Protein Science (1996), 5:2623–2637. Cambridge University Press. Printed in the USA.

Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol

CAROL A. ROHL,[1,3] AVIJIT CHAKRABARTTY,[2] AND ROBERT L. BALDWIN[2]

[1]Department of Biochemistry, Stanford University, Stanford California 94305
[2]Department of Medical Biophysics, University of Toronto, Ontario Cancer Center, Toronto, Ontario M5G 2M9, Canada

(RECEIVED July 31, 1996; ACCEPTED September 24, 1996)

Abstract

The helix propagation and N-cap propensities of the amino acids have been measured in alanine-based peptides in 40 volume percent trifluoroethanol (40% TFE) to determine if this helix-stabilizing solvent uniformly affects all amino acids. The propensities in 40% TFE are compared with revised values of the helix parameters of alanine-based peptides in water. Revision of the propensities in water is the result of redefining the capping statistical weights and evaluating the helix nucleation constant with N-capping explicitly included in the helix-coil model. The propagation propensities of all amino acids increase in 40% TFE relative to water, but the increases are highly variable. In water, all β-branched and β-substituted amino acids are helix breakers. In 40% TFE, the propagation propensities of the nonpolar amino acids increase greatly, leaving charged and neutral polar, β-substituted amino acids as helix breakers. Glycine and proline are strong helix breakers in both solvents. Free energy differences for helix propagation ($\Delta\Delta G$) between alanine and other nonpolar amino acids are twice as large in water as predicted from side-chain conformational entropies, but $\Delta\Delta G$ values in 40% TFE are close to those predicted from side-chain entropies. This dependence of $\Delta\Delta G$ on the solvent points to a specific role of water in determining the relative helix propensities of the nonpolar amino acids. The N-cap propensities converge toward a common value in 40% TFE, suggesting that differential solvation by water contributes to the diversity of N-cap values shown by the amino acids.

Keywords: alanine-based peptides; helix-breaking amino acids; helix propensities; N-cap propensities; trifluoroethanol

Alanine-based peptides have been used to determine the parameters of the helix-coil transition in water: the helix nucleation constant (Scholtz et al., 1991; Rohl et al., 1992), helix propagation propensities (Chakrabartty et al., 1994), and N-cap propensities (Doig & Baldwin, 1995). Data for helix propensities in protein helices are available from site-directed mutagenesis experiments (Horovitz et al., 1992; Blaber et al., 1993). Comparison of the peptide and protein results shows some similarities and some puzzling differences. Although the rank order of helix propensities is similar in the peptide and protein systems, the range of values is nearly ten-fold larger in alanine-based peptides (see Chakrabartty et al., 1994, and Discussion). In order to better understand the relation between peptide and protein results, it is important to compare the results of peptide helix experiments made with peptide fragments from helical segments of proteins. Recently, Muñoz and Serrano (1994) addressed this problem by using a statistical approach to derive a common set of parameters for both protein fragments and alanine-based peptides from data in the literature.

Peptide fragments from proteins show low helix contents in water, which limits the usefulness of comparing protein fragments with alanine-based peptides. A standard solution to this problem is to add TFE to induce helix formation by peptide fragments of proteins (see Nelson & Kallenbach 1986, 1989; Segawa et al., 1991; Sönnichsen et al., 1992; Storrs et al., 1992; Waterhous & Johnson, 1994; Hamada et al., 1995; and references therein). This procedure raises the problem of how the helix-coil parameters measured in water are related to those measured in TFE:H_2O mixtures. Does adding TFE uniformly enhance the helix propensities of the amino acids? Which amino acids are still helix breakers in the presence of TFE? To investigate the effects of TFE on the helix propagation and N-cap propensities of the amino acids, the helix contents of 36 alanine-based peptides have been determined in 40

Reprint requests to: Carol A. Rohl, Department of Biochemistry, University of Washington, Box 357742, Seattle, Washington 98195-7742; e-mail: rohl@u.washington.edu.

[3]Present address: Department of Biochemistry, University of Washington, Seattle, Washington 98195-7742.

Abbreviations: CD, circular dichroism; FPLC, fast protein liquid chromatography; TFE, trifluoroethanol.

volume percent TFE. For many peptides, adding TFE produces an increase in helix content up to, but not beyond, 40% TFE (Nelson & Kallenbach, 1986; Sönnichsen et al., 1992; Jasanoff & Fersht, 1994; Albert & Hamilton, 1995; Hamada et al., 1995). The reason for this behavior is not well understood. Because 40% TFE usually produces the largest obtainable increase in helix content, this TFE concentration has been chosen for our study.

The helix propagation and N-cap propensities of the amino acids in 40% TFE determined here are compared with revised values of the helix parameters in water. Although helix propagation and capping propensities of the amino acids in water have been previously determined using alanine-based peptides (Chakrabartty et al., 1994; Doig & Baldwin, 1995), two major considerations have prompted us to reevaluate the helix parameters in water. First, the values of the helix nucleation parameter determined earlier by circular dichroism (Scholtz et al., 1991) and by hydrogen exchange (Rohl et al., 1992) were found before the Lifson-Roig (Lifson & Roig, 1961) theory of the helix-coil transition was modified to include N- and C-capping (Doig et al., 1994). Explicit inclusion of capping changes the definition of the nucleation equilibrium constant and consequently affects the values of the helix parameters. We show here that when N-capping is accounted for, the helix nucleation parameter (v^2) decreases from 0.0023 to 0.0013 and the helix propagation propensities (w) are increased fairly uniformly by approximately 10%. Second, in the model used previously to describe N- and C-capping (Chakrabartty et al., 1993b; Doig et al., 1994), the capping free energy contributes to the stability of some random coil conformations as well as stabilizing helical peptide conformations. Revising the model to make the description of capping more physically reasonable causes the range of the N-cap parameters (n) for the amino acids to decrease substantially.

Helix-coil transition theory

The model used to treat the helix-coil transition is based on the Lifson-Roig theory (Lifson & Roig, 1961), which defines residues in the helical (h) conformation as those that have dihedral angles consistent with helix formation. Residues whose dihedral angles are not consistent with helix formation are defined as being in the nonhelical (c) conformation. A residue is defined as a C_α with peptide bonds on both sides. In an unblocked peptide, the number of residues, N_r, is two less than the number of amino acids. In a blocked peptide, N_r is equal to the number of amino acids. In the random coil, both the h and c conformations are allowed. When three consecutive residues ($i - 1, i, i + 1$) are in the h conformation, a helical hydrogen bond is formed between the peptide CO of residue $i - 2$ and the peptide NH of residue $i + 2$. Thus, the helix nucleus, or cooperative unit, is three residues, and a helical segment consists of an unbroken stretch of three or more residues in the h conformation. The complete partition function for the helix-coil transition is determined by summing the statistical weights of all possible conformations of the chain. The c conformation is the Lifson-Roig reference state and is assigned a statistical weight of 1. Residues that are in the helical conformation, but not hydrogen bonded, are assigned weight v. Residues in the h conformation with an associated $i - 2$, $i + 2$ hydrogen bond are assigned the propagation parameter w.

To extract thermodynamic parameters from the Lifson-Roig statistical weights, the statistical weights of helical segments must be compared with the statistical weight of the random coil. The random coil is defined as all conformations that do not include helical segments. One or two contiguous residues in the h conformation do not comprise a helical segment and instead are part of the ensemble of random coil conformations. In the random coil, consequently, a residue can be in either the c or h conformation, and the weight of a random coil residue can be approximated as $(1 + v)$ (Qian & Schellman, 1992). The equilibrium constant for adding a residue to an existing helical segment is the ratio of the weight of the residue when it is part of the helix to that when it is part of the random coil: $w/(1 + v)$. To nucleate a helix, three consecutive residues must be converted from the random coil conformation to the helical conformation. In addition, one residue on each side of the helix must be fixed in the c conformation in order to terminate the helix. The equilibrium constant for nucleating a helix is, therefore, the ratio of the weight of the helix nucleus to that of the random coil of five units: $wv^2/(1 + v)^5$ (Qian & Schellman, 1992).

The Lifson-Roig model has been previously modified to include simple N- and C-capping interactions that depend only on the identity of the residue at the N-cap or C-cap (Doig et al., 1994). In the original modification, a capping residue was defined as one in the c conformation adjacent to a residue in the h conformation. Although a nonhelical residue adjacent to a helical segment fulfills this definition, some residues in the random coil will also meet this criteria because both the c and h conformations are allowed in the random coil. According to this description, the capping statistical weight is applied not only to helical segments, but also to random coil segments. Consequently, the capping weights, n and c, are not related to interaction free energies in a straightforward manner. In addition, it becomes difficult to relate the intrinsic helix parameters w and v to propagation and nucleation equilibrium constants because the weight of a random coil residue can no longer be approximated as $(1 + v)$.

Here we have modified the description of simple capping interactions in the current Lifson-Roig-based model to eliminate the contribution of the capping weights to random coil conformations. The N-cap position is defined as the nonhelical residue immediately N-terminal to a helical segment and is the first residue with a hydrogen-bonded peptide CO. Analogously, the C-cap residue is defined as the nonhelical residue immediately C-terminal to a helical segment and is the last residue in the helix with a hydrogen-bonded peptide NH. We define the statistical weights n and c as relative equilibrium constants for placing a particular residue at the N-cap or C-cap position of a helical segment:

$$\Delta\Delta G(\text{N-cap}) = -RT \ln [n] \quad (1a)$$

$$\Delta\Delta G(\text{C-cap}) = -RT \ln [c] \quad (1b)$$

Alanine adjacent to the helical segment is the reference for both N-capping and C-capping interactions. Capping interactions are assumed to occur only when a helical segment of at least one hydrogen bond is formed. Consequently, the capping weights n and c are applied only to nonhelical residues that are adjacent to at least three consecutive residues in the h conformation. When simple capping interactions are added to the Lifson-Roig formalism according to this model, the equilibrium constant for nucleating a helix becomes $wv^2nc/(1 + v)^5$.

The advantage of the original definitions of the capping weights is that correlating the conformations of three consecutive residues is sufficient to identify both N-cap and C-cap units. Three residues are not sufficient, however, to identify the ends of a helical segment because the smallest helical segment, chhhc, is five units long. Assignment of the newly defined capping weights, which are applied only to nonhelical units adjacent to helical segments, requires correlating the conformation of five consecutive residues.

Modifying the definition of the capping weights, consequently, makes partition function calculations more computationally demanding, because a larger correlation matrix is required. We were convinced, however, to modify our treatment of capping after a discussion with Dr. Niels Andersen (pers. comm.), who had independently concluded that the original definition of the capping statistical weights allowed capping free energies to contribute to random coil conformations.

