9372 *Biochemistry* **2010,** *49,* 9372–9384
DOI: 10.1021/bi101156j

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Positional Effects on Helical Ala-Based Peptides[†]

Richard P. Cheng,*[,‡] Prashant Girinath,[§] Yuta Suzuki,[§] Hsiou-Ting Kuo,[‡] Hao-Chun Hsu,[‡] Wei-Ren Wang,[‡] Po-An Yang,[‡] Donald Gullickson,[§] Cheng-Hsun Wu,[‡] Marc J. Koyack,[§] Hsien-Po Chiu,[§] Yi-Jen Weng,[‡] Pier Hart,[∥] Bashkim Kokona,[∥] Robert Fairman,[∥] Tzu-En Lin,[‡] and Olivia Barrett[§]

[‡]*Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan,* [§]*Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, New York 14260-3000, and* [∥]*Department of Biology, Haverford College, Haverford, Pennsylvania 19041*

Received July 20, 2010

ABSTRACT: Helix−coil equilibrium studies are important for understanding helix formation in protein folding, and for helical foldamer design. The quantitative description of a helix using statistical mechanical models is based on experimentally derived helix propensities and the assumption that helix propensity is position-independent. To investigate this assumption, we studied a series of 19-residue Ala-based peptides, to measure the helix propensity for Leu, Phe, and Pff at positions 6, 11, and 16. Circular dichroism spectroscopy revealed that substituting Ala with a given amino acid (Leu, Phe, or Pff) resulted in the following fraction helix trend: KXaa16 > KXaa6 > KXaa11. Helix propensities for Leu, Phe, and Pff at the different positions were derived from the CD data. For the same amino acid, helix propensities were similar at positions 6 and 11, but much higher at position 16 (close to the C-terminus). A survey of protein helices revealed that Leu/Phe-Lys (i, $i + 3$) sequence patterns frequently occur in two structural patterns involving the helix C-terminus; however, these cases include a left-handed conformation residue. Furthermore, no Leu/Phe-Lys interaction was found except for the Lys−Phe cation−π interaction in two cases of Phe-Ala-Ala-Lys. The apparent high helix propensity at position 16 may be due to helix capping, adoption of a 3_{10}-helix near the C-terminus perhaps with Xaa−Lys (i, $i + 3$) interactions, or proximity to the peptide chain terminus. Accordingly, helix propensity is generally position-independent except in the presence of alternative structures or in the proximity of either chain terminus. These results should facilitate the design of helical peptides, proteins, and foldamers.

Helices represent one of the most common structural motifs in both proteins (*1*−*3*) and non-natural foldamers (*4*−*8*). One third of all protein residues adopt an α-helical conformation (*1*−*3*), leading to various experimental studies on α-helix formation energetics, including helix propensity (*9*−*27*), capping energetics (*18*, *22*, *28*−*39*), and intrahelical interactions (*21*, *22*, *25*, *40*−*66*). Helices are also the most prevalent structure in foldamers (*4*−*8*), which are non-natural polymers and oligomers that can adopt compact, well-defined three-dimensional conformations and structures. Foldamers such as peptoids (*67*), β-peptides (*4*, *68*, *69*), oligo-phenylene ethynylenes (*70*), and oligo-arylamides (*71*, *72*) have been designed to adopt helical conformations (*6*, *8*). Furthermore, intrahelical interactions (*73*−*77*) and helix propensity differences (*76*, *78*, *79*) have also been explored in helical foldamers. Therefore, fundamental studies of the helix−coil equilibrium are important for understanding helix formation in protein folding, and in foldamer design and applications.

The helix propensity of natural amino acids has been experimentally measured in random copolymers (*9*, *10*), monomeric helical peptides (*12*, *14*, *16*, *21*−*27*), coiled coil systems (*13*), and helices in natural proteins (*18*−*20*, *24*). In particular, experimental measurements on monomeric helical systems (*12*, *14*, *16*, *21*, *22*, *24*−*27*) have been coupled with statistical mechanical models based on either Zimm−Bragg (*80*) or Lifson−Roig (*81*) theory to obtain helix formation parameters, and thus absolute (nonrelative) helix forming energetics. Both theories assume that every amino acid can only adopt either a helical or a nonhelical conformation (*80*, *81*), effectively invoking a multistate equilibrium for helical peptides involving multiple residues. Furthermore, the probability of adopting a helical conformation for a given residue depends on its side chain (propensity) and the conformation of neighboring residues (cooperativity) (*80*, *81*). Using either statistical mechanical model, helix formation for every amino acid is represented by at least two parameters: helix initiation parameter and helix propagation parameter (i.e., helix propensity) (*80*, *81*). Helix capping parameters have also been incorporated into the modified Lifson−Roig theory (*12*, *34*, *35*, *38*). Further advancements in describing the helix−coil equilibrium have included intrahelical side chain−side chain interactions (*21*, *22*, *25*, *38*, *41*, *45*−*48*, *50*−*54*, *58*−*66*) and interaction between charged amino acids and the helix dipole (*38*, *41*, *66*, *82*). The influence of temperature (*83*−*89*) and trifluoroethanol (*87*, *88*, *90*) on the helix propensity parameter has also been studied.

The quantitative description of monomeric helix formation using statistical mechanical models has been built on a foundation based on experimentally derived helix propagation parameters (helix propensities) (*12*, *14*, *16*, *21*, *22*, *24*−*27*). In particular, seminal work by Baldwin and co-workers on such systems with minimal intrahelical side chain interactions has provided the helix propagation parameters and capping parameters for deducing the fraction helix of monomeric Ala-based peptides (*12*, *34*, *35*, *90*).

[†]This work was supported by the NYSTAR James D. Watson Investigator Program (R.P.C.), Kapoor funds (R.P.C.), The State University of New York at Buffalo (R.P.C.), the National Science Foundation (R.P.C., CHE0809633; R.F., MCB-0211754), National Taiwan University (R.P.C.), and the National Science Council (R.P.C., NSC-97-2113-M-002-019-MY2).

*To whom correspondence should be addressed. Phone: +886-2-33669789. Fax: +886-2-23636359. E-mail: rpcheng@ntu.edu.tw.

