in turns in globular proteins (*1*). This modest correlation supports previous suggestions that P_α is primarily a measure of a residue's intrinsic conformational preferences (*30*).

The complete list of $\Delta\Delta G_\alpha$ values provides a powerful tool for protein design and can be used to predict the change in conformational free energy associated with mutating a helical residue in a protein. Other effects including hydrophobicity and electrostatic interactions must also be considered, but might be small if the changes involve neutral side chains and occur on the solvent-exposed surface near the center of a helix. For example, substitution of Ala for Gly^{46} or Gly^{48} in helix 3 of lambda repressor increased the stability of the protein by 0.66 ± 0.12 and 0.87 ± 0.13 kcal/mol, respectively (*31*), within experimental error of the value expected from this work (0.77 kcal/mol; Table 2). Mutation of Val^{131}, near the COOH-terminus of a helix in T4 lysozyme, to either Ala or Thr results in a 0.23 kcal/mol stabilization or 0.08 kcal/mol destabiliza-

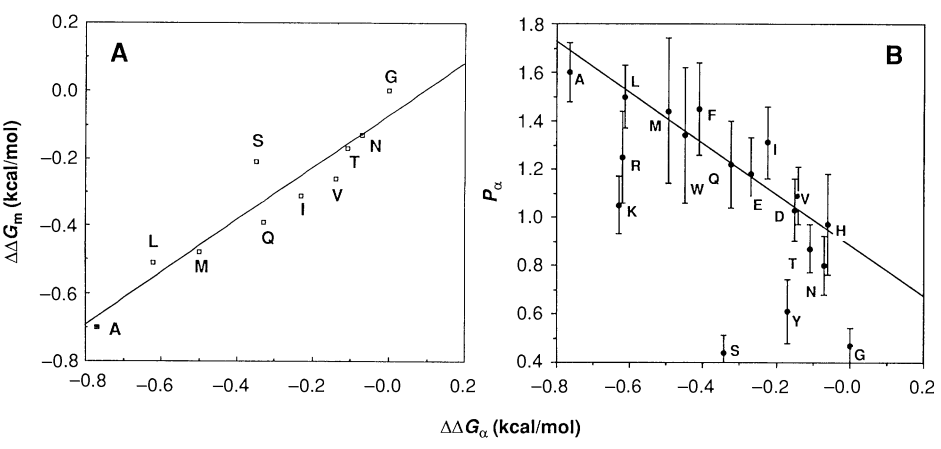

Fig. 4. (**A**) Correlation between $\Delta\Delta G_\alpha$ and $\Delta\Delta G_m$ (*11*). Individual points are identified by the single letter abbreviation for each amino acid. The line was obtained by linear regression of the data. (**B**) Correlation between $\Delta\Delta G_\alpha$ and P_α (*29*). The error bars represent the standard deviation of P_α (*29*); Cys and Pro were not included in this correlation, and Ser, Tyr, and Gly were not included in calculating the regression line.

tion, respectively (*32*). While these results agree with our results in terms of their rank order, the magnitude of the effect for the Val to Ala mutation is about one half that predicted from the differences in the corresponding values of $\Delta\Delta G_\alpha$. The explanation for this discrepancy may lie in the determination of $\Delta\Delta G_\alpha$ values for residues near the center of helices and in the nearness of the mutated residue to the COOH-terminus of the helix.

This compilation of $\Delta\Delta G_\alpha$ values should also aid in interpreting the results of site-directed mutagenesis even when the substitutions do not occur on the solvent-exposed faces of helices. For instance, the role of the hydrophobic effect in protein folding is frequently probed by mutating a large, hydrophobic side chain to a smaller, hydrophobic residue (*33*). A problem with this approach, however, is that the substitutions affect the conformational preferences in addition to hydrophobic and van der Waals interactions. Now, however, if the mutations are made near the center of a helix, the magnitude of the change in the helical potential may be estimated from the appropriate $\Delta\Delta G_\alpha$ values and subtracted from the net change in free energy.

