Helix Propensities Are Identical in Proteins and Peptides†

Jeffrey K. Myers,‡ C. Nick Pace,* and J. Martin Scholtz*

Department of Medical Biochemistry and Genetics, Department of Biochemistry and Biophysics, and Center for Macromolecular Design, Texas A&M University, College Station, Texas 77843

Received March 27, 1997; Revised Manuscript Received June 20, 1997⊗

ABSTRACT: Our understanding of the factors stabilizing α-helical structure has been greatly enhanced by the study of model α-helical peptides. However, the relationship of these results to the folding of helices in intact proteins is not well characterized. Helix propensities measured in model peptides are not in good agreement with those from proteins. In order to address these questions, we have measured helix propensities in the α-helix of ribonuclease T_1 and a helical peptide of identical sequence. We have previously demonstrated excellent agreement between peptide and protein for the nonpolar amino acids [Myers, J. K., Pace, C. N., and Scholtz, J. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 2833−2837]. Most other amino acids also show good agreement, although certain polar amino acids are exceptions. Helix propensities measured in the ribonuclease T_1 peptide/protein are compared with those measured in other systems. Reasonable agreement is found between most systems; however, our propensities differ substantially from those measured in several model peptide systems. Alanine-based peptides overestimate the propensity differences by a factor of 2, and host/guest experiments underestimate them by a factor of 2−3.

Understanding the folding of globular proteins is a major undertaking in modern biological science. A knowledge of the various molecular forces stabilizing native structures is crucial to our understanding of this process. With the development of site-directed mutagenesis, it became possible to measure the contribution of individual residues to the conformational stability of proteins. This mutational approach has provided a wealth of information on the contributions of various interactions to protein stability (*1, 2*). Mutagenesis has also been used to examine the factors contributing to the stability of secondary structure. α-Helices have proven to be of great interest, due to their common occurrence in proteins. Statistical analyses of the residues present in α-helices of proteins have shown preferences for certain residues to be located in helices, and in certain positions in helices (*3, 4*). Site-directed mutagenesis has been used to mutate residues in α-helices and observe the resulting effects on the stability and structure of the protein (*5*−*11*). Several of these studies have addressed the intrinsic preferences of the commonly occurring amino acids to be in a helical conformation (termed "helix propensity"). The consensus of these studies is that alanine is the best helix former, proline is the worst, glycine is second-worst, and that the propensity differences of the other amino acids fall within a narrow range of about 1 kcal/mol between alanine and glycine.

Short, monomeric helices composed mostly of alanine exhibit significant helix formation in water (*12*) and have been used as models for protein folding and stability (*13, 14*). Intrinsic helical propensities of the amino acids have been measured using these peptides as hosts (*15*−*17*). Another model system is the salt-bridge stabilized peptides used by Kallenbach and co-workers (*18, 19*). These peptides consist of glutamic acid and lysine residues spaced such that they form ion pairs in the helical conformation. A notable feature of these peptide investigations is the very poor *quantitative* agreement among the propensities measured in the two systems. Additionally, none of the peptide systems appear to be in particularly good agreement with the protein studies mentioned above. For example, the propensity differences from alanine-based peptides are approximately twice as large as those measured in proteins (*14, 20*). To understand the basis for this disagreement, we have measured helix propensity in a protein and a monomeric peptide containing the same helical sequence.

The mutation glycine 23 to alanine in ribonuclease T_1 (RNase T_1) causes a peptide corresponding to the α-helix in the protein (residues 13−29) to show significant helix formation in aqueous solution (*21, 22*). This allows a direct comparison of helix propensity in peptides and proteins with identical sequence. RNase T_1 is a small (104 residue), monomeric protein which has proven to be a useful model for the study of protein folding and stability (*23*). RNase T_1 is an $\alpha + \beta$ class protein with several strands of β-sheet packed against a relatively long (4.5 turn) α-helix, forming a hydrophobic core. RNase T_1 is a well-characterized protein, exhibiting a reversible two-state unfolding mechanism when unfolded using high temperatures or chemical denaturants such as urea or guanidine hydrochloride (*23*).

RNase T_1 appears to have an ideal site at which to measure helix propensities: alanine 21 is in the exact center of the helix, on the solvent exposed face, and the side chains of residues (i, $i + 3$) and (i, $i + 4$), which could interact with

† This work was supported by the National Institutes of Health (Grant GM52483 to J.M.S., GM37039 to C.N.P., and the predoctoral training Grant T32 GM08523 to J.K.M.) and the Robert A. Welch Foundation (A-1281 to J.M.S. and A-1060 to C.N.P.). C.N.P. is also supported by the Tom and Jean McMullin Professorship and J.M.S. is an American Cancer Society Junior Faculty Research Awardee (JFRA-577).

* Address correspondence to either author at the Department of Medical Biochemistry and Genetics, 440 Reynolds Medical Building, Texas A&M University, College Station, TX, 77843-1114. E-mail: pace@bioch.tamu.edu or jm-scholtz@tamu.edu. Phone: (409) 845-1788 (Pace) or (409) 845-0828 (Scholtz). Fax: (409)845-9481.

‡ Current address: Department of Biochemistry, Duke University Medical Center, Durham, NC, 27710.

⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1997.