 Published on Web 10/06/2010

One basic assumption in these statistical mechanical models is that the helix propensity for a given amino acid does not change with position or peptide length (*80*, *81*). Interestingly, studies by Kemp showed that the helix propensity of Ala is peptide length-dependent (*85*, *91*). Although the position independence of helix propensity was observed in Ala-based monomeric helices (*92*) and a protein helix (*93*), the position dependence of helix propensity was observed in another set of Ala-based monomeric helices (*94*−*96*). However, the Ala-based helices had guest positions with varying immediate neighboring residues, thereby unable to completely rule out the immediate neighboring context as the reason for the apparent position dependence of helix propensity (*94*−*96*). Herein, we present studies on a series of Ala-based peptides with neutral non-Ala residues [leucine, phenylalanine, and pentafluorophenylalanine (Pff)] at various guest positions with the same immediate surrounding residues (±2 residues), to study the effect of these amino acids at different positions of a helical peptide and to measure the helix propensity for these amino acids at these different positions.

MATERIALS AND METHODS

General. All of the chemical reagents except those indicated otherwise were purchased from Aldrich. Organic and high-performance liquid chromatography (HPLC) solvents were from EMD Science and Merck Taiwan. *N*-9-Fluorenylmethoxycarbonyl (Fmoc) amino acids, 1-hydroxybenzotriazole (HOBt), and *O*-1*H*-benzotriazol-1-yl-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were from NovaBiochem, Fmoc-PAL-PEG-PS resin was from Applied Biosystems. Analytical reverse phase (RP) HPLC was performed on an Agilent 1100 series chromatography system using a Vydac C_{18} column (4.6 mm diameter, 250 mm length). Preparative RP-HPLC was performed on a Waters Breeze chromatography system using Vydac RP C_4 and C_{18} columns (22 mm diameter, 250 mm length). Mass spectrometry of the peptides was performed on a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) spectrometer (Bruker Daltonics Biflex IV) using α-cyano-4-hydroxycinnamic acid as the matrix. Determination of peptide concentrations was performed on a UV−vis spectrophotometer (Agilent 8453 or Jasco V-650). Circular dichroism (CD) spectra were recorded on a Jasco J715 or J815 spectrometer using a 1 mm path length cell. Each reported CD value was the mean of three wavelength scans or the mean of 61 readings at 222 nm. Data were normalized in terms of per residue molar ellipticity (degrees square centimeters per decimole).

Peptide Synthesis. Peptides were synthesized by solid phase peptide synthesis using Fmoc-based chemistry (*97*). For a typical peptide synthesis, Fmoc-PAL-PEG-PS (50 μmol) was swollen in *N*,*N*-dimethylformamide (DMF, 5 mL) for 30 min before the first coupling. The resin was then washed with DMF (5 mL, 5 × 1.5 min). This was followed by Fmoc deprotection with a 20% piperidine/DMF mixture (5 mL, 3 × 8 min). The resin was subsequently washed with DMF (5 mL, 5 × 1.5 min). A mixture of 3 equiv of the appropriately protected Fmoc amino acid, HOBt, and HBTU was dissolved in DMF (1 mL). Diisopropylethylamine (DIEA, 8 equiv) was then added to the solution. The solution was then mixed thoroughly and applied to the resin. The vial that contained the solution was rinsed with DMF (2 × 1 mL), and its contents were added to the reaction mixture. The coupling reaction was typically conducted for 45 min. The coupling times varied for different amino acids depending upon their positions in the sequence. The first amino acid was coupled for 8 h. The eighth to 14th residues that were attached to the resin were coupled for 1.5 h. For capping with acetic anhydride, a solution of Ac_2O (20 equiv), DIEA (20 equiv), and DMF (3 mL) was added to the resin. The reaction mixture was shaken for 2 h. The resin was subsequently washed with DMF (5 mL, 5 × 1.5 min) and CH_2Cl_2 (5 mL, 5 × 0.5 min) and lyophilized overnight.

Peptides were deprotected and cleaved off the resin when the resin was treated with a 95:5 trifluoroacetic acid (TFA)/triisopropylsilane mixture (10 mL) for 2 h. The reaction mixture was then filtered through glass wool, and the resin was washed with TFA (3 × 3 mL). The combined filtrate was then evaporated with a gentle stream of N_2. The resulting oil was washed with hexanes, dissolved in water, and lyophilized. The peptides (1 mg/mL aqueous solution) were analyzed using analytical RP-HPLC on a 25 cm C_{18} column (diameter of 4.6 mm) using a flow rate of 1 mL/min, a linear gradient (rate of 1% per minute) from 100 to 0% A (solvent A, 99.9% water and 0.1% TFA; solvent B, 90% acetonitrile, 10% water, and 0.1% TFA). Appropriate linear solvent A/solvent B gradients were used for purification on RP-HPLC preparative C_4 and C_{18} columns; all peptides were purified to greater than 98% purity. The identity of the peptides was confirmed by MALDI-TOF.

Circular Dichroism Spectroscopy. CD data were collected using a 1 mm path length cell. The concentration of the peptide stock solution was determined by the tyrosine absorbance in 6 M guanidinium chloride (ε_{276} = 1455, ε_{278} = 1395, ε_{280} = 1285, and ε_{282} = 1220) (*98*). CD measurements were reported at peptide concentrations of 70−80 μM in 1 M NaCl, 1 mM sodium phosphate, 1 mM sodium citrate, and 1 mM sodium borate (pH 7) at 0 °C. The data were analyzed using Kaleidagraph version 3.52 (Synergy Software). Each reported CD value was the mean of at least three determinations. Data were expressed in terms of mean residue molar ellipticity (degrees square centimeters per decimole). The mean residue molar ellipticity of the peptides was independent of peptide concentration (80−160 μM). The fraction helix of each peptide (f_{helix}) was calculated from the mean residue molar ellipticity at 222 nm and the number of backbone amides (N) using eq 1.

$$f_{helix} = \frac{[\theta]_{222}}{40000\left(1 - \frac{2.5}{N}\right)} \quad (1)$$

Helix Propensity of Amino Acids. The statistical mechanical parameters for the host residues Ala, Lys, Gly, Tyr, acetyl, and carboxyamide were derived numerically on the basis of modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*) using the least-squares method by minimizing the sum of the square of the difference between the calculated and experimental values. For all residues, v was set to 0.048 (*12*). The helix propensity of the amino acid at the guest position was then numerically derived from the f_{helix} of the corresponding peptide based on modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*). The free energy for helix formation (ΔG) was calculated as $-RT \ln(w)$. To obtain the potential side chain−side chain interaction energetics, the statistical weight for each specific intrahelical side chain−side chain interaction (p) was calculated from the experimental f_{helix} on the basis of the modified nesting block method (*41*, *66*, *99*). These calculations were performed using in-house computer code written in C++. The free energy of each specific side chain−side chain interaction (ΔG) was calculated as $-RT \ln(p)$.