How large of a driving force are local conformational preferences in determining which portions of a protein chain adopt helical versus nonhelical conformations? We calculate the difference between the average value of $\Delta\Delta G_\alpha$ for sequences in helical versus nonhelical conformations in protein crystal structures (*34*) is small (0.2 kcal/mol) compared to the free energies associated with burial of a large hydrophobic side chain (1 to 4 kcal/mol) (*33, 35*). However, only about one third of the residues in a helix contribute to the packing of the apolar core, whereas every residue in a helix contributes its conformational preference. Nevertheless, even after appropriate consideration to conformational preferences, it appears that hydrophobic forces are thermodynamically more important for driving protein folding, supporting previous work showing that hydrophobic interactions are of primary importance (*5, 35*). This result also explains why secondary structure prediction algorithms are only modestly successful (*1*).

What are the factors leading to differences in $\Delta\Delta G_\alpha$ for the amino acids? Pro is known to induce a kink in helices (*36*); evidently such a kink is not energetically favorable in our model helical pair. Also, Gly is known to be helix-destabilizing because it possesses considerable conformational flexibility. Ala is far more helix-favoring because of the presence of a methyl side chain, which restricts the conformational space considerably, thereby decreasing the conformational entropy of the unfolded state. Another helix-stabilizing residue is Aib that, by virtue of its geminal dimethyl substituents, is conformationally restricted to adopt torsion angles close to those of an α helix (*22*). Interestingly, we find that Aib is about as helix stabilizing as Ala. Perhaps Aib is not more helix stabilizing because it is achiral and therefore can adopt either a left-handed or right-handed helical conformation. Dimerization requires formation of a right-handed helix and hence occurs with a concomitant decrease in conformational entropy.

Ala is the most helix favoring of the 20 commonly occurring amino acids, which suggests that its methyl side chain is sufficient to reduce the conformational freedom of the main chain in the unfolded state. Larger side chains may have a similar effect, but their own conformational space is limited in the α-helical state (*37*). Thus the side chains of residues like Leu, Met, and Phe likely experience an unfavorable change in conformational entropy upon forming a helix. Further destabilization of the helical conformation may occur for residues whose side chains are branched at the β-position (Val, Thr, and Ile) because of a potential steric clash between a substituent at the γ-position in the side chain and a carbonyl oxygen in the preceding turn of the α helix, which causes the side chain to adopt a less than optimal torsional angle (*32*). Less easy to explain are the energetic differences among some of the polar and charged side chains. These residues may form weak hydrogen-bonded and electrostatic interactions with the backbone atoms in certain nonhelical conformations, thereby biasing the equilibrium towards the unfolded state. We hope that our compilation of $\Delta\Delta G_\alpha$ values will serve as a stimulus for addressing these and related questions.

Table 2. Helix formation parameters for each of the 20 naturally occurring amino acids and Aib. The $\Delta G°$ values at 5 M urea were calculated from the relation $\Delta G° = -RT\ln K_a$ with the value of K_a determined from the concentration dependence of dimerization for each peptide (Fig. 3). The $\Delta\Delta G_\alpha$ values were calculated by subtracting $\Delta G°$ for each peptide from $\Delta G°$ for the Gly-peptide. In contrast to the neutral amino acids, the values of $\Delta\Delta G_\alpha$ for the charged amino acids (*) were somewhat dependent on ionic strength and varied up to 0.3 kcal/mol between 0 and 1 M NaCl. The values cited refer to the values obtained directly in 1.0 M NaCl or extrapolated to 1.0 M NaCl. The values for P_{mid} are from (*4*), for P_α are from (*30*), and for s are from (*38*).