The statistical weights described above can be combined in a 16 × 16 correlation matrix (not shown) that represents all possible conformations of five consecutive residues. This matrix can be reduced to a 6 × 6 matrix, shown below, in which the row labels indicate the conformation of residues $i - 2$, $i - 1$, i, and $i + 1$, and the column labels indicate the conformation of residues $i - 2$, i, $i + 1$, and $i + 2$:

$$\mathbf{M}_i = \begin{array}{l|cccccc} & \mathbf{h\bar{h}hh} & \mathbf{h\bar{h}hc} & \mathbf{c\bar{h}h(c \cup h)} & \mathbf{(c \cup h)\bar{h}c(c \cup h)} & \mathbf{(c \cup h)\bar{c}h(c \cup h)} & \mathbf{(c \cup h)\bar{c}c(c \cup h)} \\ \hline \mathbf{hh\bar{h}h} & w_i & w_i c_{i+2} & 0 & 0 & 0 & 0 \\ \mathbf{hh\bar{h}c} & 0 & 0 & 0 & v_i & 0 & 0 \\ \mathbf{ch\bar{h}(c \cup h)} & w_i n_{i-2} & w_i n_{i-2} c_{i+2} & 0 & v_i & 0 & 0 \\ \mathbf{(c \cup h)h\bar{c}(c \cup h)} & 0 & 0 & 0 & 0 & 1 & 1 \\ \mathbf{(c \cup h)c\bar{h}(c \cup h)} & 0 & 0 & v_i & v_i & 0 & 0 \\ \mathbf{(c \cup h)c\bar{c}(c \cup h)} & 0 & 0 & 0 & 0 & 1 & 1 \end{array}$$

Weights in the matrix are assigned to residue i. The subscripts on statistical weights indicate the position of the residue on which they depend. The partition function, Z, is calculated according to

$$Z = \mathbf{V}\left[\prod_{i=0}^{N_r+1} \mathbf{M}_i\right]\mathbf{V}^{\dagger}$$

where

$$\mathbf{V} = [0\,0\,0\,0\,0\,1]$$

and $\mathbf{V}^{\dagger}$ is its transpose. Properties of the system, such as the average number of hydrogen bonds formed, $\langle n_H \rangle$, or the probability that the peptide NH of residue i is hydrogen bonded, $f_B(i)$ can be calculated from the partition function.

Results

Re-evaluation of the helix parameters in water

The introduction of capping into the original Lifson-Roig model changes the equilibrium constant for nucleation from $v^2w/(1 + v)^5$ to $ncv^2w/(1 + v)^5$. Consequently, the value of v^2 must change when capping is included in the model, in order for the free energy of nucleation to remain unchanged. We have redetermined v^2 from the total NH exchange kinetics of a series of alanine-based peptides with the general sequence Ac-(AAKAA)$_m$Y-NH$_2$ and chain lengths varying from 6–51 residues. These data were used previously (Rohl et al., 1992) to determine the value of v^2 without taking capping into account. When v^2 is redetermined with capping included in the helix coil model, the determined value of v^2 decreases from 0.0023 to 0.0013 (Fig. 1). This decrease partly reflects the fact that in these peptides the helix preferentially nucleates with acetyl in the N-cap position because acetyl is a strong capping group.

The smaller value of the helix nucleation constant and the redefinition of the capping statistical weights necessitates re-evaluation of all helix propagation and N-cap propensities in water. The revised values of the helix parameters of the amino acids in water, determined using previously published mean helix contents of alanine-based peptides (Table 1), are given in Table 2. Redefinition of the N-cap statistical weight has a significant effect on the N-cap propensities of the amino acids. The order of N-cap parameters is unchanged, but the range is significantly decreased relative to the values reported by Doig & Baldwin (1995) (Fig. 2A). The C-cap parameters are not re-evaluated here because C-cap interactions have been previously observed not to vary significantly among the different amino acids in water (Chakrabartty et al., 1993b; 1994; Doig & Baldwin, 1995; see also Materials and methods).

The revised helix propagation propensities are largely unchanged, both in order and range, from those reported by Chakrabartty et al. (1994). His$^+$ is the only amino acid that shows a large change in helix propagation propensity. The helix propensity of His$^+$ was previously determined using alanine-based peptides solubilized with either Lys (AK peptides) or Gln (AQ peptides) (Armstrong & Baldwin, 1993). The AQ peptides were observed to be less destabilized by a His$^+$ substitution than the AK peptides. In redetermining the helix propagation propensity of His$^+$, we have used only the AQ peptides in which unfavorable charge–charge interactions should be minimized. For all other amino acids, the revised values of the helix propagation propensities are approximately 10% larger than previously reported values (Fig. 2B).

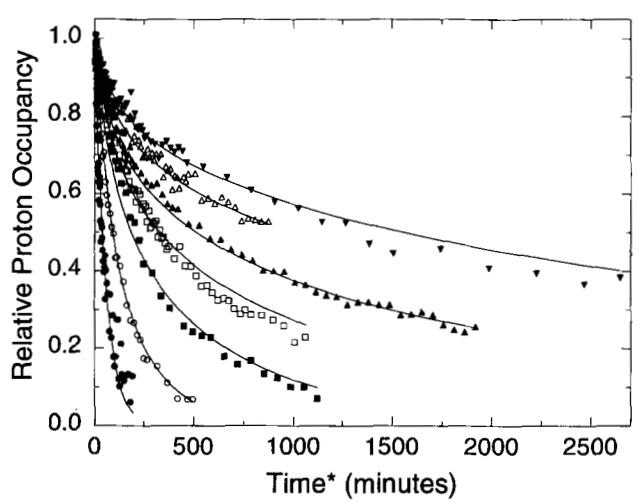

Fig. 1. Determination of the helix initiation parameter. Total NH exchange curves for peptides of the general sequence Ac-(AAKAA)$_m$Y-NH$_2$ with chain lengths of 6 (●), 16 (○), 21 (■), 26 (□), 31 (▲), 41 (△), and 51 (▼) residues. The data are taken from Rohl et al. (1992). Time * is an adjusted time scale that corrects for small differences in the measured pH* of different samples (Rohl et al., 1992). The solid lines are the best fit of the helix-coil model and Equation 3 to the data. The fitted parameters are $\langle w \rangle = 1.50$, $\langle v^2 \rangle = 0.0013$.

Table 1. *Ellipticities, helix contents, chain lengths, and sequences of peptides used to determine helix parameters in water*

Peptide	$-[\theta]_{222}$[a] (deg cm^2 dmol^{-1})	f_H[b]	N_r	Sequence
NA-GY17	12,900[c]	0.385	16	AAKAAAAKAAAAKAAGY-NH_2
NQ-GY17	10,000[c]	0.303	16	QAKAAAAKAAAAKAAGY-NH_2
NV-GY17	12,800[c]	0.382	16	VAKAAAAKAAAAKAAGY-NH_2
NM-GY17	13,700[c]	0.408	16	MAKAAAAKAAAAKAAGY-NH_2
NP-GY17	13,800[c]	0.411	16	PAKAAAAKAAAAKAAGY-NH_2
NI-GY17	14,300[c]	0.425	16	IAKAAAAKAAAAKAAGY-NH_2
NL-GY17	15,300[c]	0.453	16	LAKAAAAKAAAAKAAGY-NH_2
NT-GY17	15,600[c]	0.462	16	TAKAAAAKAAAAKAAGY-NH_2
NG-GY17	17,900[c]	0.527	16	GAKAAAAKAAAAKAAGY-NH_2
NS-GY17	17,900[c]	0.527	16	SAKAAAAKAAAAKAAGY-NH_2
NN-GY17	20,300[c]	0.595	16	NAKAAAAKAAAAKAAGY-NH_2
Ac-GY18	19,700[d]	0.578	16	Ac-AKAAAAKAAAAKAAGY-NH_2
ND-GY17	20,200[e]	0.593	16	DAKAAAAKAAAAKAAGY-NH_2
NW-GY17	17,600[e]	0.519	16	WAKAAAAKAAAAKAAGY-NH_2
NH-GY17	15,400[e]	0.456	16	HAKAAAAKAAAAKAAGY-NH_2
NE-GY17	15,300[e]	0.453	16	EAKAAAAKAAAAKAAGY-NH_2
NR-GY17	12,900[e]	0.385	16	RAKAAAAKAAAAKAAGY-NH_2
NK-GY17	12,100[e]	0.362	16	KAKAAAAKAAAAKAAGY-NH_2
NF-GY17	15,300[e]	0.453	16	FAKAAAAKAAAAKAAGY-NH_2
NC-GY17	19,400[e]	0.570	16	CAKAAAAKAAAAKAAGY-NH_2
NY-GY17	18,900[e]	0.556	16	YAKAAAAKAAAAKAAGY-NH_2
YGGAK	24,000[f]	0.676	19	Ac-YGGKAAAAKAAAAKAAAAK-NH_2
YGG-G17	16,100[f]	0.460	19	Ac-YGGKAAAAKAAAAKAAGAK-NH_2
YGG-G12	8,900[f]	0.262	19	Ac-YGGKAAAAKAAGAKAAAAK-NH_2
YGG-G7	12,900[f]	0.372	19	Ac-YGGKAAGAKAAAAKAAAAK-NH_2
YGAK	25,400[f]	0.722	18	Ac-YGKAAAAKAAAAKAAAAK-NH_2
YGGGAK	23,500[f]	0.657	20	Ac-YGGGKAAAAKAAAAKAAAAK-NH_2
AcA-GY12	8,900[c]	0.292	12	Ac-AAKAAAAKAAGY-NH_2
YGG-1Q	17,600[g]	0.501	19	Ac-YGGKAAAAKAQAAKAAAAK-NH_2
YGG-3Q	10,800[g]	0.315	19	Ac-YGGKAQAAKAQAAKAQAAK-NH_2
2Q-GGY	13,700[g]	0.394	19	Ac-KAQAAKAQAAKAAAAKGGY-NH_2
N2Q-GGY	7,600[g]	0.229	18	KAQAAKAQAAKAAAAKGGY-NH_2
Qref	18,000[g]	0.512	19	Ac-YGGQAAAAQAAAAQAAAAQ-NH_2
Q12	13,100[g]	0.377	19	Ac-YGGQAAAAQAQAAQAAAAQ-NH_2
YG-ZC17	18,200[e]	0.579	16	Ac-YGAAKAAAAKAAAAKA-NH_2
YGG-1L	19,000[g]	0.538	19	Ac-YGGKAAAAKALAAKAAAAK-NH_2
YGG-3L	17,100[g]	0.487	19	Ac-YGGKALAAKALAAKALAAK-NH_2
2L-GGY	16,900[g]	0.482	19	Ac-KALAAKALAAKAAAAKGGY-NH_2
N2L-GGY	10,100[g]	0.298	18	KALAAKALAAKAAAAKGGY-NH_2
YGG-1M	17,600[g]	0.501	19	Ac-YGGKAAAAKAMAAKAAAAK-NH_2
YGG-3M	12,600[g]	0.365	19	Ac-YGGKAMAAKAMAAKAMAAK-NH_2
2M-GGY	14,000[g]	0.402	19	Ac-KAMAAKAMAAKAAAAKGGY-NH_2
N2M-GGY	8,500[g]	0.253	18	KAMAAKAMAAKAAAAKGGY-NH_2
YGG-1I	16,000[g]	0.456	19	Ac-YGGKAAAAKAIAAKAAAAK-NH_2
YGG-3I	7,500[g]	0.224	19	Ac-YGGKAIAAKAIAAKAIAAK-NH_2
I-GGY	16,800[g]	0.479	19	Ac-KAAAAKAIAAKAAAAKGGY-NH_2
NI-GGY	10,000[g]	0.295	18	KAAAAKAIAAKAAAAKGGY-NH_2
YGG-1C	15,700[g]	0.449	19	Ac-YGGKAAAAKACAAKAAAAK-NH_2
YGG-1S	14,800[g]	0.424	19	Ac-YGGKAAAAKASAAKAAAAK-NH_2
YGG-3S	7,700[g]	0.229	19	Ac-YGGKASAAKASAAKASAAK-NH_2
S-GGY	15,600[g]	0.445	19	Ac-KAAAAKASAAKAAAAKGGY-NH_2
NS-GGY	7,400[g]	0.223	18	KAAAAKASAAKAAAAKGGY-NH_2
YGG-1F	14,800[g]	0.424	19	Ac-YGGKAAAAKAFAAKAAAAK-NH_2
YGG-1N	13,800[g]	0.396	19	Ac-YGGKAAAAKANAAKAAAAK-NH_2
YGG-3N	5,700[g]	0.174	19	Ac-YGGKANAAKANAAKANAAK-NH_2
N-GGY	14,900[g]	0.426	19	Ac-KAAAAKANAAKAAAAKGGY-NH_2
NN-GGY	7,300[g]	0.220	18	KAAAAKANAAKAAAAKGGY-NH_2
YGG-1T	10,200[g]	0.299	19	Ac-YGGKAAAAKATAAKAAAAK-NH_2
YGG-3T	2,700[g]	0.093	19	Ac-YGGKATAAKATAAKATAAK-NH_2