Table 1: Sequences and Experimentally Measured Fraction Helix (f_{helix}) Values of Ala-Based Peptides

peptide	sequence[a]	f_{helix} Baldwin[b]	f_{helix}
KAla	Ac-YGG KAAAA KAAAA KAAAA K-NH_2	0.677	0.549 ± 0.006
KGly6	Ac-YGG KA**G**AA KAAAA KAAAA K-NH_2		0.349 ± 0.010
KGly11	Ac-YGG KAAAA KA**G**AA KAAAA K-NH_2		0.176 ± 0.005
KGly16	Ac-YGG KAAAA KAAAA KA**G**AA K-NH_2		0.289 ± 0.004
KLeu6	Ac-YGG KA**L**AA KAAAA KAAAA K-NH_2		0.513 ± 0.006
KLeu11	Ac-YGG KAAAA KA**L**AA KAAAA K-NH_2	0.546	0.502 ± 0.006
KLeu16	Ac-YGG KAAAA KAAAA KA**L**AA K-NH_2		0.616 ± 0.006
KLeu611	Ac-YGG KA**L**AA KA**L**AA KAAAA K-NH_2		0.467 ± 0.006
KLeu616	Ac-YGG KA**L**AA KAAAA KA**L**AA K-NH_2		0.557 ± 0.006
KLeu1116	Ac-YGG KAAAA KA**L**AA KA**L**AA K-NH_2		0.561 ± 0.010
KLeu61116	Ac-YGG KA**L**AA KA**L**AA KA**L**AA K-NH_2	0.492	0.469 ± 0.006
KPhe6	Ac-YGG KA**F**AA KAAAA KAAAA K-NH_2		0.455 ± 0.007
KPhe11	Ac-YGG KAAAA KA**F**AA KAAAA K-NH_2	0.426	0.426 ± 0.008
KPhe16	Ac-YGG KAAAA KAAAA KA**F**AA K-NH_2		0.494 ± 0.006
KPff6	Ac-YGG KA**Z**AA KAAAA KAAAA K-NH_2		0.320 ± 0.008
KPff11	Ac-YGG KAAAA KA**Z**AA KAAAA K-NH_2		0.264 ± 0.006
KPff16	Ac-YGG KAAAA KAAAA KA**Z**AA K-NH_2		0.391 ± 0.013
NCapAla	Ac-**A**A KAAAA KAAAA KAA GGY-NH_2		0.573 ± 0.011
CCapAla	Ac YGG AA KAAAA KAAAA KA**A**-NH_2		0.576 ± 0.017
NCapGly	Ac-**G**A KAAAA KAAAA KAA GGY-NH_2		0.512 ± 0.009
CCapGly	Ac YGG AA KAAAA KAAAA KA**G**-NH_2		0.421 ± 0.012
NCapLeu	Ac-**L**A KAAAA KAAAA KAA GGY-NH_2		0.459 ± 0.009
CCapLeu	Ac YGG AA KAAAA KAAAA KA**L**-NH_2		0.543 ± 0.013
NCapPhe	Ac-**F**A KAAAA KAAAA KAA GGY-NH_2		0.502 ± 0.011
CCapPhe	Ac YGG AA KAAAA KAAAA KA**F**-NH_2		0.388 ± 0.008
NCapPff	Ac-**Z**A KAAAA KAAAA KAA GGY-NH_2		0.377 ± 0.011
CCapPff	Ac YGG AA KAAAA KAAAA KA**Z**-NH_2		0.402 ± 0.013

[a]The amino acids are represented using the standard one-letter code; Z is pentafluorophenylalanine. [b]Values reported by Baldwin and co-workers (*12*).

Survey of Natural Protein Structures. The survey was performed on PDBselect (April 2009, 25% threshold) (*100*, *101*), a database of nonredundant protein chains. The α-helical conformation for each residue was defined by backbone dihedrals as described by Balaram and co-workers (*102*, *103*). The residues were selected using in-house code written in ActivePerl 5.8.8.819, and the dihedral angles were compiled using DSSP (*104*). Segments of six or more α-helical residues were considered to avoid end effects. The occurrence was compiled for Leu-Lys and Phe-Lys (i, $i + 3$) residue patterns, and for the specific sequences Leu-Ala-Ala-Lys and Phe-Ala-Ala-Lys. These occurrences were compiled using in-house code written in ActivePerl 5.8.8.819. The propensity of each residue pattern was calculated by dividing the occurrence of the sequence pattern in α-helices by the expected occurrence for the sequence pattern based on the structural context for each residue in the database. The expected occurrence and the corresponding standard deviation were obtained by bootstrapping (*105*) the sequence pattern against the appropriate contexts across PDBselect. The bootstrapping was performed using in-house code written in C++. Dividing the difference between the occurrence and the expected occurrence by the standard deviation gave the Z value, which was used to obtain the P value based on a normal distribution (*106*, *107*). The structures were examined and overlaid to generate the superimposed figures using Discovery Studio version 2.1 (Accelrys).