Amino acid	$\Delta\Delta G_\alpha$ (kcal/mol)	P_{mid}	P_α	s
Ala	−0.77	1.8	1.60	1.07
Aib	−0.69			
Arg	−0.68*	1.3	1.25	1.03
Lys	−0.65*	1.1	1.05	0.94
Leu	−0.62	1.2	1.50	1.14
Met	−0.50	1.5	1.44	1.20
Trp	−0.45	1.5	1.34	1.11
Phe	−0.41	1.3	1.45	1.09
Ser	−0.35	0.6	0.44	0.76
Gln	−0.33	1.3	1.22	0.98
Glu	−0.27*	0.8	1.18	0.97
Cys	−0.23	0.7	0.66	0.99
Ile	−0.23	1.2	1.31	1.14
Tyr	−0.17	0.8	0.61	1.02
Asp	−0.15*	1.0	1.03	0.68
Val	−0.14	1.2	1.09	0.95
Thr	−0.11	1.0	0.87	0.82
Asn	−0.07	0.9	0.80	0.78
His	−0.06*	1.0	0.97	0.69
Gly	0.00	0.5	0.47	0.59
Pro	~3	0.3	0.19	0.19

REFERENCES AND NOTES

1. P. Y. Chou and G. D. Fasman, *Adv. Enzymol.* **47**, 45 (1978); G. D. Fasman, in *Prediction of Protein Structure and the Principles of Protein Conformation*, G. D. Fasman, Ed. (Plenum, New York, 1989), p. 193; P. Prevelige, Jr., and G. D. Fasman, *ibid* , p. 391; J. Garnier and B. Robson, *ibid* , p. 417.
2. S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8898 (1987).
3. K. R. Shoemaker *et al* , *Biopolymers* **29**, 1 (1990).
4. J. S. Richardson and D. C. Richardson, *Science* **240**, 1648 (1988).
5. W. F. DeGrado and J. D. Lear, *J. Am. Chem Soc.* **107**, 7684 (1985).
6. H. Scheraga, *Pure Appl. Chem.* **50**, 315 (1978); M. Sueki *et al* , *Macromolecules* **17**, 148 (1984); B. H. Vasquez, M. R. Pincus, H. A. Scheraga, *Biopolymers* **26**, 351 (1987); K.-H. Altmann, J. Wojcik, M. Vasquez, H. A. Scheraga, *ibid.* **30**, 107 (1990).
7. B. H. Zimm and W. K. Bragg, *J. Chem. Phys.* **31**, 526 (1984).
8. S. Padmanabhan *et al.*, *Nature* **344**, 268 (1990).
9. G. Merutka and E. Stellwagen, *Biochemistry* **29**, 894 (1990).
10. P. C. Lyu, L. A. Marky, N. R. Kallenbach, *J. Am. Chem Soc* **111**, 2733 (1989).
11. ———, *Science* **250**, 669 (1990).
12. J. B. Udgaonkar and R. L. Baldwin, *Nature* **335**, 694 (1988); D. N. Brems and R. L. Baldwin, *J. Biol. Chem.* **180**, 1141 (1984); A. M. Labhardt *et al.*, *Biochemistry* **22**, 321 (1983); A. M. Labhardt, *Proc. Natl Acad. Sci U.S.A.* **81**, 7674 (1984).
13. H. Roder, G. A. Elove, S. W. Englander, *Nature* **335**, 700 (1988).
14. L. Serrano and A. R. Fersht, *ibid.* **342**, 296 (1989).
15. B. W. Matthews, *Biochemistry* **336**, 651 (1988).
16. W. F. DeGrado, Z. R. Wasserman, J. D. Lear, *Science* **243**, 622 (1989).
17. E. K. O'Shea, R. Rutkowski, P. S. Kim, *ibid.*, p. 538; E. K. O'Shea, R. Rutkowski, W. F. Stafford, III, P. S. Kim, *ibid.* **245**, 646 (1989); C. Cohen and D. A. D. Parry, *Proteins* **7**, 1 (1990).
18. C. N. Pace, *Trends Biochem. Sci.* **15**, 14 (1990); W. J. Becktel and J. A. Schellman, *Biopolymers* **26**, 1859 (1987); J. A. Schellman, *Annu Rev. Biophys. Biophys. Chem.* **16**, 115 (1987); D. W. Bolen and M. M. Santoro, *Biochemistry* **27**, 8069 (1988).