S0006-2960(97)00718-6 CCC: $14.00

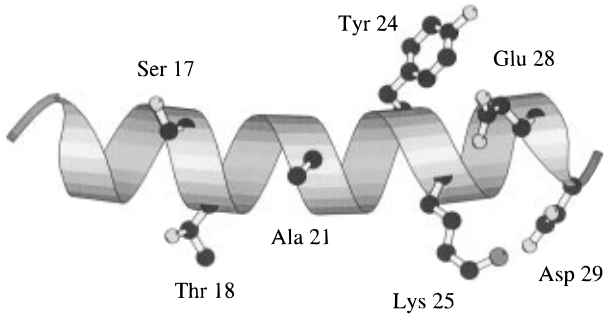

FIGURE 1: α-Helix from the X-ray crystal structure of RNase T_1 (*52*). α-Helix spans residues 13−29 in the protein. The side chain of alanine 21, our substitution site, is shown as a ball and stick. Other relevant side chains are also shown. The ribbon drawing was made using MOLSCRIPT (*53*).

residues at position 21, all appear to be involved in other interactions (Figure 1). No nonhelix residues appear to be close enough to interact with side chains at position 21. All 19 of the other common amino acids were substituted at position 21 using site-directed mutagenesis. Additionally, 20 peptides corresponding to the helical sequences of the 20 mutant proteins were synthesized. Changes in stability of the 20 proteins and helical peptides were measured and compared. Results for the nonpolar amino acids have been reported previously (*22*).

EXPERIMENTAL PROCEDURES

Mutagenesis. The mutant glycine 23 to alanine (G23A) was constructed by the polymerase chain reaction (PCR) single mutagenic primer technique (*24*). The mutant RNase T_1 gene was sequenced to confirm the presence of the mutation and the absence of any extraneous mutations. For the series of mutants at position 21, a modification of the procedure was used. Instead of mutagenic primers of known sequence, a set of primers with codon 21 randomized was used in the PCR to give a chance of all 20 amino acids. The G23A plasmid was used as a template, so that all mutants at position 21 also contained the G23A mutation. The clones were sequenced to identify the mutants at position 21. Twelve of the 20 mutants were made this way; the remaining seven were made with primers of known sequence, like the G23A mutant.

Protein Expression and Purification. The mutant proteins were expressed and purified as described previously (*25*) with the following modification. The release of the protein from the cells was performed in 15 mM Tris + 3 mM EDTA buffer, pH 7.4, instead of at pH 2.4. The combined supernatants were then loaded onto a DEAE-cellulose (Sigma) anion-exchange column. This column was pre-equilibrated in Tris + EDTA buffer (pH 7.4). After loading, the column was washed with ca. 300 mL of this same buffer. The protein was eluted from the column with a gradient of 0 to 0.6 M NaCl (1 L total volume) in Tris + EDTA buffer. Active fractions were pooled and lyophilized. Final purification was performed on a G50 size-exclusion column. Protein purity was confirmed using SDS−polyacrylamide gel electrophoresis.

Urea Denaturation Curves. The stability of the mutant proteins was determined using urea denaturation. Samples were prepared containing 0.01 mg/mL protein in 30 mM glycine, pH 2.5, or 30 mM MOPS buffer, pH 7, along with various concentrations of urea (Sigma ultrapure). A urea stock solution was prepared fresh for each curve; the urea concentration was measured by refractive index (*26*). These tubes were incubated at 25 °C for at least 16 h before measurements were taken. This was sufficient time for each sample to reach equilibrium. The intrinsic fluorescence of each sample (kept at 25.0 ± 0.1 °C by a circulating water bath and stirred using a magnetic stirring apparatus) was measured by exciting at 278 nm and monitoring emission at 320 nm for a period of 1 min in a SLM AB2 fluorescence spectrometer (Aminco-Bowman).

Analysis of the denaturation curves was performed using a two-state unfolding model and the linear extrapolation method (*26*). These were combined into a single equation to describe the shape of the denaturation curve (*27*, *28*):

$$Y = \frac{Y_f + m_f[\text{urea}] + (Y_u + m_u[\text{urea}])e^{-m(D_{1/2}+[\text{urea}])/RT}}{1 + e^{-m(D_{1/2}+[\text{urea}])/RT}} \quad (1)$$

where Y is the observed fluorescence (after subtracting out the intrinsic fluorescence of buffer and urea), m_f and Y_f are the slope and intercept, respectively, of the pretransition baseline, m_u and Y_u are the slope and intercept, respectively, of the post-transition baseline, m is the dependence of free energy of unfolding on urea concentration, and $D_{1/2}$ is the midpoint of the denaturation curve. The free energy of unfolding in the absence of denaturant (defined as the conformational stability of the protein) is the product of m and $D_{1/2}$. The experimental curves were fit to the above equation using Origin data analysis software (Microcal). Denaturation curves were performed at least twice for each mutant.

Peptide Synthesis. The 20 A21X peptides were synthesized using solid-phase Fmoc/OPfp activated ester chemistry (*29*). Peptides were synthesized with acetylated N-termini and amidated C-termini. Lyophilized peptides were dissolved in water and purified on a Whatman C_{18} reverse-phase fast-performance liquid chromatography (FPLC) column on a Pharmacia FPLC system. A smaller reverse-phase C_{18} analytical FPLC column (Pharmacia) was used to confirm peptide purity. Peptide identity was confirmed using MALDI time-of-flight mass spectrometry.