RESULTS

Peptide Design and Synthesis. A series of Ala-based peptides was designed with one Gly, Leu, Phe, or Pff incorporated at position 6, 11, or 16 or at the terminus (Table 1), analogous to those studied by Baldwin and co-workers (*12*, *35*). The three positions (6, 11, and 16) have the same immediate neighboring amino acids (±2 residues) but differ in the overall position in the helical peptide. Position 6 is close to the N-terminus; position 16 is close to the C-terminus, and position 11 is close to the middle of the peptide. In addition to these Ala-based peptides with a single Gly, Leu, Phe, or Pff substitution, Ala-based peptides with varying numbers of Leu residues incorporated at positions 6, 11, and 16 were also investigated. These uncharged non-hydrogen bonding residues were chosen because such side chains will have minimal interactions with the helix dipole and no hydrogen bonding with the backbone. We also synthesized the peptide with Ala at all three positions (6, 11, and 16) to serve as a control. Furthermore, 10 Ala-based peptides were investigated with Ala, Gly, Leu, Phe, or Pff at the N- or C-terminus to derive proper statistical mechanical capping parameters for Gly, Leu, Phe, and Pff based on modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*). The N-termini of all peptides were acetylated, and the C-termini were designed to be a carboxyamide so there would be no bias (created by charged termini) on the statistical mechanical capping parameters derived for these residues. Tyr was incorporated to facilitate the determination of the concentration by UV−vis (*98*, *108*), and the Gly-Gly intervening sequence was included to minimize interference in the circular dichroism signal by the Tyr chromophore (*109*). Multiple Lys residues were evenly distributed within the sequence to increase solubility, minimize aggregation, and balance the attractive and repulsive interaction with the helix macrodipole. All peptides were synthesized by solid phase peptide synthesis using Fmoc-based chemistry (*97*). Upon cleavage with concomitant side chain deprotection, the peptides were purified by RP-HPLC to >98% purity. The concentration of the peptides was determined by the Edelhoch method (*98*, *108*). Analogous peptides have been shown to be monomeric in solution (*12*, *26*, *27*, *35*, *110*), and the CD spectrum of each peptide did not change significantly between 80 and 160 μM.

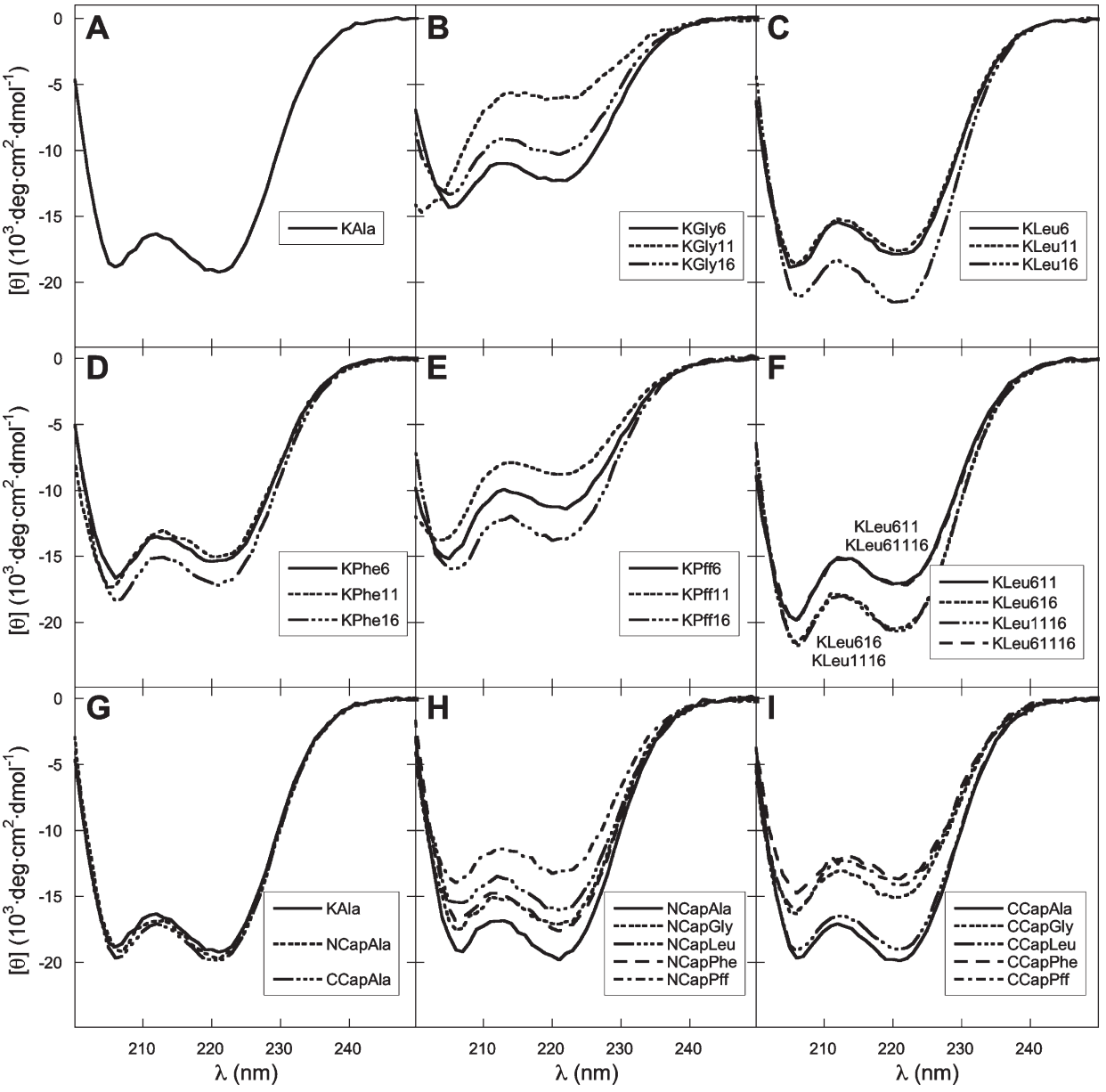


FIGURE 1: Circular dichroism spectra of the peptides at pH 7 (273 K) in 1 mM phosphate, borate, and citrate buffer with 1 M NaCl: (A) KAla; (B) KGly6, KGly11, and KGly16; (C) KLeu6, KLeu11, and KLeu16; (D) KPhe6, KPhe11, and KPhe16; (E) KPff6, KPff11, and KPff16; (F) KLeu611, KLeu616, KLeu1116, and KLeu61116; (G) KAla, NCapAla, and CCapAla; (H) NCapAla, NCapGly, NCapLeu, NCapPhe, and NCapPff; and (I) CCapAla, CCapGly, CCapLeu, CCapPhe, and CCapPff.

Importantly, the KXaa6, KXaa11, and KXaa16 peptides were monomeric in solution based on sedimentation equilibrium experiments, suggesting that intermolecular interactions do not contribute significantly to the helical content of these peptides in solution.