(*continued*)

Table 1. *Continued*

Peptide	$-[\theta]_{222}$[a] (deg cm^2 dmol^{-1})	f_H[b]	N_r	Sequence
T-GGY	14,000[g]	0.401	19	Ac-KAAAAKATAAKAAAAKGGY-NH_2
NT-GGY	6,300[g]	0.192	18	KAAAAKATAAKAAAAKGGY-NH_2
YGG-1P	0[g]	0.018	19	Ac-YGGKAAAAKAPAAKAAAAK-NH_2
YGG-3V	3,200[g]	0.106	19	Ac-YGGKAVAAKAVAAKAVAAK-NH_2
V-GGY	15,000[g]	0.430	19	Ac-KAAAAKAVAAKAAAAKGGY-NH_2
NV-GGY	6,800[g]	0.206	18	KAAAAKAVAAKAAAAKGGY-NH_2
W16	15,600[g]	0.462	16	Ac-KAAAAKAWAAKAAAAK-NH_2
Y16	18,800[g]	0.553	16	Ac-KAAAAKAYAAKAAAAK-NH_2
AQ16	16,000[h]	0.483	16	Ac-AAQAAAAQAAAAQAAY-NH_2
E5	11,700[h,l]	0.359	16	Ac-AAQAEAAQAAAAQAAY-NH_2
E6	12,100[h,l]	0.368	16	Ac-AAQAAEAQAAAAQAAY-NH_2
E9	9,300[h,l]	0.289	16	Ac-AAQAAAAQEAAAQAAY-NH_2
E11	9,000[h,l]	0.279	16	Ac-AAQAAAAQAAEAQAAY-NH_2
$E^0$5	12,100[h,m]	0.368	16	Ac-AAQAEAAQAAAAQAAY-NH_2
$E^0$6	12,400[h,m]	0.378	16	Ac-AAQAAEAQAAAAQAAY-NH_2
$E^0$9	11,400[h,m]	0.349	16	Ac-AAQAAAAQEAAAQAAY-NH_2
$E^0$11	12,100[h,m]	0.368	16	Ac-AAQAAAAQAAEAQAAY-NH_2
D4	14,600[i,l]	0.441	16	Ac-AAQDAAAQAAAAQAAY-NH_2
D6	9,200[i,l]	0.285	16	Ac-AAQAADAQAAAAQAAY-NH_2
D9	8,300[i,l]	0.259	16	Ac-AAQAAAAQDAAAQAAY-NH_2
D10	6,100[i,l]	0.195	16	Ac-AAQAAAAQADAAQAAY-NH_2
$D^0$4	11,600[i,m]	0.354	16	Ac-AAQDAAAQAAAAQAAY-NH_2
$D^0$6	8,400[i,m]	0.262	16	Ac-AAQAADAQAAAAQAAY-NH_2
$D^0$9	8,300[i,m]	0.259	16	Ac-AAQAAAAQDAAAQAAY-NH_2
$D^0$10	7,800[i,m]	0.244	16	Ac-AAQAAAAQADAAQAAY-NH_2
AQH5	5,900[j,n]	0.187	16	Ac-AAQAHAAQAAAAQAAY-NH_2
AQH6	4,500[j,n]	0.146	16	Ac-AAQAAHAQAAAAQAAY-NH_2
AQH10	7,400[j,n]	0.229	16	Ac-AAQAAAAQAHAAQAAY-NH_2
AQH11	9,100[j,n]	0.277	16	Ac-AAQAAAAQAAHAQAAY-NH_2
$AQH^0$5	9,100[j,o]	0.276	16	Ac-AAQAHAAQAAAAQAAY-NH_2
$AQH^0$6	8,100[j,o]	0.249	16	Ac-AAQAAHAQAAAAQAAY-NH_2
$AQH^0$10	8,500[j,o]	0.261	16	Ac-AAQAAAAQAHAAQAAY-NH_2
$AQH^0$11	9,100[j,o]	0.277	16	Ac-AAQAAAAQAAHAQAAY-NH_2
R5	14,400[k,m]	0.427	16	Ac-AAQARAAQAAAAQAAY-NH_2
R6	13,200[k,m]	0.390	16	Ac-AAQAARAQAAAAQAAY-NH_2
R10	16,600[k,m]	0.492	16	Ac-AAQAAAAQARAAQAAY-NH_2
R11	15,600[k,m]	0.463	16	Ac-AAQAAAAQAARAQAAY-NH_2

[a] 1 M sodium chloride, pH 7.0 for acetylated peptides, pH 9.55 for unblocked peptides, unless otherwise noted.
[b] Calculated from Equation 3.
[c] Chakrabartty et al., 1993b.
[d] Doig et al., 1994.
[e] Doig & Baldwin, 1995.
[f] Chakrabartty et al., 1993a.
[g] Chakrabartty et al., 1994.
[h] Scholtz et al., 1993.
[i] Huyghues-Despointes et al., 1993.
[j] Armstrong & Baldwin, 1993.
[k] B.M.P. Huyghues-Despointes, thesis.
[l] 10 mM sodium chloride, pH 7.0.
[m] 10 mM sodium chloride, pH 2.5.
[n] 10 mM sodium chloride, pH 5.0.
[o] 10 mM sodium chloride, pH 8.5.

Peptide design and helix contents in TFE

Most of the peptides used in earlier studies of helix formation in water have helix contents in 40% TFE that are too high to be useful for determining helix or N-cap propensities (see Materials and methods). The peptides studied here in 40% TFE are generally shorter than sequences examined in water, and are designed to have helix contents in the range of 20–80%, so that small changes in free energy result in significant changes in ellipticity. All the sequences studied are alanine-based with charged residues included for solubility. A single tyrosine is included for concentration determination, and is separated from the rest of the sequence by glycine to eliminate aromatic contributions to the CD (Chakrabartty et al., 1993a). The design of such peptides has been dis-

Table 2. *Helix propagation and N-cap propensities and free energies of amino acid residues in water at 273 K*

Residue	Helix propagation		N-cap	
	w	$\Delta G°$(helix)[a] (kcal/mol)	n	$\Delta\Delta G°$(N-cap)[b] (kcal/mol)
Ala	1.70	−0.27		
Glu^0	0.70	0.21		
Cys^0	0.32	0.64		
Cys^-			5.4	−0.92
Asp^-	0.38	0.54	6.6	−1.0
Glu^-	0.54	0.35	2.06	−0.39
Phe	0.27	0.73	2.06	−0.39
Gly	0.048	1.7	3.9	−0.74
His^+	0.22	0.84		
Ile	0.46	0.44	1.57	−0.25
His^0	0.36	0.57	2.12	−0.41
Lys^+	1.00	0.019	0.72	0.18
Leu	0.87	0.095	2.06	−0.39
Met	0.65	0.25	1.31	−0.15
Asn	0.29	0.69	6.8	−1.0
Asp^0	0.40	0.52		
Pro	<0.001	>3.8	1.35	−0.16
Gln	0.62	0.28	0.12	1.2
Arg^+	1.14	−0.052	1.00	0.0
Ser	0.40	0.52	3.9	−0.74
Thr	0.18	0.95	2.23	−0.44
Val	0.25	0.77	0.96	0.022
Trp	0.29	0.69	3.6	−0.70
Tyr	0.48	0.42	4.9	−0.96
Acetyl			5.9	−0.86

[a] $\Delta G(\text{helix}) = -RT\ln[w/(1+v)]$, where $v = 0.036$.
[b] $\Delta\Delta G(\text{N-cap}) = -RT\ln[n]$. N-cap propensities and free energies are relative to Ala.

cussed previously (Chakrabartty et al., 1994). Guest amino acids are substituted at one or more positions in each peptide. Tyr and Trp were not examined in 40% TFE because optical contributions from their aromatic side chains introduce large errors into measurements of helix content by CD (Chakrabartty et al., 1993a).

The positions in which guest residues are substituted are selected both to minimize potential side chain–side chain interactions and to be sensitive to the parameter being determined. Helix contents of peptides with N-terminal substitutions are used to determine the N-cap parameter. This position is most sensitive to the N-capping propensity of an amino acid. Similarly, the helix contents of peptides with a centrally substituted guest residue are most sensitive to changes in the propagation parameter, w, because central residues have the highest probability of being embedded in a helical segment. The sequences and ellipticities of peptides studied in 40% TFE are given in Table 3. The helix propagation and N-cap propensities in 40% TFE, determined by fitting the helix-coil model to the ellipticities of these peptides, are given in Table 4.