Circular Dichroism. Helicity of the peptides was measured using circular dichroism (CD) on an Aviv model 62DS CD spectropolarimeter. Samples contained 30 μM peptide in CD buffer (1 mM each of potassium phosphate, borate, and citrate) in a 0.5 cm path length cuvette, maintained at 0 °C by a built-in temperature controlling unit. The cysteine-containing peptide was measured in CD buffer containing 10 mM DTT to inhibit intermolecular disulfide formation. Peptide stock solutions were made in water and the peptide concentration was determined by UV absorbance, using extinction coefficients at 276 nm of 1390 M^{-1} cm^{-1} for tyrosine and 5455 M^{-1} cm^{-1} for tryptophan (*30*). The CD signal at 222 nm was monitored until it reached a plateau and averaged for 100 s. CD spectra of the peptides were recorded in a 0.1 cm cuvette in CD buffer.

Raw CD signal (in millidegrees), after subtracting the blank signal of cuvette and buffer, was converted to mean residue ellipticity (in degrees squared centimeters per decimole) using the equation

$$[\theta] = \frac{100(\text{signal})}{Cnl} \quad (2)$$

where C is the peptide concentration in millimolarity, $n = 17$ (the number of residues in the peptide), and l is the pathlength in centimeters. This ellipticity ($[\theta]_{obs}$) was converted to fraction helix values using

$$F_{helix} = \frac{[\theta]_{obs} - [\theta]_{coil}}{[\theta]_{helix} - [\theta]_{coil}} \quad (3)$$

where $[\theta]_{helix}$ and $[\theta]_{coil}$ represent the mean residue ellipticity of a complete helix [$-42500(1 - (3/n))$ where n = number of residues in the peptide] and complete random coil (+640), respectively (*17*).

Helix-Coil Theory. To convert measured helicity into a free energy scale, we used the homopolymer version of the standard Lifson−Roig helix/coil model (*31, 32*). We define the wt* peptide as the "host" peptide and all the other peptides contained guest residues at position 21. We use the single sequence approximation, meaning that only one stretch of helical residues is allowed to exist in any one peptide (one nucleation site per molecule). For short peptides, this is equivalent to the more complex full version of the theory (*33*). The nucleation constant, v^2, was taken to be 0.0023 (*34*). By treating the host peptide as a homopolymer, we can assign one w value to the whole peptide based on the measured helicity. The w value is the helix propagation parameter and represents the equilibrium of a residue between a random coil and a helical conformation. The w parameter is thus the helix propensity for the residue. The w values of the guest residues were determined from the fractional helicity of the respective peptides. Free energy changes can be calculated using $\Delta G = -RT \ln(w/(1 + v))$. Relative changes in ΔG were calculated using $\Delta\Delta G = -RT \ln(w_{guest}/w_{host})$, so that a positive $\Delta\Delta G$ indicates destabilization of the helix.

RESULTS

A peptide of the sequence of the helix in wild-type RNase T_1 (residues 13−29),

S S D V S T A Q A A G Y K L H E D

shows a circular dichroism (CD) spectrum characteristic of a random coil conformation (Figure 2) containing at most only a few percent helix. In order to study helical stability in peptides, more helical structure needs to be present. The sequence of the helix contains a glycine residue at position 23 (protein numbering, underlined in sequence above). Glycine is known to be a poor helix former, while alanine appears to be the best helix former (*16*). A peptide was therefore synthesized with this glycine replaced with alanine. The CD spectrum of this Gly 23 to Ala peptide is characteristic of a mixture of α-helical and random coil structures (Figure 2). Its fractional helicity is 15% at pH 7.0 and 30% at pH 2.5, which is enough structure to serve as a model for the helix in the protein. This helicity shows no dependence on peptide concentration, which suggests that the peptide is monomeric (data not shown). To have a peptide and protein helix with identical sequence, we made the same Gly 23 to Ala substitution in the RNase T_1 protein. The G23A protein and peptide will serve as references in which the position 21 mutations will be compared. Therefore, we will hereafter designate the G23A protein/peptide as wild-type*, or just wt*.

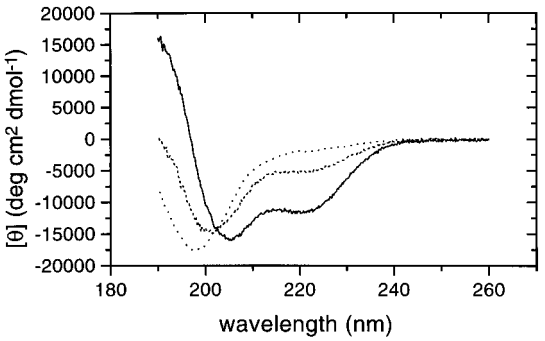

FIGURE 2: Far-UV CD spectra of peptides from the helix of RNase T_1: wild-type, pH 2.5 (dotted line), wild-type* (G23A) at pH 7.0 (dashed line) and pH 2.5 (solid line). All three were measured at 0 °C in CD buffer (see text).

Table 1: Protein Urea Denaturation Data for A21X Mutants

	pH 2.5			pH 7.0		
residue at 21	$D_{1/2}$[a] (M)	m[b] (kcal/mol M)	$\Delta\Delta G$[c] (kcal/mol)	$D_{1/2}$[a] (M)	m[b] (kcal/mol M)	$\Delta\Delta G$[c] (kcal/mol)
A (wt*)	2.78	1.58	0	4.32	1.28	0
C	2.33	1.56	0.74			
D	2.98	1.59	−0.33	3.73	1.24	0.71
E	2.81	1.52	−0.05	3.75	1.14	0.69
F	2.43	1.69	0.57			
G	2.23	1.71	0.90			
H	2.44	1.79	0.56	4.18	1.27	0.17
I	2.51	1.57	0.44			
K	2.47	1.70	0.51			
L	2.70	1.64	0.13			
M	2.69	1.60	0.15			
N	2.99	1.72	−0.34			
Q	2.62	1.60	0.26	3.99	1.12	0.40
R	2.53	1.65	0.41			
S	2.48	1.67	0.49	3.99	1.22	0.40
T	2.43	1.60	0.57			
V	2.38	1.77	0.66			
W	2.60	1.56	0.30			
Y	2.54	1.69	0.39			