Circular Dichroism Spectroscopy and Fraction Helix (f_{helix}). The circular dichroism (CD) spectrum for each peptide was recorded at pH 7 in the presence of 1 M NaCl to derive the fraction helix (f_{helix}) (Table 1 and Figure 2); these were the conditions for the experiments performed by Baldwin and co-workers (*12*). The mean residue ellipticity at 222 nm reflects the helical content of a given peptide. More specifically, the more negative the CD signal at 222 nm, the more helical the peptide. For peptides with single substitutions at a given position (with Gly, Leu, Phe, or Pff), the helical content followed the trend Leu > Phe > Pff > Gly. This is consistent with previously published helix propensity for these four amino acids (*10*, *12*−*14*, *19*−*22*, *24*−*26*). However, incorporating Leu at different positions (6, 11, and 16) resulted in peptides with different helical content following the trend KLeu11 < KLeu6 < KLeu16 (Figure 1C). Similar positional trends were observed for Phe and Pff (Figure 1D,E). Interestingly, peptides with two or three Leu residues incorporated were distinctly divided into two groups (Figure 1F): peptides KLeu611 and KLeu61116 exhibited lower helical content than peptides KLeu616 and KLeu1116.

The fraction helix (f_{helix}) for each peptide was calculated from the corresponding mean residue ellipticity at 222 nm (Table 1). Our f_{helix} values appeared to be consistently lower than those reported by Baldwin and co-workers (*12*). To ensure the accuracy of our values, we determined the peptide concentration and measured the f_{helix} of peptide KAla multiple times from separate batches of this peptide; other peptides were also measured at least three times independently. On the basis of our repeated experimentation, concentration determination was critical for determining the mean

Table 2: Statistical Mechanical Helix Formation Parameters for Host Peptide Residues Derived from Experimentally Measured Fraction Helix Values Based on Modified Lifson−Roig Theory[a]

residue	*w*	*n*	*c*
Ala	1.44 ± 0.01	1.00[b]	0.801 ± 0.343
Gly	0.00119 ± 0.00948	2.85 ± 0.40	0.88[b]
Lys	1.06 ± 0.03	0.79[b]	2.85 ± 1.45
acetyl		9.52 ± 1.47	
carboxyamide			1.30[b]

[a]These statistical mechanical helix formation parameters are derived from experimental data from eight peptides: KAla, KGly6, KGly11, KGly16, NCapAla, CCapAla, NCapGly, and CCapGly; for sequences, see Table 1. [b]The parameter initially converged to a negative probability (which carries no physical meaning); therefore, the value was set to that published by Baldwin (*35*).

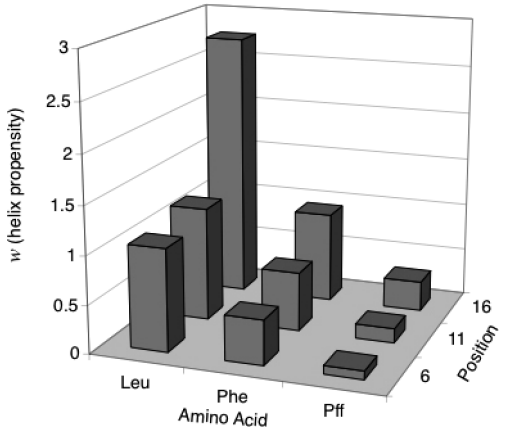


FIGURE 2: Helix propensity for Leu, Phe, and Pff at various positions (6, 11, and 16) derived from circular dichroism data at 222 nm of the corresponding Ala-based peptides based on the modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*).

residue ellipticity, and thus f_{helix}, of a peptide. Interestingly, the molar absorptivity (ε, also known as the extinction coefficient) is known to differ by 15% in different aqueous solutions (*98*). To ensure consistency, we have performed concentration determination of all peptide stock solutions in 6 M guanidinium chloride (*98*) and used aliquots of the stock to prepare the corresponding CD samples. Importantly, our experimentally determined mean residue ellipticity at 222 nm and f_{helix} for KAla are within error of the values reported by Waters (*25*), despite being different from the values published previously by Baldwin (*12*).

Helix Propensity of Leu, Phe, and Pff. The helix propensity (*w*), N-capping parameter (*n*), and C-capping parameter (*c*) for Ala, Gly, and Lys were derived from globally fitting the f_{helix} of peptides KAla, KGly6, KGly11, KGly16, NCapAla, CCapAla, NCapGly, and CCapGly based on modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*) using the least-squares method (Table 2). For parameters that converged to negative probabilities (which carry no physical meaning), the values were set to those published by Baldwin (*35*). The apparent helix propensities for Leu, Phe, and Pff at the various positions were initially derived from the f_{helix} values of the corresponding peptides based on modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*) (Figure 2). These initial derivations were performed using one peptide for each amino acid at one given position. For example, the apparent helix propensities for Leu at positions 6, 11, and 16 were derived from peptides KLeu6, KLeu11, and KLeu16, respectively. Interestingly, the helix propensity of Leu (w_{Leu}) appeared to be the same at positions 6 and 11 but was considerably higher at position 16. Similarly, the helix propensity of Phe (and Pff) appeared to be the same at positions 6 and 11 (for each amino acid) but was significantly higher at position 16. Therefore, we derived a global $w_{6,11}$ for positions 6 and 11 and a separate w_{16} for position 16 for the three amino acids Leu, Phe, and Pff (Table 3). The resulting helix propensity $w_{6,11}$ for the amino acids followed the trend Leu > Phe > Pff (consistent with known helix propensity trends Leu > Phe and Phe > Pff) (*10*, *12*−*14*, *19*−*22*, *24*−*26*). The same trend was observed for the values of w_{16}.

Table 3: Apparent Helix Propensities (*w*), Free Energies of Helix Formation ($\Delta G_{helix\ formation}$), and Putative Xaa16−Lys (*i*, *i* + 3) Interaction Energetics for Leu, Phe, and Pff

amino acid	*w*	$\Delta G_{helix\ formation}$ (cal/mol)[a]	putative Xaa16−Lys (*i*, *i* + 3) interaction energy (cal/mol)[b]
Leu (6,11)	1.09 ± 0.04	−46.8 ± 20.3	
Leu (16)	2.41 ± 0.23	−477 ± 54	−804 ± 48
Phe (6, 11)	0.573 ± 0.037	302 ± 36	
Phe (16)	0.940 ± 0.047	33.6 ± 27.8	−521 ± 43
Pff (6, 11)	0.135 ± 0.013	1090 ± 50	
Pff (16)	0.310 ± 0.052	636 ± 99	−790 ± 130

[a]$\Delta G_{helix\ formation} = -RT \ln(w)$. [b]The putative Xaa16−Lys (*i*, *i* + 3) interaction energetics are derived from the corresponding experimentally measured f_{helix} values (Table 1) by the nesting block method (*41*, *66*, *99*), using the helix propensity at positions 6 and 11 as the helix propensity at position 16.