19. S. Y. M. Lau *et al.*, *J. Biol. Chem* **259**, 13253 (1984); R. S. Hodges *et al.*, **256**, 1214 (1981).
20. H. Edelhoch, *Biochemistry* **6**, 1948 (1967).
21. L. G. Presta and G. D. Rose, *Science* **240**, 1591 (1988).
22. B. V. Prasad and P. Balaram, *CRC Crit. Rev Biochem* **16**, 307 (1989); C. Toniolo and E. Benedetti, *ISI Atlas of Science*, p. 225 (1988).
23. H. R. Wolfe and R. Wilke, *Peptide Res.* **2**, 352 (1989).
24. J. Schneider and S. B. H. Kent, *Cell* **54**, 363 (1988); W. F. DeGrado and J. D. Lear, *Biopolymers* **29**, 205 (1990).
25. Y.-H. Chen, J. T. Yang, K. H. Chau, *Biochemistry* **13**, 3350 (1974); N. Greenfield and G. D. Fasman, *ibid.* **8**, 4108 (1969).
26. R. J. Pollet, B. A. Haase, M. L. Standaert, *J Biol. Chem* **254**, 30 (1979).
27. C. N. Pace, *Methods Enzymol.* **131**, 266 (1986); J. U. Bowie and R. T. Sauer, *Biochemistry* **28**, 7139 (1989).
28. These free energy differences refer to the values expected at 300°C, based on an examination of the van't Hoff plots in (*9*).
29. R. W. Williams, A. Chang, D. Juretíc, S. Loughran, *Biochem. Biophys. Acta* **916**, 200 (1987).
30. P. Y. Chou and G. D. Fasman, *J. Mol. Biol* **74**, 263 (1973); *Biochemistry* **13**, 211 (1974).
31. M. H. Hecht, J. M. Sturtevant, R. T. Sauer, *Proteins* **1**, 43 (1986).
32. S. Dao-Pin, W. A. Baase, B. W. Matthews, *ibid.* 7, 198 (1990).
33. M. Matsumura, M. Becktel, B. W. Matthews, *J. Biol Chem* **264**, 16059 (1989); J. T. Kellis, Jr., K. Nyberg, A. R. Fersht, *Biochemistry* **28**, 4914 (1989).
34. The difference between $\Delta\Delta G_\alpha$ for residues in the center of the helices versus in nonhelical conformations was calculated from the equation $\gamma = \Sigma[f_\alpha(i)(\Delta\Delta G_\alpha)] - \Sigma[f_{n\alpha}(i)(\Delta\Delta G_\alpha)]$, where Σ refers to the sum over all 20 naturally occurring amino acids, and $f_\alpha(i)$ and $f_{n\alpha}(i)$ refer to the probability of finding a given type of amino acid in the center of a helix, or a nonhelical conformation, respectively. The quantities $f_\alpha(i)$ and $f_{n\alpha}(i)$ for all 20 amino acids were calculated from the data compiled in (*4*).
35. J. U. Bowie, J. F. Reidhaar-Olson, W. A. Lim, R. T. Sauer, *Science* **247**, 1306 (1990); R. Kauzmann, *Adv Prot. Chem.* **14**, 1 (1959).
36. L. Piela, G. N. Némethy, H. A. Scheraga, *Biopolymers* **26**, 1587 (1987).
37. ———, *ibid*, p. 1273; M. J. McGregor, S. A. Islam, M. J. E. Sternberg, *J Mol Biol* **198**, 295 (1987); J. Janin, S. Wodak, M. Levitt, B. Maigret, *ibid.* **125**, 357 (1978).
38. J. Wojcik, K.-H. Altmann, H. A. Scheraga, *Biopolymers* **30**, 12 (1990).
39. We gratefully acknowledge J. Bryson for his preliminary work on this project. We thank H. A. Scheraga for sharing unpublished data, R. Houghton for peptide synthesis, T. Devine for peptide purification, L. Janvier for amino acid analysis, and J. Lazaar for mass spectrometry.