[a] Urea concentration at which the protein is half-denatured. Error is estimated to be 0.04 M. [b] Linear dependence of ΔG of unfolding on urea concentration. Error is estimated to be 0.06 kcal/mol M. [c] Difference in free energy of unfolding, calculated using the difference in $D_{1/2}$ times the average m value for all the mutants (1.64 ± 0.08 kcal/mol M at pH 2.5 and 1.21 ± 0.07 kcal/mol M at pH 7.0). Error estimated to be 0.1 kcal/mol. Positive values indicate the mutation is destabilizing.

All 19 other amino acids were substituted into position 21 of the wt* (G23A) protein. All of these mutations except proline resulted in stable, active ribonucleases. The stability parameters determined by urea denaturation are given in Table 1. The proline mutant expressed in extremely low yields and appears to be denatured under ambient conditions. Urea denaturation had little effect on the intrinsic fluorescence of the mutant, and no cooperative transition was seen. Consequently, we could not obtain stability parameters for A21P.

As for the other mutants, Table 1 shows that the stability changes are all within 1 kcal/mol of wt*, which has alanine at position 21. Wild-type* is one of the most stable proteins, while glycine is the least stable; most of the other mutants fall between alanine and glycine. Uncharged aspartic acid and asparagine, two closely related amino acids, are more stable than alanine, as is uncharged glutamic acid.

Table 2: Peptide Helicity and Stability[a]

residue at 21	pH 2.5 $-[\theta]_{222}$[b]	pH 2.5 F_{helix}[c]	pH 2.5 $\Delta\Delta G$[d] (kcal/mol)	pH 7.0 $-[\theta]_{222}$[b]	pH 7.0 F_{helix}[c]	pH 7.0 $\Delta\Delta G$[d] (kcal/mol)
A (wt*)	9900	0.30	0	4700	0.15	0
C	5600	0.17	0.53	2300	0.083	0.51
D	4800	0.15	0.66	1900	0.070	0.68
E	8500	0.26	0.17	3100	0.10	0.31
F	5100	0.16	0.61	2100	0.077	0.59
G	3200	0.11	0.98	1300	0.055	0.95
H	2600	0.09	1.2	1900	0.071	0.67
I	6600	0.20	0.38	3100	0.11	0.29
K	6100	0.19	0.45	3100	0.11	0.30
L	7700	0.24	0.25	3600	0.12	0.19
M	8300	0.25	0.18	4000	0.13	0.12
N	4700	0.15	0.66	2100	0.078	0.58
P	2700	0.09	1.1	1100	0.048	1.1
Q	7200	0.22	0.31	3200	0.11	0.29
R	5400	0.17	0.56	2800	0.097	0.38
S	5700	0.18	0.51	2600	0.091	0.42
T	4500	0.14	0.71	2100	0.076	0.59
V	4800	0.15	0.66	2000	0.075	0.61
W	8600	0.26	0.14	4600	0.15	0.02
Y	6100	0.19	0.45	3100	0.11	0.31

[a] Potassium phosphate, 1 mM; potassium citrate, 1 mM; potassium borate, 1 mM; 0 °C. [b] Mean residue ellipticity at 222 nm, calculated from measured CD signal as described in Experimental Procedures. The units of mean residue ellipticity are degrees centimeter squared per decimole. [c] Fractional helicity (or fraction helix), calculated from mean residue ellipticity as described in Experimental Procedures. [d] Change in free energy of helix formation, determined using Lifson−Roig helix/coil theory as described in Experimental Procedures.

Twenty peptides with all the amino acids at position 21 were synthesized. The helical content of these peptides was calculated from the CD mean residue ellipticity at 222 nm. Ellipticity and fraction helix values are given in Table 2. Fractional helicities of the peptides range from 5 to 15% at pH 7.0 and from 10 to 30% at pH 2.5. Alanine is found to give the most stable helix. Proline is found to be the worst helix former at pH 7.0, although charged histidine is found to be even worse than proline at pH 2.5. Glycine is also found to be a poor helix former. All peptides were approximately twice as helical at low pH, due to the interaction of two acidic residues at the C-terminus of the helix (Figure 1) with the helix macrodipole (*21*). Most amino acids are nearly twice as helical at low pH, meaning that their relative helix propensities for these nonionizable residues do not depend on pH, as expected. Three amino acids, histidine, aspartic acid, and glutamic acid, have different protonation states at pH 7 and 2.5. The propensity of histidine is much lower when the side chain is protonated, which is also observed in alanine-based peptides (*35*). Glutamic acid is a worse helix former at pH 7.0 than pH 2.5; this is also in agreement with studies on alanine-based peptides, showing that the intrinsic propensity of glutamic acid increases upon protonation (*36*). The helix propensity of aspartic acid does not seem to depend on pH, again mirroring observations made in alanine-based peptides (*37*).