The calculated fraction helix values using the parameters determined in this study fit all the experimentally measured data to a similar degree compared to Baldwin parameters (*12*, *35*, *90*) based on the sum of the squares of the difference between calculated and experimental values (Table 4). The lower the sum of squares, the better the calculated values matched the experimental values. The experimentally measured f_{helix} values of the peptides studied by Baldwin (*12*, *35*, *90*) were more accurately calculated using Baldwin parameters compared to our parameters. However, the f_{helix} values of peptides investigated in this study were more accurately calculated using our parameters compared to Baldwin parameters (*12*, *35*, *90*) (Table 4).

The apparent high helix propensity for all three amino acids when they are placed near the C-terminus could be due to specific stabilizing Xaa−Lys (*i*, *i* + 3) interactions with the terminal Lys, because intrahelical (*i*, *i* + 3) interactions can be energetically favorable (*21*, *22*, *25*, *40*−*66*). Also, intrahelical (*i*, *i* + 4) interactions have been observed between aromatic and basic residues in analogous Ala-based peptides (*50*, *51*, *53*, *60*, *65*). Accordingly, Xaa−Lys (*i*, *i* + 3) interaction could be possible in our studies. Using the helix propensity at positions 6 and 11 as the helix propensity at position 16, we calculated the putative Xaa−Lys (*i*, *i* + 3) interaction energetics near the C-terminus (Table 3). The putative stabilizing interaction appears to follow the trend Pff ∼ Leu > Phe. Hydrophobic interactions between the four hydrophobic methylenes on the Lys side chain and these residues would be one possible explanation for the putative interaction trend. The hydrophobicity of Pff is higher than that of Phe, whereas the aliphatic Leu side chain is slightly more hydrophobic than Phe, with more flexibility to interact with (and pack against) the Lys methylenes. However, it was unclear if such Phe-Lys (*i*, *i* + 3) and Leu-Lys (*i*, *i* + 3) sequence patterns are prevalent in natural protein structures.

Table 4: Sum of Squares of the Differences between Experimentally Measured and Calculated f_{helix} Values for Leu-, Phe-, and Ala-Containing Ala-Based Peptides Based on Modified Lifson−Roig Theory Using Various Parameter Sets

peptide	Baldwin 1994 parameters, calculated sum of squares[e]	Baldwin 1995 parameters, calculated sum of squares[e]	Baldwin 1996 parameters, calculated sum of squares[f]	parameters from this study, calculated sum of squares[e]
Baldwin 1994 (10 peptides[a])	0.0290	0.0156	0.0148	0.0686
Baldwin 1995 (8 peptides[b])	0.0261	0.0458	0.0063	0.1035
Baldwin 1996 (11 peptides[c])	0.0293	0.0156	0.0154	0.0692
This study (22 peptides[d])	0.1017	0.1113	0.1897	0.0274
all (34 peptides[a,b,c,d])	0.1354	0.1596	0.2051	0.1348

[a]Ten Ala-based peptides containing Ala, Lys, Leu, and Phe studied by Baldwin (*12*): YG, YGG, YGGG, YGG-G8, YGG-G13 YGG-G18, YGG-1L, YGG-1F, YGG-3L, and 2L-GGY. [b]Eight Ala-based peptides containing Ala, Lys, and Leu studied by Baldwin (*35*): YG, YGG, YGGG, YGG-G8, YGG-G13, YGG-G18, XAK, and CCZ. [c]Eleven Ala-based peptides containing Ala, Lys, Leu, and Phe studied by Baldwin (*90*): YGAK, YGGAK, YGGGAK, YGG-G7, YGG-G12, YGG-G17, YG-ZC17, YGG-1L, YGG-3L, 2L-GGY, and YGG-1F. [d]Twenty-two Ala-based peptides containing Leu, Phe and Ala investigated in this study: KAla, NCapAla, CCapAla, KGly6, KGly11, KGly16, NCapGly, CCapGly, KPhe6, KPhe11, KPhe16, NCapPhe, CCapPhe, KLeu6, KLeu11, KLeu16, KLeu611, KLeu616, KLeu1116, KLeu61116, NCapLeu, and CCapLeu (Table 1). [e]The experimental f_{helix} was derived from the CD data at 222 nm as determined by Baldwin (*12*). [f]The experimental f_{helix} was derived from the CD data at 222 nm as determined by Baldwin (*90*).