Comparison of helix parameters in water and TFE

The helix propagation propensities of the amino acids, expressed as $\Delta G(\text{helix}) = -RT\ln[w/(1+v)]$, in 40% TFE and in water are compared in Figure 3. As a group, the nonpolar amino acids show the largest increases, but the extent of the increase from water to 40% TFE is not uniform. Some of the charged (Glu^-, His^+) and neutral polar (Gln, Asn, Thr) amino acids also show large increases. The two strongest helix breakers in water, Pro and Gly, retain this property in 40% TFE and are not shown in Figure 3. None of the nonpolar amino acids is a helix breaker in 40% TFE. Phe and Val, which are helix breakers in water, become helix neutral [w near 1, ΔG(helix) near 0] in 40% TFE. Among the charged and neutral polar amino acids, Lys^+ and Arg^+ are helix neutral in both water and 40% TFE, and Asn also becomes helix neutral in 40% TFE. Glu and Gln become helix formers, while the other charged and neutral polar amino acids (Cys, Ser, Thr, His^0,

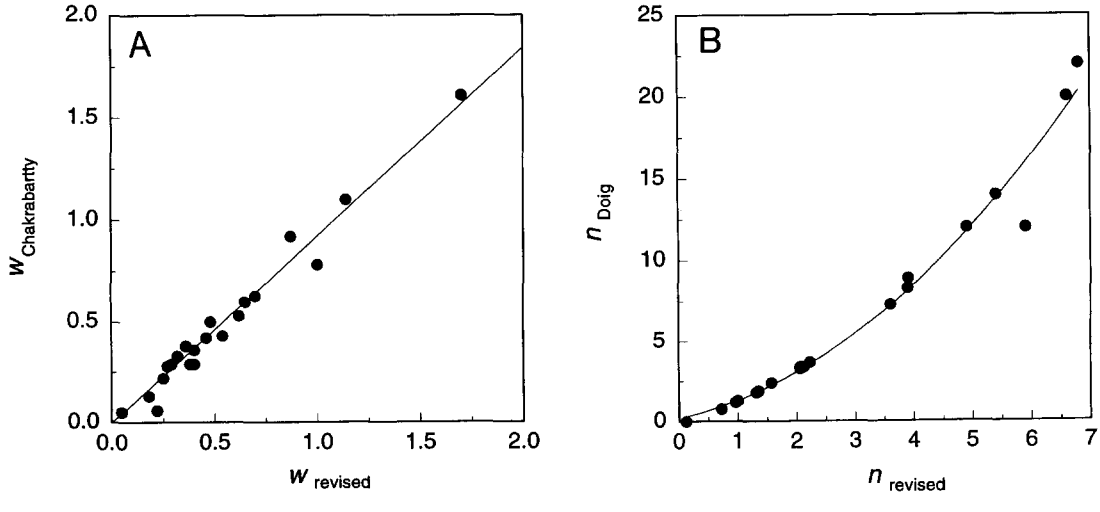

Fig. 2. Comparison of revised and previously reported helix parameters in water. The revised values of helix parameters determined in water are compared with the N-capping propensities reported by Doig and Baldwin (1995) (panel A) and the helix propagation propensities reported by Chakrabartty et al. (1994) (panel B). The solid line in panel A is the best correlation, with the intercept fixed at 0: $w_{Chakrabartty} = 0.92w_{revised}$, $R = 0.984$. The solid line in panel B is drawn to guide the eye.

Table 3. *Ellipticities, helix contents, chain lengths, and sequences of peptides used to determine helix parameters in 40% TFE*

Peptide	$-[\theta]_{222}$ (deg cm^2 dmol^{-1})	f_H[a]	N_r	Sequence
NA-AK11[b]	15,800	0.521	11	AAKAAAAKAAGY-NH_2
NQ-AK11[b]	15,200	0.502	11	QAKAAAAKAAGY-NH_2
NV-AK11[b]	16,700	0.550	11	VAKAAAAKAAGY-NH_2
NI-AK11[b]	17,500	0.575	11	IAKAAAAKAAGY-NH_2
NP-AK11[b]	15,600	0.515	11	PAKAAAAKAAGY-NH_2
NL-AK11[b]	17,400	0.572	11	LAKAAAAKAAGY-NH_2
NT-AK11[b]	17,900	0.588	11	TAKAAAAKAAGY-NH_2
NG-AK11[b]	18,600	0.610	11	GAKAAAAKAAGY-NH_2
NS-AK11[b]	19,500	0.638	11	SAKAAAAKAAGY-NH_2
NN-AK11[b]	20,900	0.683	11	NAKAAAAKAAGY-NH_2
NM-AK11	18,900	0.619	11	MAKAAAAKAAGY-NH_2
N(Ac)-AK11	20,800	0.680	11	Ac-AKAAAAKAAGY-NH_2
NK-AK11	18,600	0.610	11	KAKAAAAKAAGY-NH_2
NR-AK11	16,200	0.534	11	RAKAAAAKAAGY-NH_2
A-AR12	21,600	0.684	12	Ac-AARAAAARAAGY-NH_2
Q-AR12	21,100	0.669	12	Ac-AARAQAARAAGY-NH_2
L-AR12	21,000	0.666	12	Ac-AARALAARAAGY-NH_2
I-AR12	20,200	0.641	12	Ac-AARAIAARAAGY-NH_2
M-AR12	20,000	0.635	12	Ac-AARAMAARAAGY-NH_2
V-AR12	19,300	0.613	12	Ac-AARAVAARAAGY-NH_2
N-AR12	18,400	0.586	12	Ac-AARANAARAAGY-NH_2
S-AR12	15,900	0.509	12	Ac-AARASAARAAGY-NH_2
T-AR12	16,300	0.521	12	Ac-AARATAARAAGY-NH_2
F-AR12	18,700	0.595	12	Ac-AARAFAARAAGY-NH_2
H^+-AR12	15,600	0.499	12	Ac-AARAHAARAAGY-NH_2
C-AR12	14,000	0.450	12	Ac-AARACAARAAGY-NH_2
H^0-AR12	17,100	0.546	12	Ac-AARAHAARAAGY-NH_2
YGG-P11[c]	12,700	0.366	19	Ac-YGGKAAAAKAPAAKAAAAK-NH_2
YGG-G12[d]	26,000	0.731	19	Ac-YGGKAAAAKAAGAKAAAAK-NH_2
YGG-G7[d]	24,500	0.690	19	Ac-YGGKAAGAKAAAAKAAAAK-NH_2
A-AQ12	25,700	0.810	12	Ac-AAQAAAAQAAGY-NH_2
R-AQ12	24,900	0.785	12	Ac-AAQARAAQAAGY-NH_2
K-AQ12	24,800	0.782	12	Ac-AAQAKAAQAAGY-NH_2
AK12	21,600	0.684	12	Ac-AAKAAAAKAAGY-NH_2
AD12	14,300	0.459	12	Ac-AADAAAADAAGY-NH_2
AE12	24,500	0.773	12	Ac-AAEAAAAEAAGY-NH_2

[a]Calculated from Equation 3.
[b]Synthesis and purification described by Chakrabartty et al., 1993b.
[c]Synthesis and purification described by Chakrabartty et al., 1994.
[d]Synthesis and purification described by Chakrabartty et al., 1993c.

Asp^-, and His^+) remain helix breakers in 40% TFE. Thus, except for Gly and Pro, the helix breakers in 40% TFE contain β-substituted side chains with charged or neutral polar groups.

The N-cap propensities of the amino acids in 40% TFE and in water are compared in Figure 4. N-cap propensities in both water and 40% TFE are expressed relative to the propensity of Ala and are given as $\Delta\Delta G(\text{N-cap}) = -RT\ln[n]$ (see Discussion). Although Gln is the worst N-cap residue in both water and TFE, it shows the largest change of all the amino acids, and in 40% TFE it is not significantly worse than Pro, which is a reasonably good N-cap residue in water. Ala remains near the bottom of the ranking of relative N-cap propensities in both solvents. The polar amino acids, which are strong N-cap residues in water (Asn, Ser, Thr), show fairly uniform changes in 40% TFE, as do Gly and the acetyl group. The changes in N-cap value from water to 40% TFE are quite variable among the amino acids, however, both in the magnitude and direction of the change. In one respect, the amino acids do show a common behavior: the N-cap values of all amino acids converge toward a common value in 40% TFE, yielding a substantially decreased range of propensities in 40% TFE compared with water.

Discussion

Helix parameters in water

There are two general effects of changing the definition of the capping weights and reevaluating all the helix parameters in water with capping explicitly included. The range of N-cap values is significantly decreased, but the relative order is unaffected. This change is expected because the capping free energy now contributes only to helical conformations, not to random coil conforma-

Table 4. *Helix propagation and N-cap propensities and free energies of amino acid residues in 40% TFE at 273 K*

Residue	Helix propagation		N-capping	
	w	$\Delta G°$(helix)[a] (kcal/mol)	n	$\Delta\Delta G°$(N-cap)[b] (kcal/mol)
Ala	2.70	−0.52		
Cys^0	0.42	0.49		
Asp^-	0.40	0.52		
Glu^-	1.69	−0.27		
Phe	1.00	0.02		
Gly	0.098	1.3	2.09	−0.40
His^+	0.56	0.33		
Ile	1.42	−0.17	1.56	−0.24
His^0	0.73	0.19		
Lys^+	1.15	−0.06	2.09	−0.40
Leu	1.77	−0.29	1.52	−0.23
Met	1.37	−0.15	2.26	−0.44
Asn	0.98	0.03	4.1	−0.77
Pro	0.0016	3.5	0.95	0.03
Gln	1.80	−0.30	0.85	0.09
Arg^+	1.19	−0.08	1.11	−0.06
Ser	0.60	0.30	2.67	−0.53
Thr	0.63	0.27	1.74	−0.30
Val	1.14	−0.05	1.27	−0.13
Acetyl			4.0	−0.075

[a] $\Delta G(\text{helix}) = -RT \ln [w/(1 + v)]$, where $v = 0.036$.
[b] $\Delta\Delta G(\text{N-cap}) = -RT \ln [n]$. N-cap propensities and free energies are relative to Ala.

tions. The revised N-cap values are related to previously determined N-cap parameters in a predictable way. Consequently, redefining the N-cap statistical weights does not change any predictions of the helix-coil model, but merely changes the energies assigned to N-cap interactions. Because the redefined capping statistical weights contribute only to helical conformations, the revised N-cap parameters reflect real free energies for stabilization of the peptide helix, relative to the random coil, by N-cap interactions. When N-capping is accounted for, the helix nucleation parameter (v^2) decreases from 0.0023 to 0.0013 and the helix propagation propensities (w) are increased fairly uniformly by approximately 10%. Because the helix propagation propensities are determined from peptides with substitutions at central sites, changes in the N-cap propensities do not significantly affect the determined propagation propensities. Instead, the changes in the w-values are dominated by the revised value of v^2.

The relative helix-forming tendencies of the amino acids in water are well-determined in the set of peptides in Table 1. The observed helix contents of the alanine-based peptides indicate that the helix propagation propensity of Ala is substantially larger than that of Lys. This conclusion is also supported by the high helix contents observed for alanine-based peptides solubilized by either Gln or Arg. Kemp and coworkers have concluded, however, from short, template-nucleated helices that the helix propensity of Ala is close to 1, and that alanine-based peptides are helical because the solubilizing Lys residues have a high helical propensity (Groebke et al., 1996). The helix content is not measured directly in the nucleated helices of Kemp et al. but is inferred from the *cis/trans* ratio of the proline-based template. The measured temperature dependence of the cis/trans ratio of the template linked to a six residue alanine peptide gives the surprising result that the alanine helix does not melt with increasing temperature (Kemp et al., 1996). Although the explanation for the differences observed between alanine-based helices and template-nucleated helices studied by Kemp and coworkers is unknown, the discrepancy is unlikely to be a direct result of helix nucleation by a template. Zhou et al. (1994) measured NH protection factors in a helix nucleated by a covalent i, $i + 4$ side-chain bridge and found the Zimm-Bragg helix propagation propensity of alanine to be 1.7 ± 0.2, in good agreement with results from alanine-based peptides.