By applying helix/coil transition theory (*31, 32*), a free energy scale of the helix forming tendencies of the amino acids was determined from the observed helicity of each peptide (shown in Table 2; the scale is relative to alanine which has been set to zero). The range of free energies between alanine and glycine is around 1 kcal/mol, just like the A21X proteins. A plot of $\Delta\Delta G$ (the first Δ denotes a free energy change relative to another amino acid, in this case alanine) comparing peptide and protein data is shown in Figure 3. Most of the amino acids are near a line with a slope of unity and an intercept of zero, indicating good correspondence between the peptide and protein. The magnitudes of the $\Delta\Delta G$s measured in the two systems are generally in excellent agreement. However, several outliers are noticeable on the plot. Histidine is much more stable in the protein than in the peptide. This effect is present regardless of the protonation state of the side chain. Uncharged aspartic acid and asparagine are also much more stable in the protein than in the peptide. This is curious since uncharged glutamic acid and glutamine show good agreement between peptide and protein. Another interesting result is that charged glutamate is more destabilizing in the protein than in the peptide. These are discussed below.

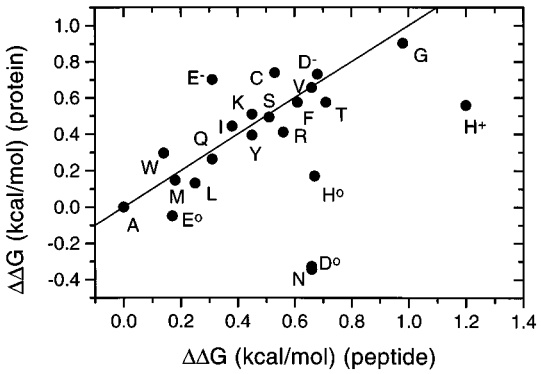

FIGURE 3: Comparison of helix propensity ($\Delta\Delta G$ relative to alanine) measured in the wt* protein and peptide. Single letter amino acid codes identify the residue at position 21. Solid line is $Y = X$.

DISCUSSION

Helix Propensity in Peptides and Proteins. We have measured helical propensities of the common amino acids in a protein and a peptide model with identical helical sequence. We have found that for most amino acids the propensities are very similar in each system. The nonpolar amino acids in particular show very good agreement (reported previously in ref *22*). The correlation coefficient between the peptide and protein data for the seven nonpolar amino acids is 0.98, and a best fit regression line has a slope of 1 and an intercept of 0. The outliers on Figure 3 are all polar residues, namely histidine, the acidic residues, and asparagine. Comparing protein and peptide results for the 16 other residues results in a linear correlation coefficient of 0.92; the slope of the regression line is 0.96, and the intercept is 0.01. Thus, for most of the amino acids, the helix propensities measured in the two systems are in excellent agreement. This shows that measurements of propensity in model helical peptides are, in certain cases, directly applicable to proteins.

The helix propensities of several polar amino acids show significant disagreement between the peptide and protein. Can we understand these outliers? It can be seen from Table 3 that the protein is more stable than the peptide by 0.5−1.0 kcal/mol for histidine, asparagine, and uncharged aspartic acid. This suggests that a specific interaction involving these residues exists which stabilizes the protein more than the peptide. Since charged aspartate does not exhibit this effect, our best guess is that the side chains of these residues are acting as hydrogen bond donors. Serine 17 is in close proximity to position 21 (Figure 1). In the wild-type

Table 3: Comparison of Helix Propensities in Different Systems[a]

	RNase T_1 peptide[b]	RNase T_1 protein[b]	difference[c]	alanine peptides[d]	T4 lysozyme[e]	barnase[f]
Ala	0	0	0	0	0	0
Gly	0.96	0.90	0.06	1.97	0.96	0.91
Val	0.64	0.66	−0.02	1.04	0.33	0.88
Ile	0.34	0.44	−0.10	0.71	0.12	0.81
Leu	0.22	0.13	0.09	0.37	0.02	0.35
Met	0.15	0.15	0.0	0.52	0.10	0.31
Phe	0.60	0.57	0.03	1.0	0.37	0.69
Ser	0.46	0.45	−0.01	0.79	0.43	0.41
Cys	0.52	0.74	−0.22	0.90	0.54	0.82
Thr	0.65	0.57	0.08	1.22	0.42	0.79
Tyr	0.38	0.39	−0.01	0.69	0.72	0.82
Trp	0.08	0.30	−0.22	0.96	0.38	0.84
Lys^+	0.38	0.51	−0.13	0.29	0.23	0.19
Arg^+	0.47	0.41	0.06	0.22	0.19	0.14
Gln	0.30	0.33	−0.03	0.55	0.16	0.48
Glu^0	0.17	−0.05	0.22	0.48	0.43	nd
Glu^-	0.31	0.69	−0.38	0.62	nd	0.55
Asn	0.62	−0.34	0.96	0.96	0.57	0.66
Asp^0	0.66	−0.33	0.99	0.79	0.54	nd
Asp^-	0.68	0.71	−0.03	0.81	nd	0.71
His^0	0.67	0.17	0.50	0.84	nd	nd
His^+	1.20	0.56	0.64	1.11	0.39	0.78
Pro	1.1	nd	nd	>3.8	3.5	nd

[a] In kilocalories per mole, relative to alanine. [b] Data at pH 2.5 and 7 were averaged, with the exception of His, Asp, and Glu. [c] Difference in propensities measured in the wt* protein and wt* peptide. [d] Data from ref 17. [e] Data from ref 8. [f] Data from ref 6.

structure it forms a hydrogen bond with a carbonyl oxygen which points the side chain away from position 21. However, mutagenesis results suggest that this hydrogen bond is quite weak [the stability of the protein increases when this serine is replaced with alanine (*38*)]. Other candidates for interaction are threonine 18, tyrosine 24, and lysine 25. Threonine 18 appears to be farther away from position 21 than serine 17. The side chain of tryosine 24 is mostly buried, so forming a hydrogen bond with side chains at position 21 would require disrupting significant hydrophobic interactions. Lysine 25 is pointing to the C-terminus of the helix, where it interacts with two acidic residues 28 and 29 (Figure 1). Thus, we suggest that serine 17 accepts a hydrogen bond from some side chains at position 21 in the protein. In a peptide, a hydrogen bond to serine 17 would be expected to contribute less to the stability because end-fraying would significantly weaken the interaction.