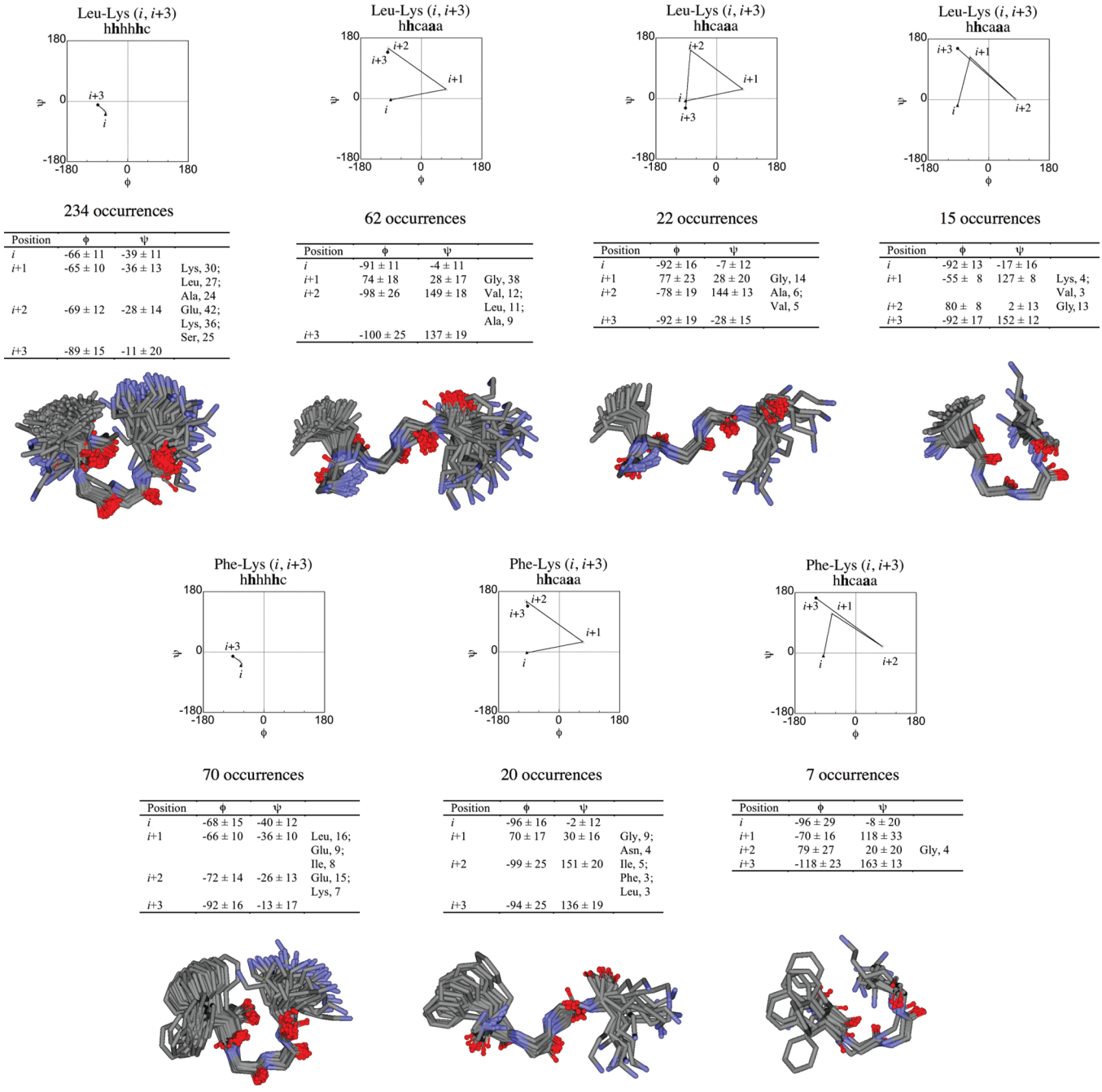
Exploring Natural Protein Structures. The nonredundant protein structure database PDBselect (April 2009, 25% threshold) (*100*, *101*) was surveyed for Leu-Lys (i, $i + 3$) and Phe-Lys (i, $i + 3$) sequence patterns in the context of various helix-related structures (Table 5). The definition of the helical conformation was based on backbone dihedrals (*102*, *103*), and only helices with more than six residues were considered to be a helix to avoid "helices" with fewer than one turn (*66*). A total of 4418 protein chains and 666086 residues were considered involving 17622 helices and 236790 helical residues. Approximately 33.7% of the residues in the database were helical (236790 of 666086), a result similar to those of earlier analyses (*1*−*3*). To gauge the significance of the occurrences for Leu-Lys (i, $i + 3$) and Phe-Lys (i, $i + 3$), the pair propensities for the sequence patterns in the various structures were derived by dividing the occurrence by the corresponding expected occurrence. The expected occurrence was obtained by bootstrapping the residues in the individual structural context of each amino acid. Bootstrapping was performed 100000 times for these cases to yield standard deviations for the expected occurrence. This enabled the calculation of standard deviations for the pair propensities and Z values (Table 5), which was used to derive the P values. Propensities greater than unity represent occurrences that are higher than expected on the basis of residue occurrences in the corresponding structural context, whereas propensities less than unity indicate occurrences lower than the expected value. To survey Leu-Lys (i, $i + 3$) and Phe-Lys (i, $i + 3$) sequence patterns in various helix-related structures, we considered the conformation of six-residue segments: $i - 1$ through $i + 4$. The conformation of each residue in the segment was categorized as helix (h), nonhelix (or coil, c), or any structure (including helix and nonhelix, a). The conformation of the six-residue segment is designated with the conformation of residues i (Leu or Phe) and $i + 3$ (Lys) boldfaced and underlined. For example, the conformation of the sequence pattern Leu-Lys (i, $i + 3$) in all protein structures would be designated a**a**aaa**a**a; the conformation of the same sequence pattern within only helical structures (including termini) would be designated a**h**hh**h**a.

The pair propensity for both Leu-Lys (i, $i + 3$) and Phe-Lys (i, $i + 3$) sequences was close to unity when considering all protein structures in the survey [a**a**aaa**a**a (Table 5)], suggesting indifference for both sequence patterns in protein structures. In the context of natural helices (a**h**hh**h**a), the two patterns exhibited pair propensities less than unity, suggesting a less than expected occurrence for both patterns in helices. Because the exceptionally high statistical mechanical helix propensity (w) was observed near the C-terminus of the helical peptides for both Leu and Phe (vide supra), we then focused on surveying the C-terminal ends of natural helices. Five different helix terminating scenarios were considered: h**h**hh**h**c, h**h**hh**c**a, h**h**hc**a**a, h**c**a**a**a, and h**c**aa**a**a (Table 5). Interestingly, only helices terminating immediately after Lys_{i+3} (h**h**hh**h**c) or immediately after Leu_i/Phe_i (h**c**a**a**a) exhibited pair propensity values significantly greater than unity, suggesting

Table 5: Statistical Analysis for Leu-Lys and Phe-Lys (i, $i + 3$) Sequence Patterns in Various Structural Patterns

i	$i+3$	structure[a]	occurrence[b]	pair propensity[c]	Z value[d]	P value[e]
Leu	Lys	a**a**aaa**a**a	3476	0.949 ± 0.016	−3.06	2.21×10^{-3}
Leu	Lys	a**h**hh**h**a	1285	0.848 ± 0.022	−6.14	8.16×10^{-10}
Leu	Lys	h**h**hh**h**c	234	**1.68 ± 0.14**	8.01	1.15×10^{-15}
Leu	Lys	h**h**hh**c**a	117	1.12 ± 0.11	1.20	2.30×10^{-1}
Leu	Lys	h**h**hc**a**a	131	1.23 ± 0.12	2.37	1.78×10^{-2}
Leu	Lys	h**h**ca**a**a	188	**2.05 ± 0.21**	10.1	6.77×10^{-24}
Leu	Lys	h**c**aa**a**a	69	1.19 ± 0.16	1.42	1.56×10^{-1}
Phe	Lys	a**a**aa**a**a	1737	1.02 ± 0.02	1.02	3.08×10^{-1}
Phe	Lys	a**h**hh**h**a	508	0.908 ± 0.043	−1.46	1.44×10^{-1}
Phe	Lys	h**h**hh**h**c	70	**1.43 ± 0.20**	3.01	2.61×10^{-3}
Phe	Lys	h**h**hh**c**a	31	0.837 ± 0.136	−1.00	3.17×10^{-1}
Phe	Lys	h**h**hc**a**a	32	0.856 ± 0.139	−0.890	3.73×10^{-1}
Phe	Lys	h**h**ca**a**a	70	**1.63 ± 0.25**	4.15	3.32×10^{-5}
Phe	Lys	h**c**aa**a**a	49	1.19 ± 0.19	1.23	2.19×10^{-1}