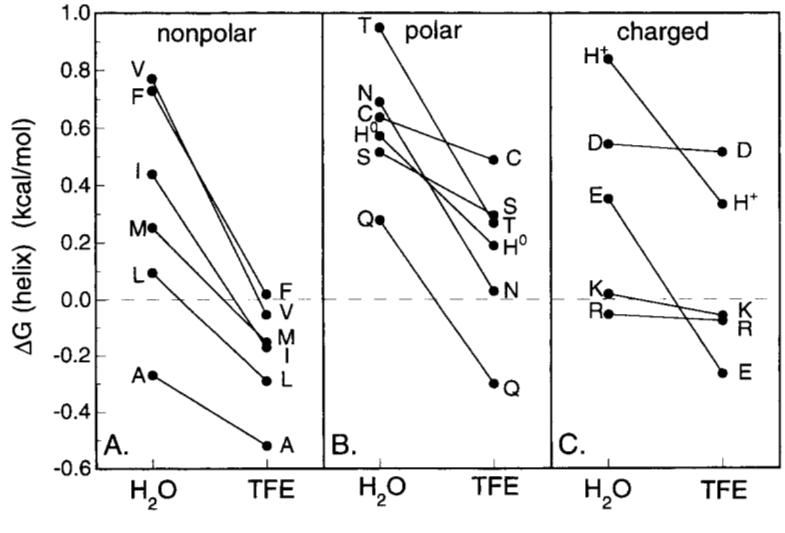

Fig. 3. Comparison of helix propagation propensities in water and 40% TFE. The free energies of helix propagation are indicated for the nonpolar (panel A), polar (panel B), and charged (panel C) amino acids. Solid lines are drawn to guide the eye and do not imply a linear dependence of helix propensity on TFE concentration. Gly and Pro are strong helix breakers in both water and 40% TFE and are not shown.

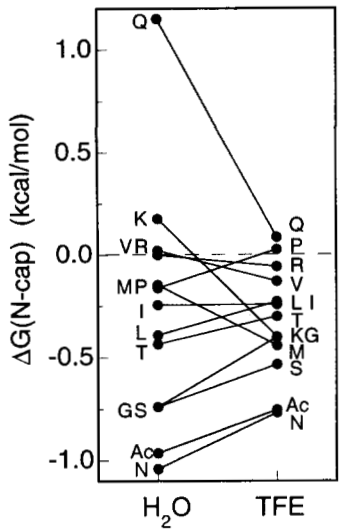

Fig. 4. Comparison of N-capping free energies in water and 40% TFE. Solid lines are drawn to guide the eye and do not imply a linear dependence of N-cap values on TFE concentration. N-cap propensities are relative to alanine in both water and 40% TFE.

Non-uniform increases in helix propensity among the amino acids

Figure 3 illustrates that many amino acids show substantial increases in helix propensity in 40% TFE relative to water. While most amino acids are helix breakers in water, only charged and neutral polar, β-substituted residues (in addition to Gly and Pro) remain helix breakers in 40% TFE. These large increases in propensity, however, cannot completely account for the extent of helix formation shown by protein fragments in 40% TFE. Table 5 gives the ellipticities, both in water and in 40% TFE, of three peptide fragments from β-lactoglobulin studied by Hamada et al. (1995); the observed values are compared with ones predicted from the helix-coil parameters given here. The observed ellipticities in water are low but they are several times larger than the predicted values. This behavior occurs commonly and it is caused by the presence of helix-stabilizing side-chain interactions, including H-bonds, salt bridges, and nonpolar interactions (see Muñoz & Serrano, 1994; Chakrabartty & Baldwin, 1995; Creamer & Rose, 1995). Table 5 indicates that some of these side-chain interactions still contribute to helix stability in 40% TFE. Measurements by Albert and Hamilton (1995) suggest that nonpolar side-chain interactions are still present in 40% TFE. Thus, the problem of relating helix formation in 40% TFE and in water involves side-chain interactions as well as the intrinsic helix propensities.

Although all amino acids show some increase in helix propensity in TFE relative to water, the increases are highly variable, indicating that the helix propensities of the amino acids in 40% TFE are not related in a simple manner to their values in water. This result is not surprising. Several factors are involved in determining values of helix propensity and most of these factors undergo changes from water to 40% TFE (see below). The nonpolar amino acids provide a logical starting point for analyzing the factors that determine helix propensity. Neutral polar amino acids can interfere with helix formation by H-bonding to backbone CO and NH groups. Charged and neutral polar amino acids with short side chains are stronger helix breakers than ones with longer side chains, probably for this reason (Padmanabhan et al., 1996).

Role of water in determining relative helix propensities of nonpolar amino acids

Three factors that change upon helix formation have been used to rationalize the differences in helix propensities of the amino acids: loss of side-chain conformational entropy, burial of nonpolar surface, and van der Waals interactions between the side chain and the helix matrix. The loss of side-chain entropy has been estimated in different ways by Creamer and Rose (1992, 1994), Blaber et al. (1994), Lee et al. (1994), and Wang and Purisima (1996) for the nonpolar amino acids. These different estimates agree with each other within a factor of 2, and the mean values agree within a factor of 2 with the observed values of $\Delta\Delta G$ in water determined from alanine-based peptides (Fig. 5). The amount of nonpolar surface area buried upon helix formation shows only small differences for the various amino acids beyond Cβ (Richmond & Richards, 1978; Wang & Purisima, 1996). Blaber et al. (1994) and Yang and Honig (1995) argue nevertheless that the differences are significant, while Wang and Purisima (1996) find that they are small compared with either the changes in side-chain entropy or the observed values of $\Delta\Delta G$.

The novel approach of computing the entire partition functions of the helix and the coil (Wang & Purisima, 1996) enables each of the individual factors contributing to the values of $\Delta\Delta G$ to be dissected. On the whole, Wang and Purisima find close agreement between their predicted values of the differences between helix propensities in alanine-based helices and the observed values, al-

Table 5. *Measured and predicted ellipticities of sequences from β-lactoglobulin in water and 40% TFE*

	$-[\theta]_{222}$ (deg cm^2 dmol^{-1})			
	Water		40% TFE	
Sequence	Measured[a]	Predicted[b]	Measured[a]	Predicted[b]
Ac-DIQKVAGTWYSLAMAASD-NH_2	4,000	300	28,000	7,000
Ac-WENGECAQKKIIAEKTK-NH_2	3,000	0	26,500	10,900
Ac-YEVDDEALEKFDKALKA-NH_2	6,500	100	22,500	6,900

[a]Hamada et al., 1995.

[b]Calculated from Equations 3 and 4, using the the helix-coil model and helix parameters reported in either Table 3 (water) or Table 5 (40% TFE).

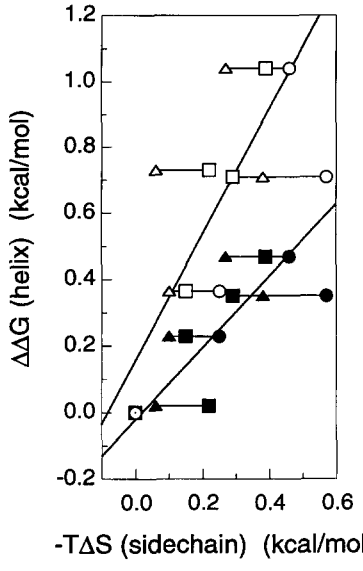

Fig. 5. Correlation between $\Delta\Delta G$(helix) and side-chain conformational entropy of the nonpolar amino acids. Helix propagation free energies, relative to alanine, in water (open symbols) and 40% TFE (filled symbols) are given on the *y*-axis. The values of $-T\Delta S$ are those reported by Wang and Purisima (1996) (circles), Blaber et al. (1994) (squares), and Creamer and Rose (1994) (triangles). Solid lines indicate the best correlation: $\Delta\Delta G(\text{helix})_{H_2O} = 0.16 - 1.9\ T\Delta S$, $R = 0.81$; $\Delta\Delta G(\text{helix})_{TFE} = -0.02 - 1.1\ T\Delta S$, $R = 0.90$. Data for methionine are not included because the side chain is partly polar. Data for tyrosine and tryptophan are not included because aromatic contributions to the ellipticity lead to uncertainty in the determined parameters (Chakrabartty et al., 1993a; see also Results).

though the agreement is not close in the case of $\Delta\Delta G$ for Ala–Gly. They point out that the observed values of $\Delta\Delta G$ are two-fold larger than the estimated values of $-T\Delta S$(conf), and they attribute the difference to van der Waals interactions between the side chains and the helix matrix. Figure 5 provides additional evidence for the conclusion that the observed values of $\Delta\Delta G$ in water are significantly larger than the estimated values of $-T\Delta S$(conf). The observed values of $\Delta\Delta G$ are two-fold smaller in 40% TFE than in water. Thus, the choice of solvent affects the observed values of $\Delta\Delta G$, which is not expected if the $\Delta\Delta G$ values depend only on side-chain entropies.

Why should the choice of solvent affect the differences between the helix propensities of the nonpolar amino acids? There have been two recent suggestions about how water might affect the relative helix propensities of the nonpolar amino acids. Ben-Naim (1991) and Yang et al. (1992) point out that the peptide CO group can still hydrogen bond to water after formation of an isolated helix, and this behavior affects the enthalpy of helix formation. Avbelj and Moult (1995) note that the dipoles of the peptide CO and NH groups interact unfavorably in the α-helical compared with the β-strand conformation. Water dipoles interact with peptide dipoles in the backbone so as to relieve the unfavorable helical interactions. This behavior depends on access of water to the helix backbone, which is sterically hindered by β-branched and bulky nonpolar side chains. Avbelj and Moult (1995) argue that this effect should be an important determinant of the relative helix propensities of the nonpolar amino acids. Calculations by Wang and Purisima (1996) do not, however, support this suggestion.

Both effects, hydrogen bonding between water and peptide CO groups in the helix and screening of unfavorable peptide dipolar interactions in the helix by water dipoles, should make the helix stronger in water than in 40% TFE. The opposite behavior is observed, however, indicating that factors that favor helix formation in 40% TFE must dominate. A main favorable factor is strengthening the peptide hydrogen bond by reducing the extent of solvation by water of the peptide CO and NH groups in the coil form (reviewed by Cammers-Goodwin et al., 1996). The fact that TFE has a lower basicity than water (Llinas & Klein, 1975) is an important factor in this behavior. Although strengthening of the peptide hydrogen bond is undoubtedly a main factor by which TFE promotes helix formation, it may affect the relative helix propensities of the amino acids, as well as causing general increases in helix propensity, because the access of water to the peptide backbone in the coil form depends on the size and β-branching of the side chain.