In the case of charged glutamate, the peptide is more stable than the protein by 0.4 kcal/mol. This suggests that a specific interaction occurs in the peptide that is more important than the same interaction in the protein. Here, our best guess is that the glutamate is interacting with lysine 25. As noted above, in the wild-type structure, the amino group of lysine 25 forms ion pairs with glutamic acid 28 and aspartic acid 29 (Figure 1). In the peptide, these interactions are probably less important due to end-fraying, so the lysine would be availible for a favorable interaction with glutamic acid. In model peptides, it has been shown that the interaction between charged glutamate and lysine contributes ≈0.5 kcal/mol to helix stability (*36*); thus, the increased stability of the glutamate peptide is probably due to an interaction with lysine 25. Of course, only detailed structural information can confirm these speculations.

The A21P mutant is the only protein that has a major change in stability. We were unable to completley characterize the A21P protein, since it expresses in very poor yields. The behavior of A21P is probably due to structural reasons. Proline residues in α-helices are unique in that the unusual structure of the residue causes a pronounced kink in the helix (*39*). While some proteins have evolved with proline in helices, for RNase T_1 the angle created in the helix would probably disrupt the packing of the helix against the rest of the protein. This is likely the cause of the inability of A21P to fold.

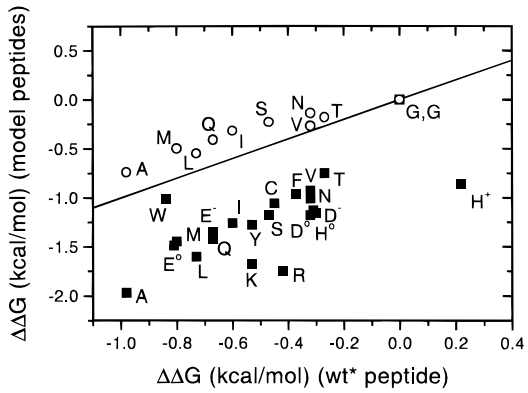

FIGURE 4: Comparison of helix propensity (relative to glycine) measured in the wt* peptide and alanine-based model peptides (solid squares, data from ref *17*) and salt-bridge stabilized model peptides (open circles, data from ref *18*). Solid line is $Y = X$.

In contrast, the wt* peptide appears to accommodate proline quite well. The A21P peptide is only slightly less helical than the A21G peptide. This translates into around 0.1 kcal/mol less than glycine, and 1.1 kcal/mol less than alanine. In most other systems, proline is worse than alanine by a much greater margin: 3−4 kcal/mol in alanine-based peptides (*16*, *17*), 3.5 kcal/mol in T4 lysozyme (*8*), 3 kcal/mol in coiled coils (*9*), and 3 kcal/mol by molecular dynamics calculations (*40*). It is possible that we cannot accurately measure extraordinarily low helix propensities in the wt* peptide. Fluctuating helices could exist in the ends of the peptide regardless of how low the propensity is at residue 21. Therefore, a basal level of helicity exists, and the CD signal at 222 nm would feature some minimum value, independent of the propensity of residue 21. Model peptides with a higher initial helicity would be more suitable for measuring residues with very low helix propensities.

Comparison of RNase T_1 and Other Systems. The first systematic attempt to measure intrinsic helical propensities of the amino acids was made by Scheraga and co-workers in random host polymers (*41*, *42*), work which spanned about 20 years. These polymers consisted of mainly hydroxypropyl/butyl-glutamine, and helical structure was stabilized by hydrophobic interactions between these very large side chains. Comparing our peptide results with data from host/guest experiments shows serious discrepancies. The correlation between the two sets is not bad ($R = 0.77$), but the magnitude of $\Delta\Delta G$ in host/guest peptides is 2−3 times smaller than in our peptide (not shown).

A comparison of our results (sans proline) with those from two other model helical peptides is shown in Figure 4. There is reasonable qualitative agreement with the rank order of propensities measured in alanine-based peptides (*16*, *17*). However, the $\Delta\Delta G$ values for many of the amino acids are almost twice as large as those measured in our peptide (also see Table 3). This is essentially the same result found when

alanine-based peptides have been compared to other protein systems like T4 lysozyme (*20*). There is in better agreement with the salt-bridge stabilized peptides of Kallenbach and co-workers (*18*). Helix propensities of the amino acids reported in ref 18 are in good qualitative agreement with the wt* peptide; however, they are around 30% smaller in magnitude.

Two proteins have been used previously to measure helix propensities. These are T4 lysozyme (*7*, *8*) and barnase (*6*). The approach was the same: substitute all the amino acids at a solvent-exposed helical site and measure the conformational stabilities of the resulting mutants. The $\Delta\Delta G$ values found in wt* protein and peptide are similar in magnitude to the $\Delta\Delta G$s measured in T4 lysozyme and barnase (see Table 3). Similarly good agreement is found between our results and propensities measured in the coiled coil peptides of O'Neil and DeGrado (*9*), and propensities calculated by the program AGADIR, developed by Muñoz and Serrano (*43*) to predict the helical content of peptides.