1 May 1990; accepted 5 September 1990

Protein Splicing Converts the Yeast *TFP1* Gene Product to the 69-kD Subunit of the Vacuolar H^+-Adenosine Triphosphatase

PATRICIA M. KANE, CARL T. YAMASHIRO, DAVID F. WOLCZYK, NORMA NEFF, MARK GOEBL, TOM H. STEVENS*

The *TFP1* gene of the yeast *Saccharomyces cerevisiae* encodes two proteins: the 69-kilodalton (kD) catalytic subunit of the vacuolar proton-translocating adenosine triphosphatase (H^+-ATPase) and a 50-kD protein. The 69-kD subunit is encoded by the 5′ and 3′ thirds of the *TFP1* coding region, whereas the 50-kD protein is encoded by the central third. Evidence is presented that both the 69-kD and 50-kD proteins are obtained from a single translation product that is cleaved to release the 50-kD protein and spliced to form the 69-kD subunit.

VACUOLAR-TYPE PROTON-TRANSLOCATING ADENOSINE TRIphosphatases (H^+-ATPases), which acidify certain intracellular compartments in eukaryotic cells (*1, 2*), have been purified from various sources and show an overall structural similarity (*1*). All of the enzymes purified are multisubunit complexes containing at least two peripheral membrane subunits with molecular masses of ~70 and 60 kD and at least one integral membrane subunit of ~15 kD. The 70-kD subunit contains the catalytic adenosine triphosphate (ATP) binding site. The vacuolar H^+-ATPase of the yeast *Saccharomyces cerevisiae* consists of eight different subunits, including a 69-kD catalytic subunit and 60- and 17-kD subunits similar to those found in other cells (*3, 4*). Genes encoding the 70- and 60-kD subunits have been cloned and characterized from a number of different species, and the sequences of both subunits from plant, animal, and fungal sources are remarkably conserved (*5–7*). The predicted amino acid sequence of the yeast 60-kD subunit is 82 percent identical to that of the *Neurospora crassa* 57-kD subunit and 74 percent identical to that of the human 60-kD subunit (*8*). The functional roles of the yeast vacuolar H^+-ATPase have been assessed by disrupting the gene for the 60-kD subunit, which is termed *VAT2* (*8*). Cells lacking *VAT2* grow more slowly than wild-type cells, fail to grow at neutral pH, and lack the ability to acidify their vacuoles. Isolated vacuoles from these cells lack ATPase activity.

We now present evidence that the 69-kD subunit of the yeast vacuolar H^+-ATPase is one of two proteins encoded by the previously identified *TFP1* gene (*9*). The vacuolar H^+-ATPase subunit is specified by the two ends of *TFP1*, and the central one-third encodes a 50-kD "spacer" protein. Our results indicate that the 69- and 50-kD proteins are formed from a single translation product by post-translational cleavage and splicing.

Encoding of the catalytic subunit of the yeast vacuolar H^+-ATPase by *TFP1*. A dominant allele of *TFP1* was previously isolated from a mutant yeast strain resistant to the drug trifluopera-

P. M. Kane, C. T. Yamashiro, D. F. Wolczyk, and T. H. Stevens are at the Institute of Molecular Biology, University of Oregon, Eugene, OR 97403. P. M. Kane is now in the Department of Chemistry, College of William and Mary, Williamsburg, VA 23185 and C. T. Yamashiro is in the Department of Biological Sciences, Stanford University, Stanford, CA 94305. N. Neff is at the Memorial Sloan-Kettering Cancer Center, New York, NY 10021. M. Goebl was in the Department of Genetics, University of Washington, Seattle, WA 98195, but is now in the Department of Biochemistry, Indiana University School of Medicine, Indianapolis, IN 46202.

*To whom correspondence should be addressed.