Comparison of peptide and protein helix propensities

It is relevant to consider whether peptide helix propensities measured in 40% TFE correlate better with protein results than values measured in water. Results are shown here only for T4 lysozyme (Blaber et al., 1993) because the results measured in alanine-based peptides correlate better with the T4 lysozyme data than with the barnase data (see Chakrabartty et al., 1994). At first glance, there is little difference when protein data are compared with peptide helix propensities in either water or TFE (Fig. 6A). The correlation coefficient for values determined in 40% TFE ($R = 0.89$) is a trifle better than for those determined in water ($R = 0.85$), but the range of peptide values in 40% TFE is larger than the range of

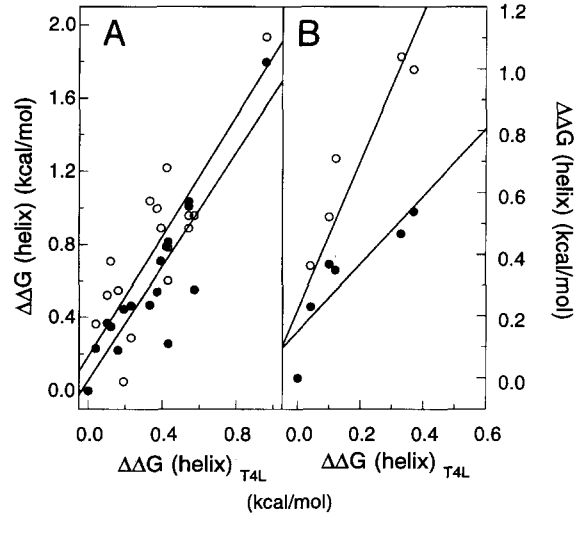

Fig. 6. Correlation between $\Delta\Delta G$(helix) measured in alanine-based peptides and in T4 lysozyme. Helix propagation free energies, relative to alanine, in water (●) and 40% TFE (○) are given on the *y*-axis for all amino acids (panel A) or the nonpolar amino acids (panel B). $\Delta\Delta G$(helix) values reported by Blaber et al. (1994) for substitutions in T4 lysozyme are given on the *x*-axis. Solid lines indicate the best correlation: Panel A: $\Delta\Delta G(\text{helix})_{H_2O} = 0.18 + 1.6\ \Delta\Delta G(\text{helix})_{T4L}$, $R = 0.85$; $\Delta\Delta G(\text{helix})_{TFE} = 0.06 + 1.6\ \Delta\Delta G(\text{helix})_{T4L}$, $R = 0.89$. Panel B: $\Delta\Delta G(\text{helix})_{H_2O} = 0.22 + 2.4\ \Delta\Delta G(\text{helix})_{T4L}$, $R = 0.93$; $\Delta\Delta G(\text{helix})_{TFE} = 0.15 + 1.1\ \Delta\Delta G(\text{helix})_{T4L}$, $R = 0.88$.

protein values, just as in water, and the slope of the correlation line (1.6, versus 1.0 expected for exact correlation) is the same in 40% TFE as in water. A different picture emerges, however, when only the nonpolar amino acids are compared with protein helix results (Fig. 6B). The slope of the correlation line relating the peptide and protein results is larger than 1 (2.4) in water, but in 40% TFE it is close to 1 (1.1). The correlation coefficients in 40% TFE and in water are nearly the same when only the nonpolar amino acids are considered. Figure 6B suggests that 40% TFE does mimic in part the mixture of buried and solvent-exposed environments experienced by protein helices (compare Waterhous & Johnson, 1994).

Convergence of N-cap propensities in 40% TFE

In the Zimm and Bragg (1959) and Lifson and Roig (1961) theories, the N-cap residue is classified as a coil residue because it is half inside, half outside, the helix. As a coil residue, the N-cap amino acid does not affect the stability of the helix in traditional helix-coil models. Measurements of the frequencies of different amino acids at the N-cap position (Richardson & Richardson, 1988) show, however, that certain amino acids occur with high frequencies. Presta and Rose (1988) proposed that side chain–main chain hydrogen bonds at the ends of helices are important in stabilizing protein helices, and this effect can explain the high frequencies of Asn, Ser, and Thr at the N-cap position. Mutagenesis experiments on barnase (Serrano et al., 1992) show that Gly is considerably more helix-stabilizing than Ala at the N-cap position, which demonstrates that the N-capping phenomenon is not restricted to polar amino acids and suggests that access of water to unsatisfied peptide NH groups at the N-terminal end of the helix may be important.

The measured N-cap propensities in alanine-based helices are closely correlated with amino acid frequencies at the N-cap position in protein helices; the largest N-cap propensities observed for Asn, Ser, Thr, and Gly (Chakrabartty et al. 1993b, 1994; Doig & Baldwin, 1995). Each of the amino acids has a specific N-cap propensity. In 40% TFE, the hydrogen-bonding effectiveness of small polar residues is expected to increase, and thus the capping propensities of residues that form side chain–main chain hydrogen bonds (Asn, Ser, Thr) should increase. Conversely, the effect of differential access of water to the N terminus should decrease in 40% TFE, reducing the capping propensity of Gly. Because the N-cap free energy is defined relative to alanine in both water and 40% TFE, changes in N-cap propensity between water and 40% TFE are given relative to the change for alanine. Because we cannot determine whether alanine is a stronger N-cap residue in water or TFE, we cannot interpret the direction of the changes observed in the N-cap propensities. Relative changes among the amino acids, as well as the order and range of N-cap propensities observed in 40% TFE and water can, however, be directly compared. Relatively uniform behavior is observed for Asn, Ser, and Thr, suggesting that their hydrogen-bonding effectiveness may be a factor in the observed changes in N-cap value. The N-cap propensities of these small polar residues all decrease, however, relative to alanine. The convergence toward a common N-cap value supports the view (compare Jasanoff & Fersht, 1994) that water plays a central role in determining the diverse N-cap values shown by the amino acids. To test this postulate, N-cap values should be measured in the absence of water, for example, in pure TFE.

Materials and methods

Peptide synthesis and circular dichroism measurements

Peptides were synthesized by standard solid phase methods, using pentafluorophenyl esters of 9-fluorenylmethoxycarbonyl amino acids for all coupling reactions except for coupling reactions of Ser and Thr, where 3,4-dihydro-4-oxo-1,2,3-benzotriazine esters were used. The N termini were either acetylated with acetic anhydride or left unblocked. Rink resin (Advanced Chemtech, Louisville, KY) was used to provide carboxyamidated C termini. The peptides were cleaved from the resin with 95% trifluoroacetic acid, 5% anisole. Ethanedithiol and thioanisole were added if Cys was present. Peptides were purified by reverse phase FPLC. Peptide identity was confirmed by FAB mass spectrometry. The synthesis and purification of some peptides, as indicated in Table 3, have been previously described (Chakrabartty et al., 1993a; 1993b; 1994).

CD measurements were made on an Aviv 60DS spectropolarimeter in quartz cuvettes with 1.0 cm pathlengths and are reported as mean residue ellipticity in units of deg cm^2 $dmol^{-1}$. The concentration of peptide stock solutions was determined by measuring tyrosine absorbance at 275 nm in 6.0 M guanidine hydrochloride, 20 mM potassium phosphate, pH 6.5, using ϵ_{275nm} = 1,450 M^{-1} cm^{-1} (Brandts & Kaplan, 1973; Chakrabartty et al., 1993b). Ellipticity at 222 nm, $-[\theta]_{222}$, was measured at 0 °C in 40% TFE, 100 mM sodium chloride, and 1 mM each of sodium borate, sodium citrate, and sodium phosphate. 0.1 mM dithiothreitol was added when Cys-containing peptides were examined. Acetylated peptides were measured at pH 7.0; unacetylated peptides were measured at pH 8.5 where the α-amino group is neutral and the ϵ-amino groups of Lys residues are charged. The protonation state of the ionizable groups was verified by performing pH titrations on several unacetylated peptides. pK_a values of the α-amino groups were found to vary between 7.0 and 8.0, depending on the identity of the N-terminal amino acid residue. The pK_a value of the ϵ-amino group of Lys residues was found to be relatively constant at 9.0.

Determination of the helix initiation parameter

The total NH exchange kinetics reported by Rohl et al. (1992) for a series of peptides of sequence Ac-$(AAKAA)_n$Y-NH_2 and chain lengths varying from 6–51 residues were used to redetermine the helix initiation parameter, v^2, in water when capping is introduced into the helix-coil model. Equation 2, describing the relative proton occupancy, $f_o(t)$, as a function of time, was fitted to the exchange curves.

$$f_0(t) = \sum_{i=1}^{N_r} (1/N_r) \exp[-(k_c[1 - f_B(i)])t] \quad (2)$$

The chemical exchange rate constant, k_c, was held constant at 0.018 min^{-1}, the value reported by Rohl et al. (1992) for the unstructured six-residue peptide in this series. The probability that residue i is hydrogen bonded, $f_B(i)$, is calculated from the helix coil model. n(Ac) was fixed at 5.9, the value determined below. n(K) was arbitrarily set to 1.0 because the calculated exchange curves are insensitive to the value of this parameter. All C-cap parameters were set to 1.0 (see below). The peptides were treated as homopolymers with respect to the helix nucleation and propagation parameters, and average values of $\langle v^2 \rangle$ = 0.0013 ± 0.0001 and $\langle w \rangle$ = 1.50 ± 0.01 in water were fitted from the data using the program NONLIN (Johnson et al., 1981) on a Silicon Graphics

Indigo. Reported errors are the 67% confidence intervals determined by NONLIN.

Use of helix-coil theory to predict ellipticity

The helix propagation and N-cap propensities for the amino acids in water and 40% TFE were determined by fitting the helix-coil model to the observed ellipticity of the peptides in Tables 1 and 3. Ellipticity at 222 nm, θ_{obs}, is assumed to be linearly related to mean helix content, f_H:

$$\theta_{obs} = f_H * (\theta_H * (1 - x/N_r) - \theta_C) + \theta_C, \quad (3)$$

where $\theta_C = 640$ deg cm^2 dmol^{-1} is the ellipticity of the random coil at 0 °C (Scholtz et al., 1991b). $\theta_H = -42{,}500$ m deg cm^2 dmol^{-1} is the ellipticity of a complete helix of infinite length at 0 °C (Scholtz et al., 1995). The term $(1 - x/N_r)$ is a correction to θ_H for end effects, where x is the number of peptide COs that are not hydrogen bonded in a complete helix (Chen et al., 1974; Scholtz et al., 1991). The N-terminal acetyl and C-terminal carboxamide blocking groups are assumed to be capable of forming helical hydrogen bonds. For carboxyamidated peptides, consequently, $x = 3$. The mean helix content of a peptide, f_H, is defined as:

$$f_H = \langle n_H \rangle / (N_r - 2), \quad (4)$$

where $\langle n_H \rangle$ is the average number of helical hydrogen bonds formed and $N_r - 2$ is the number of hydrogen bonds possible in a peptide with N_r residues. Mean helix content, f_H, is calculated from the helix-coil partition function and is a function of the chain length, N_r, and the helix parameters (w, v, n, and c) of the component amino acids.

C-cap interactions have been previously observed not to vary significantly among the different amino acids in water (Chakrabartty et al., 1993b; 1994; Doig & Baldwin, 1995). In addition, most amino acids do not show strong preferences for the C-cap position in protein helices (Richardson & Richardson, 1988). Consequently, all C-cap weights have been set to 1.0. The value of $v^2 = 0.0013$ determined above was used for all amino acids in both water and 40% TFE. The helix-coil transitions of alanine-based peptides with chain lengths of 7, 12, 17, and 22 residues, from 0–35% trifluoroethanol, can be fitted approximately with a constant nucleation parameter (P. Luo & R.L. Baldwin, pers. comm.). This result indicates that the helix nucleation parameter does not change significantly between water and 40% TFE.