Most of the systems agree fairly well with one another. The range of propensities agrees particularly well between the wt* peptide/protein and barnase and T4 lysozyme (within 0.1 kcal/mol). However, model peptides do not appear to be in such good agreement. Alanine-based peptides overestimate the propensities found in proteins by a factor of 2. Salt-bridge stabilized peptides give propensity differences which are about 30% smaller than proteins. Host/guest studies underestimate the propensities by a factor of $2-3$.

Our results suggest that the differences observed between some model peptides and wt* protein/peptide as well as other systems are due to specific sequence effects. One possibility is the effect of larger side chains on the solvation of the helix backbone. The range of helix propensities observed in model systems appears to depend inversely on the average size of the residues neighboring the guest position. Alanine is a small side chain and would not be expected to hinder the access of solvent to the peptide group in either the helical or the coil conformation. Alanine-based peptides are usually designed such that the site of substitution is adjacent to only alanine residues and has alanines at most (i, $i + 3$) and (i, $i + 4$) positions (which are on adjacent turns of the helix). In peptides and proteins of natural sequence (like the wt* peptide), a variety of other, larger side chains are present at the (i, $i + 1$), (i, $i + 3$), and (i, $i + 4$) positions. Salt-bridge stabilized model peptides consist of glutamic acid and lysine, which are large side chains. Host/guest peptides in particular contain very large, bulky host side chains (hydroxypropyl or hydroxybutyl glutamine). These side chains adjacent to the site of mutation could effect the solvation of the peptide backbone and therefore the measured propensity of the amino acid being substituted. Blaber et al. (*44*) have examined the crystal structures of several variants of T4 lysozyme containing alanine substitutions in an α-helix. They directly observe better solvation of the backbone amides and carbonyls when large side chains are truncated to alanine.

Thermodynamic Explanations of Helix Propensity. Helix formation is driven by favorable enthalpy (*45*), probably resulting from both the formation of backbone intramolecular hydrogen bonds and van der Waals interactions between the atoms in the helix. Opposing this favorable enthalpy is the unfavorable entropy of fixing the backbone in the helical conformation. Backbone factors clearly affect the helix propensities of proline and glycine. Proline lacks a free backbone NH group, so the helical hydrogen bonding pattern is broken. Additionally, steric clashes occur between the side chain and the previous residue in the helix. The lack of a β-carbon affects the propensity of glycine. Glycine residues are able to occupy a much larger area of ϕ-ψ space than other residues in the random coil. Therefore, the entropic cost of fixing the residue in a helical conformation is much greater than for other residues like alanine. For other residues, the reason for different helix propensities must lie in the side chain. Some side-chain rotamers, which are available to the side chain in the random coil conformation, are restricted in the helical conformation, and therefore, their entropy losses would vary from side chain to side chain, contributing to the measured helix propensities (*46, 47*).

Amino acid side chains also have different steric properties and will bury varying amounts of surface area upon helix formation. Thus the hydrophobic effect, resulting from burial of nonpolar surface (the β-carbon of the side chain, for example), could contribute to helix propensities (*7*). Side chains could also restrict access of solvent to nearby backbone atoms. Therefore, it seems likely that solvation contributes to the energetics of helix formation and differs among the side chains. Additionally, when backbone atoms are shielded from solvent, their partial charges can have stronger interactions with one another. This electrostatic effect has been proposed to contribute to helical propensities (*48*).

Several attempts have been made to rationalize helix propensities based solely on the differences in side-chain entropy (*46, 47*). However, all the factors discussed above probably contribute to some extent. Recently, by taking into account solvation thermodynamics from accessible surface area models and calculating backbone and side-chain conformational entropy, helix propensity differences were predicted fairly successfully in four different systems by Luque et al. (*49*).

CONCLUSIONS

The differences in helix propensity are small, yet significant as proteins on average have 30% of their residues in helical conformations. Additionally, intrinsic propensities are likely to be important to the stability of early folding intermediates, which often have secondary structure but lack specific tertiary interactions (*50, 51*). Helix propensities of the amino acids make equivalent energetic contributions in proteins and peptides. Considering all the measurments in different systems, a consensus scale of helix propensities emerges. Alanine has the highest helix propensity, and in absolute terms, is the only amino acid to contribute favorably to helix stability (*17*). Alanine is followed by amino acids with long hydrophobic side chains, like methionine and leucine, then by residues with polar groups several carbons removed from the backbone, like glutamine and lysine. Aromatic amino acids are in the middle of the scale. β-Branched amino acids and amino acids with polar groups only one carbon from the backbone, like histidine and asparagine, have low helix propensities, followed by glycine. There is then a substantial drop to proline, which clearly has the lowest propensity. These differences in helix propensity are less than 1 kcal/mol for all amino acids except proline.

Measurements of helix propensity in some cases are sensitive to the helical sequence; this appears to be due to

solvation effects. The local sequence of each system probably causes solvation differences, leading to differences in the measured propensities.

ACKNOWLEDGMENT

We thank Robert L. Baldwin and Carol Rohl for communicating their results before they were published, Professor Baldwin and the members of the Pace and Scholtz lab groups for helpful discussions, Geoff Horn for DNA sequencing, Kevin Shaw for generating Figure 1, and Theronica Gray for help in expression and purification of the mutants.