The N-cap parameters and helix propagation propensities of the amino acids are fitted from the peptide sequences given in Tables 1 and 3. The peptide sequences used to determine the helix parameters in water or 40% TFE do not represent all amino acids equally. To minimize the effect of a biased data set, each helix parameter was fitted from the subset of data that is most sensitive to the parameter being measured as described below. In addition, some peptides are treated as homopolymers or partial homopolymers, with the host residues assigned an average helix propagation propensity. This approximation allows the helix parameters of a guest residue to be accurately determined without propagating errors in the helix parameters of the host residues into the parameters determined for the guest residues. Although some approximations have been used in evaluating the helix parameters, the set of determined parameters and the complete model without approximations reproduce the helix contents of the peptides in the data set well with a RMS deviation between observed and predicted f_H of 0.039 in water and 0.043 in 40% TFE.

All nonlinear least-squares fitting was accomplished using NONLIN, and 67% confidence intervals were evaluated when such error analysis could be performed. Confidence intervals were generally within ±15% of the fitted parameter value. When only a single measurement of a particular parameter was made, the error was estimated by calculating the effect of uncertainty in both the measured ellipticities and the value of θ_H. An uncertainty of a ±500 deg cm^2 dmol^{-1} error in the ellipticity of a test peptide corresponds to a 10% error in the determined value of the helix propagation propensity of a guest residue substituted in a central position of a reference peptide with mean helix content of 70–80%. An uncertainty in the value of θ_H of ±1,000 deg cm^2 dmol^{-1} translates into an error of ±1–2% in helix content for peptides with helix contents between 40 and 60%. Accounting for this additional source of error, we estimate the error of the helix parameters is approximately ±15% of the determined value.

As peptides approach 100% helix content, the error introduced by uncertainty in the value of θ_H results in much larger uncertainties in helix content. In addition, at very high helix contents, the values of the helix parameters are less sensitive to changes in helix content. These two factors make it very difficult to obtain accurate helix parameters from peptides with high helix content, and it is for this reason that the helix parameters in 40% TFE are determined from a different set of peptides than those used to determine the helix parameters in water.

Determination of helix parameters in water

Helix propagation and N-cap propensities in water were determined from the mean helix contents of the peptides in Table 1. N-cap parameters were determined from the peptides in the series NX-GY17. NA-GY17 serves as the reference and substitutions were made at the most N-terminal position in the test peptides. The peptides in the NX-GY17 series were treated as partial homopolymers with w(A,K) = 1.503 and w(G) = 0.048. The average value of w(A,K) was determined by fitting the ellipticity of the reference peptide, NA-GY17. The N-cap parameter of Lys$^+$ is arbitrarily set to 1.0 because the calculated helix contents are insensitive to the value of this parameter. w(G) was determined as described below. The N-cap parameter of the substituted amino acid, or acetyl group, was fitted from the measured ellipticity of the test peptide.

w(G) was determined from the helix contents of peptides YGGAK, YGG-G17, YGG-G12, and YGG-G7. YGGAK serves as the reference peptide in this series, and the test peptides contain single glycine substitutions at interior positions. The peptides were treated as partial homopolymers with w(A,K) = 1.580, n(G) = 3.9, n(Ac) = 5.9, and n(K) = 1.0. The average value of w(A,K) was determined by fitting the helix content of the reference peptide YGGAK. The N-cap parameter of Lys$^+$ is arbitrarily set to 1.0 because the calculated helix contents are insensitive to the value of this parameter. The N-cap parameters of glycine and the acetyl group were determined as described above.

The determination of the best values of v^2, w(G), n(G), and n(Ac) are interdependent: the fitting of the exchange data to determine v requires knowledge of n(Ac); determination of the N-cap parameters requires knowledge of v^2 and w(G); and determination of w(G) requires knowledge of the value of v^2, n(Ac), and n(G). These parameters, consequently, were iteratively fitted according

to the procedures described above until a self-consistent parameter set was obtained.

All of the peptides in the data set are alanine-based, with Lys and Gln residues included for solubility. Consequently, the w-values of Ala, Lys^+, and Gln are relatively well determined and can be fitted from the subset of peptides that contains only Ala, Lys^+, and Gln at central positions: YGAK, YGGGAK, AcA-GY12, YGG-1Q, YGG-3Q, 2Q-GGY, N2Q-GGY, Qref, Q12, YG-ZC17, YGGAK, YGG-G17, YGG-G12, YGG-G7, NQ-GY17, NA-GY17, Ac-GY18, YG-ZC17, NG-GY17, NK-GY17, and NY-GY17. These peptides were treated as heteropolymers. N-cap parameters were fixed at the values in Table 2 and the helix propensities of Ala, Lys^+, and Gln were simultaneously fitted from the data.

Helix propagation propensities for Cys^0, Phe, Ile, Leu, Met, Asn, Pro, Gln, Arg^+, Ser, Thr, Val, Trp, and Tyr were fitted using only the helix contents of peptides containing single or multiple Ala → X substitutions at central sites. For example, the helix propagation propensity of Leu was fitted from the helix contents of peptides YGG-1L, YGG-3L, 2L-GGY, and N2L-GGY. The peptides were treated as heteropolymers with all parameters except the w-value of the substituted amino acid set to the values given in Table 2. Because the N-cap parameter of Cys^0 was not determined, it was arbitrarily set to 1.0. The reported helix propagation parameter for Cys^0, consequently, is an apparent value. The helix propensity of Pro cannot be determined from this data set because the only peptide with a proline substitution (YGG-1P) shows no measurable helix formation. Consequently, only an upper limit for the value of w(P) can be estimated.

The helix propensities of Asp^-, Asp^0, Glu^-, Glu^0, His^+, His^0, and Arg^+ were determined from the helix contents of peptides that contain a single charged residue for Ala substitution in a neutral alanine and glutamine host peptide. The interaction between the charged residue and the helix dipole is not explicitly included in the model described here, although this interaction is partially taken into account by the N-capping propensities. In determining the helix propagation propensities of these charged residues, we used only peptides in which the charged residue is substituted at a central position in order to minimize contributions from charge-helix dipole interactions. These peptides do not have the tyrosine separated from the helix by glycine residues, and consequently the tyrosine is expected to contribute to the ellipticity at 222 nm (Chakrabartty et al., 1993a).

The peptides were treated as homopolymers with an average w value assigned to all residues except the single charged residue. N-cap propensities for residues other than the acetyl group and the single charged residue are arbitrarily set to 1.0 because the calculated helix contents are insensitive to these parameters. An average value of w(A,Q) = 1.359 was determined by fitting the helix content of the reference peptide AQhost. The helix propensity of the charged residue was fitted from helix contents of the peptides containing single substitutions. The N-cap parameters for His^+, Glu^0, and Asp^0 were not determined and were arbitrarily set to 1.0. The fitted propagation parameters for these residues consequently are apparent values of w.

Determination of helix parameters in TFE

Helix propagation and N-cap propensities in 40% TFE were determined from the mean helix contents of the peptides in Table 3. N-cap parameters were determined from the peptides in the series NX-GY12. NA-GY12 serves as the reference and substitutions were made at the most N-terminal position in the test peptides. The peptides in the NX-GY12 series were treated as partial homopolymers with w(A,K) = 2.323, w(G) = 0.098, and n(A,K) = 1.00. The average value of w(A,K) was determined by fitting the ellipticity of the reference peptide, NA-GY12. w(G) was determined as described below. The N-capping parameter of the substituted amino acid, or acetyl group, was fitted from the measured ellipticity of the test peptide.

Helix propagation propensities for Leu, Ile, Met, Asn, Gln, Ser, Thr, Val, Phe, His^+, His^0, and Cys^0 were fitted from peptides in the X-AR12 series. Test peptides contain single Ala → X substitutions at a central site. The peptides were treated as partial homopolymers with w(A,K) = 2.112 and n(A,K) = 1.00. The average w(A,K) was determined from the ellipticity of the reference peptide A-AR12. The propagation parameter, w, of the substituted amino acid was fitted from the measured ellipticity of the test peptide. For Phe, His^+, His^0, and Cys^0, N-capping parameters are not determined and were arbitrarily set to 1.0. The reported propagation parameters consequently are apparent values of w.

The helix propagation propensities for Arg^+ and Lys^+ were determined from peptides in the X-AQ12 series. The peptides were treated as partial homopolymers with w(A,Q) = 2.699 and n(A,Q) = 1.0. w(G), n(Ac), and the N-cap parameter of the substituted amino acids were held constant at the values given in Table 4. The average w(A,Q) was determined from the ellipticity of the reference peptide A-AQ12. The propagation parameter, w, of the substituted amino acid was fitted from the measured ellipticity of the test peptide.

w(A) was fitted from the helix contents of the reference peptides NX-GY, A-AR12, A-AQ12, and AK12. The peptides were treated as heteropolymers with all helix parameters except w(A) held constant at the values in Table 4. Apparent w-values for Asp^- and Glu^- were determined from the peptides AE12 and AD12, respectively. The peptides were treated as heteropolymers with all parameters except for w(D) or w(E) held constant at the values given in Table 4. N-cap parameters for Asp^- and Glu^- were arbitrarily set equal to 1.0. The reported propagation parameters are consequently apparent values of w.

Because Gly and Pro are strong helix breakers, 12-residue alanine-based peptides with single Gly or Pro substitutions are expected to have very low helix contents. The helix propensities of Gly and Pro consequently cannot be well determined from peptides of this design. Instead, single Gly or Pro substitutions were made at central positions in longer, more helical peptides. The peptides in the YGG-X series were treated as heteropolymers with all helix parameters except for w(G) or w(P) held constant at the values in Table 4. n(Y) is not determined by the peptides in this data set and is arbitrarily set equal to 1.0.

The determination of w(G) requires the knowledge of n(Ac), n(K), n(G), w(K), and w(A). w(G), however, is a constant in the determination of these helix parameters. Furthermore, n(Q), n(R), w(Q), and w(R) are constants in the determination of w(A). Consequently, the propagation parameters for Ala, Gly, Lys^+, Gln, and Arg^+ and the N-cap parameters for acetyl, Gly, Lys^+, Gln, and Arg^+ were iteratively fitted until self-consistent values were obtained.

Acknowledgments

We thank Dr. Andrew Doig for extensive discussion and critical review of the modifications to the helix-coil model. Dr. Niels Andersen indepen-

dently concluded that the description of capping should be modified, and we acknowledge his discussion. FAB mass spectra were provided by the Mass Spectrometry Facility, University of San Francisco, supported by NIH grant RR 01614. This work was supported by a grant from the National Institutes of Health (GM 31475).

References

Albert JS, Hamilton AD. 1995. Stabilization of helical domains in short peptides using hydrophobic interactions. *Biochemistry 34*:984–990.

Armstrong KM, Baldwin RL. 1993. Charged histidine affects α-helix stability at all positions in the helix by interacting with the backbone charges. *Proc Natl Acad Sci USA 90*:11337–11340.

Avbelj F, Moult J. 1995. Role of electrostatic screening in determining protein main chain conformational preferences. *Biochemistry 34*:755–764.

Ben-Naim A. 1991. The role of hydrogen bonds in protein folding and protein association. *J Phys Chem 95*:1437–1444.

Blaber M, Zhang X, Matthews BW. 1993. Structural basis of amino acid α-helix propensity. *Science 260*:1637–1640.

Blaber M, Zhang X, Lindstrom, JD, Pepiot SD, Baase WA, Matthews BW. 1994. Determination of α-helix propensity within the context of a folded protein. Sites 44 and 131 in bacteriophage T4 lysozyme. *J Mol Biol 235*:600–624.

Brandts JF, Kaplan LJ. 1973. Derivative spectroscopy applied to tyrosyl chromophores. studies on ribonuclease, lima bean inhibitors, insulin, and pancreatic trypsin inhibitor. *Biochemistry 12*:2011–2024.

Cammers-Goodwin A, Allen TJ, Oslick SL, McClure KF, Lee JH, Kemp DS. 1996. Mechanism of stabilization of helical conformations of polypeptides by water-containing trifluoroethanol. *J Am Chem Soc 118*:3082–3090.

Chakrabartty A, Baldwin RL. 1995. Stability of α-helices. *Adv Prot Chem 46*:141–176.

Chakrabartty A, Doig AJ, Baldwin RL. 1993a. Helix capping propensities in peptides parallel those in proteins. *Proc Natl Acad Sci USA 90*:11332–11336.

Chakrabartty A, Kortemme T, Baldwin RL. 1994. Helix propensities of the amino acids measured in alanine–based peptides without helix-stabilizing side-chain interactions. *Prot Sci 3*:843–852.

Chakrabartty A, Kortemme T, Padmanabhan S, Baldwin RL. 1993b. Aromatic side-chain contributions to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry 32*:5560–5565.

Chen Y-H, Yang JT, Chau KH. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry 13*:3350–3359.

Creamer TP, Rose GD. 1992. Side-chain entropy opposes α-helix formation but rationalizes experimentally determined helix-forming propensities. *Proc Natl Acad Sci USA 89*:5937–5941.

Creamer TP, Rose GD. 1994. α-Helix-forming propensities in peptides and proteins. *Proteins Struct Funct Genet 19*:85–97.

Creamer TP, Rose GD. 1995. Interactions between hydrophobic side chains within α-helices. *Prot Sci 4*:1305–1314.

Doig AJ, Baldwin RL. 1995. N- and C-capping preferences for all 20 amino acids in α-helical peptides. *Protein Sci 4*:1325–1336.

Doig AJ, Chakrabartty A, Klingler TM, Baldwin RL. 1994. Determination of free energies of N-capping in α-helices by modification of the Lifson-Roig helix-coil theory to include N- and C-capping. *Biochemistry 33*:3396–3403.

Groebke K, Renold P, Tsang KY, Allen TJ, McClure KF, Kemp DS. 1996. Template-nucleated alanine-lysine helices are stabilized by position-dependent interactions between the lysine side chain and the helix barrel. *Proc Natl Acad Sci USA 93*:4025–4029.

Hamada D, Kuroda Y, Tanaka T, Goto Y. 1995. High helical propensity of the peptide fragments derived from β-lactoglobulin, a predominantly β-sheet protein. *J Mol Biol 254*:737–746.

Harper ET, Rose GD. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry 32*:7605–7609.

Horovitz A, Matthews JM, Fersht AR. 1992. α-Helix stability in proteins. II. Factors that influence stability at an internal position. *J Mol Biol 227*:560–568.

Huyghues-Despointes BMP, Scholtz JM, Baldwin RL. 1993. Effect of a single aspartate on helix stability at different positions in a neutral alanine-based peptide. *Protein Sci 2*:1604–1611.

Jasanoff A, Fersht AR. 1994. Quantitative determination of helical propensities from trifluoroethanol titration curves. *Biochemistry 33*:2129–2135.

Johnson ML, Correia JJ, Yphantis DA, Halvorson HR. 1981. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophysical J 36*:575–588.

Kemp DS, Allen TJ, Oslick SL, Boyd JG. 1996. The structure and energetics of helix formation by short templated peptides in aqueous solution. 2. Characterization of the helical structure of Ac-Hel_1-Ala_6-OH. *J Am Chem Soc 118*:4240–4248.

Lee KH, Xie D, Freire E, Amzel LM. 1994. Estimation of changes in side chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins Struct Funct Genet 20*:68–84.

Lifson S, Roig A. 1961. On the theory of helix-coil transition in polypeptides. *J Chem Phys 34*:1963–1974.

Llinas M, Klein MP. 1975. Charge relay at the peptide bond. A proton magnetic resonance study of solvation effects on the amide electron density distribution. *J Am Chem Soc 97*:4731–4737.

Muñoz V, Serrano L. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct Biol 6*:399–409.

Nelson JW, Kallenbach NR. 1986. Stabilization of the ribonuclease S-peptide α-helix by trifluoroethanol. *Proteins: Struct Funct Genet 1*:211–217.

Nelson JW, Kallenbach NR. 1989. Persistence of the α-helix stop signal in the S-peptide in trifluoroethanol solutions. *Biochemistry 28*:5256–5261.

Padmanabhan S, York, EJ, Stewart JM, Baldwin RL. 1996. Helix propensities of basic amino acids increase with the length of the side-chain. *J Mol Biol 257*:726–734.

Presta LG, Rose GD. 1988. Helix signals in proteins. *Science 240*:1632–1641.

Qian H, Schellman JA. 1992. Helix-coil theories: A comparative study for finite length polypeptides. *J Phys Chem 96*:3987–3994.

Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of α helices. *Science 240*:1648–1652.

Richmond TJ, Richards FM. 1978. Packing of alpha-helices: Geometrical constraints and contact areas. *J Mol Biol 119*:537–555.

Rohl CA, Scholtz JM, York EJ, Stewart JM, Baldwin RL. 1992. Kinetics of amide proton exchange in helical peptides of varying chain lengths. Interpretation by the Lifson-Roig equation. *Biochemistry 31*:1263–1269.

Scholtz JM, Barrick D, York EJ, Stewart JM, Baldwin RL. 1995. Urea unfolding of peptide helices as a model for interpreting protein unfolding. *Proc Natl Acad Sci USA 92*:185–189.

Scholtz JM, Qian H, Robbins VH, Baldwin RL. 1993. The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry 32*:9668–9676.

Scholtz JM, Qian H, York EJ, Stewart JM, Baldwin RL. 1991. Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers 31*:1463–1470.

Segawa S-I, Fukuno T, Fujiwara K, Noda Y. 1991. Local structures in unfolded lysozyme and correlation with secondary structures in the native conformation: Helix-forming or -breaking propensity of peptide segments. *Biopolymers 31*:497–509.

Serrano L, Neira J-L, Sancho J, Fersht AR. 1992. Effect of alanine versus glycine in α-helices on protein stability. *Nature 356*:453–455.

Sönnichsen FD, Van Eyk JE, Hodges RS, Sykes BD. 1992. Effect of trifluoroethanol on protein secondary structure: An NMR and CD study using a synthetic actin peptide. *Biochemistry 31*:8790–8798.

Stapley BJ, Rohl CA, Doig AJ. 1995. Addition of side chain interactions to modified Lifson-Roig helix-coil theory. Application to energetics of phenylalanine-methionine interactions. *Protein Sci 4*:2383–2391.

Storrs RW, Truckses D, Wemmer DE. 1992. Helix propagation in trifluoroethanol solutions. *Biopolymers 32*:1695–1702.

Wang J, Purisima EO. 1996. Analysis of thermodynamic determinants in helix propensities of nonpolar amino acids through a novel free energy calculation. *J Am Chem Soc 118*:995–1001.

Waterhous DV, Johnson WC Jr. 1994. Importance of environment in determining secondary structure in proteins. *Biochemistry 33*:2121–2128.

Yang A-S, Honig B. 1995. Free energy determinants of secondary structure formation: I. α-Helices. *J Mol Biol 252*:351–365.

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

Zhou HX, Hull LA, Kallenbach NR, Mayne L, Bai Y, Englander SW. 1994. Quantitative evaluation of stabilizing interactions in a prenucleated α-helix by hydrogen exchange. *J Am Chem Soc 116*:6482–6483.

Zimm BH, Bragg JK. 1959. Theory of the phase transition between helix and random coil in polypeptide chains. *J Chem Phys 31*:526–535.

Supplementary material in electronic appendix

A computer program implementing the helix-coil model described here has been written in FORTRAN for the Silicon Graphics Irix platform and is available in the electronic appendix and by anonymous FTP from cmgm.stanford.edu in the directory /pub/helix/helix2. In the calculations described in this work, sequence specific interactions are assumed not to contribute to helix stability. To accommodate future work, however, we have incorpo-

rated some i, $i + 3$, and i, $i + 4$ side-chain interactions into the implementation of the Lifson-Roig based formalism. An independent implementation of this model has also been derived (B.J. Stapley & A.J. Doig, pers. comm.).

Addition of side-chain interactions to the Lifson-Roig theory has been previously described (Stapley et al., 1995; Scholtz et al., 1993). The statistical weights p and q are defined as equilibrium constants for i, $i + 4$, and i, $i + 3$ side-chain interactions, respectively, such that

$$\Delta G(i, i + 4) = -RT \ln[p] \tag{5a}$$

$$\Delta G(i, i + 3) = -RT \ln[q] \tag{5b}$$

Side-chain interactions are assumed to occur when the interacting residues and all intervening residues are in the h conformation: $p_{i-2,i+2}$ is assigned to the central residue, i, in the quintet hhhhh and $q_{i-2,i+1}$ is assigned to the central residue, i, in the quintets hhhhh and hhhhc. Interactions occurring between the residues at N-cap and N3 positions (capping boxes; Harper & Rose, 1993) are incorporated into the helix model (A. Chakrabartty, A.J. Doig, B.J. Stapley, CAR unpubl.) through the statistical weight, r, which is defined as the equilibrium constant for the capping interaction in the background of a complete helix:

$$\Delta G(\text{capping box}) = -RT \ln[r] \tag{6}$$

The statistical weight $r_{i-2,i+1}$ is assigned to the central residue, i, in the quintets chhhc and chhhh. Residues $i - 2$ and $i + 1$ are the N-cap and N3 residues, respectively. The reference for all sequence-specific interactions is alanine at both interacting positions. The assignment of statistical weights to particular quintets is accomplished by making the following replacements of elements in matrix M_i above:

$$(1,1) = w_i\, p_{i-2,i+2} q_{i-2,i+1}$$

$$(1,2) = w_i\, q_{i-2,i+1}\, c_{i-1}$$

$$(3,1) = w_i\, r_{i-2,i+1} n_{i-1}$$

$$(3,2) = w_i\, r_{i-2,i+1}\, n_{i-1}\, c_{i-1}$$

When all interaction parameters (p, q, r) are set to 1.0, this model reduces to the model described in the Helix-coil transition theory section above. The complete parameterization of the model, including the pairwise interactions described above, is currently underway in the laboratories of A.J. Doig and J.M. Scholtz.