REFERENCES

1. Fersht, A. R., & Serrano, L. (1993) *Curr. Opin. Struct. Biol. 3*, 75−83.
2. Matthews, B. W. (1993) *Annu. Rev. Biochem. 62*, 139−160.
3. Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251−276.
4. Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648−1652.
5. Serrano, L., Sancho, J., Hirshberg, M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 544−559.
6. Horovitz, A., Matthews, J. M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 560−568.
7. Blaber, M., Zhang, X. J., & Matthews, B. W. (1993) *Science 260*, 1637−1640.
8. Blaber, M., Zhang, X. J., Lindstrom, J. D., Pepoit, S. D., Baase, W. A., & Matthews, B. W. (1994). *Mol. Biol. 235*, 600−624.
9. O'Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646−651.
10. Bell, J. A., Becktel, W. J., Sauer, U., Baase, W. A., & Matthews, B. W. (1992) *Biochemistry 31*, 3590−3596.
11. Thapar, R., Nicholson, E. M., Rajagopal, P., Waygood, E. B., Scholtz, J. M., & Klevit, R. E. (1996) *Biochemistry 35*, 11268−11277.
12. Marqusee, S., & Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8898−8902.
13. Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95−118.
14. Chakrabartty, A., & Baldwin, R. L. (1995) *Adv. Protein Chem. 46*, 141−176.
15. Park, S. H., Shalongo, W., & Stellwagen, E. (1993) *Biochemistry 32*, 7048−7053.
16. Chakrabartty, A., Kortemme, T., & Baldwin, R. L. (1994) *Protein Sci. 3*, 843−852.
17. Rohl, C. A., Chakrabartty, A., & Baldwin, R. L. (1996) *Protein Sci. 5*, 2623−2637.
18. Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669−673.
19. Kallenbach, N. R., Lyu, P., & Zhou, H. (1996) *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp 201−259, Plenum Press, New York.
20. Qian, H., & Chan, S. I. (1996) *J. Mol. Biol. 261*, 279−288.
21. Myers, J. K., Smith, J. S., Pace, C. N., & Scholtz, J. M. (1996) *J. Mol. Biol. 263*, 390−395.
22. Myers, J. K., Pace, C. N., & Scholtz, J. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 2833−2837.
23. Pace, C. N., Heinemann, U., Hahn, U., & Saenger, W. (1991) *Angew. Chem., Int. Ed. Engl. 30*, 343−360.
24. Landt, O., Grunert, H. P., & Hahn, U. (1990) *Gene 96*, 125−128.
25. Shirley, B. A., & Laurents, D. V. (1990) *J. Biochem. Biophys. Methods 20*, 181−188.
26. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
27. Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
28. Scholtz, J. M. (1995) *Protein Sci. 4*, 35−43.
29. Fields, G. B., & Noble, R. L. (1990) *Int. J. Pept. Res. 35*, 161−214.
30. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., & Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
31. Lifson, R., & Roig, A. (1961) *J. Chem. Phys. 34*, 1963−1974.
32. Qian, H. (1993) *Biopolymers 33*, 1605−1616.
33. Qian, H., & Schellman, J. A. (1992) *J. Phys. Chem. 96*, 3987−3994.
34. Rohl, C. A., Scholtz, J. M., York, E. J., Stewart, J. M., & Baldwin, R. L. (1992) *Biochemistry 31*, 1263−1269.
35. Armstrong, K. M., & Baldwin, R. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11337−11340.
36. Scholtz, J. M., Qian, H., Robbins, V. H., & Baldwin, R. L. (1993) *Biochemistry 32*, 9668−9676.
37. Huyghues-Despointes, B. M. P., Scholtz, J. M., & Baldwin, R. L. (1993) *Protein Sci. 2*, 1604−1611.
38. Shirley, B. A., Stanssens, P., Hahn, U., & Pace, C. N. (1992) *Biochemistry 31*, 725−732.
39. MacArthur, M. W., & Thornton, J. M. (1991) *J. Mol. Biol. 218*, 397−412.
40. Hermans, J., Anderson, A. G., & Yun, R. H. (1992) *Biochemistry 31*, 5646−5653.
41. Wojcik, J., Altman, K. H., & Scheraga, H. A. (1990) *Biopolymers 30*, 121−134.
42. Sueki, M., Lee, S., Powers, S. P., Denton, J. B., Konishi, Y., & Scheraga, H. A. (1984) *Macromolecules 17*, 148−155.
43. Muñoz, V., & Serrano, L. (1995) *J. Mol. Biol. 245*, 275−296.
44. Blaber, M., Baase, W. A., Gassner, N., & Matthews, B. W. (1995) *J. Mol. Biol. 246*, 317−330.
45. Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M., & Bolen, D. W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2854−2858.
46. Lee, K. H., Xie, D., Freire, E., & Amzel, L. M. (1994) *Proteins*: *Struct., Funct., Genet. 20*, 68−84.
47. Creamer, T. P., & Rose, G. D. (1994) *Proteins: Struct., Funct., Genet. 19*, 85−97.
48. Avbelj, F., & Moult, J. (1995) *Biochemistry 34*, 755−764.
49. Luque, I., Mayorga, O. L., & Freire, E. (1996) *Biochemistry 35*, 13681−13688.
50. Ptitsyn, O. B. (1995) *Adv. Protein Chem. 47*, 83−229.
51. Kuwajima, K., Yamaya, H., & Sugai, S. (1996) *J. Mol. Biol. 264*, 806−822.
52. Martinez-Oyanedel, J., Heineman, U., & Saenger, W. (1991) *J. Mol. Biol. 222*, 335−352.
53. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI9707180