[a]The various structural patterns for a six-residue segment from residues $i - 1$ through $i + 4$; residues i and $i + 3$ are underlined. The structure for each amino acid is determined by the backbone dihedrals (ϕ and ψ) to be either helix (h) or nonhelix (c). When the structure of the residue is not restricted, the letter a is used to represent any structure. The ϕ and ψ definition for helix follows a relative inclusive criterion (*102*, *103*). [b]The occurrence of the sequence and structure pattern in the nonredundant protein structure database PDBselect (April 2009, 25% threshold) (*100*, *101*). [c]The occurrence of the Leu-Lys or Phe-Lys (i, $i + 3$) pairs divided by the corresponding expected value. The expected value was obtained by bootstrapping the corresponding individual structural pattern for positions i and $i + 3$, thereby removing bias due to amino acid usage in the structure of interest. Because complete bootstrapping was performed 100000 times, this enabled the calculation of a standard deviation for the expected value and thus propensity. [d]The difference between the occurrence and the expected value divided by the standard deviation for the expected value. In other words, the number of standard deviations that separate the occurrence and the expected value; a positive number means the occurrence is larger than the expected value, whereas a negative number means the opposite. [e]The probability that the occurrence and expected occurrence are the same based on the standard deviation obtained from bootstrapping assuming a Gaussian distribution.



Position	ϕ	ψ	
i	-66 ± 11	-39 ± 11	
i+1	-65 ± 10	-36 ± 13	Lys, 30; Leu, 27; Ala, 24
i+2	-69 ± 12	-28 ± 14	Glu, 42; Lys, 36; Ser, 25
i+3	-89 ± 15	-11 ± 20	

Position	ϕ	ψ	
i	-91 ± 11	-4 ± 11	
i+1	74 ± 18	28 ± 17	Gly, 38
i+2	-98 ± 26	149 ± 18	Val, 12; Leu, 11; Ala, 9
i+3	-100 ± 25	137 ± 19	

Position	ϕ	ψ	
i	-92 ± 16	-7 ± 12	
i+1	77 ± 23	28 ± 20	Gly, 14
i+2	-78 ± 19	144 ± 13	Ala, 6; Val, 5
i+3	-92 ± 19	-28 ± 15	

Position	ϕ	ψ	
i	-92 ± 13	-17 ± 16	
i+1	-55 ± 8	127 ± 8	Lys, 4; Val, 3
i+2	80 ± 8	2 ± 13	Gly, 13
i+3	-92 ± 17	152 ± 12	

Position	ϕ	ψ	
i	-68 ± 15	-40 ± 12	
i+1	-66 ± 10	-36 ± 10	Leu, 16; Glu, 9; Ile, 8
i+2	-72 ± 14	-26 ± 13	Glu, 15; Lys, 7
i+3	-92 ± 16	-13 ± 17	

Position	ϕ	ψ	
i	-96 ± 16	-2 ± 12	
i+1	70 ± 17	30 ± 16	Gly, 9; Asn, 4
i+2	-99 ± 25	151 ± 20	Ile, 5; Phe, 3; Leu, 3
i+3	-94 ± 25	136 ± 19	

Position	ϕ	ψ	
i	-96 ± 29	-8 ± 20	
i+1	-70 ± 16	118 ± 33	
i+2	79 ± 27	20 ± 20	Gly, 4
i+3	-118 ± 23	163 ± 13	

FIGURE 3: Two-helix C-terminal structural scenarios (with occurrences higher than expected) for Leu-Lys (*i*, *i* + 3) (top row) and Phe-Lys (*i*, *i* + 3) (bottom row) sequence patterns. There is only one type of backbone conformation for the h**h**hh**h**c structures, whereas there are several different types of backbone conformations for the h**h**ca**a**a structures. The backbone $\phi-\psi$ plots, the number of occurrences, average backbone dihedral angles, the most frequently observed amino acids for each position, and the overlay of the three-dimensional structures for each type are depicted.

Phe-Lys (*i*, *i* + 3) and Leu-Lys (*i*, *i* + 3) sequences occurred more than expected in these two structural contexts. Furthermore, these four sequence structural patterns were particularly significant on the basis of the considerably highly positive *Z* values and small *P* values. Accordingly, these sequence structural patterns were examined in more detail.

There were 234 occurrences of the Leu-Lys (*i*, *i* + 3) sequence patterns with Lys as the most C-terminal residue with helical dihedrals (h**h**hh**h**c) (Table 5 and Figure 3); however, no Leu-Lys interactions were apparent in these cases (Figure 3). Nonetheless, there were three cases of Leu-Ala-Ala-Lys sequence, which is the exact sequence in our experimental studies (Figure 4). There were 70 occurrences of the Phe-Lys (*i*, *i* + 3) sequence patterns with Lys as the most C-terminal residue with helical dihedrals (h**h**hh**h**c); however, none of the sequences were the Phe-Ala-Ala-Lys sequence. Furthermore, no Phe−Lys interactions were apparent (Figure 3). Interestingly, the relative placement of the Leu/Phe four residues upstream from Gly in h**h**hh**h**c is consistent with previously observed Schellman motifs (*102*, *111*).

There were 188 occurrences of the Leu-Lys (*i*, *i* + 3) sequence patterns with Leu as the most C-terminal residue with helical dihedrals (h**h**ca**a**a) (Table 5 and Figure 3); however, none of the sequences were Leu-Ala-Ala-Lys. The most prevalent structure (84 of 188) involved the Schellman motif with a residue in the left-handed helix conformation (Figure 3). Unfortunately, this is unlikely the case for the peptides in this study with two

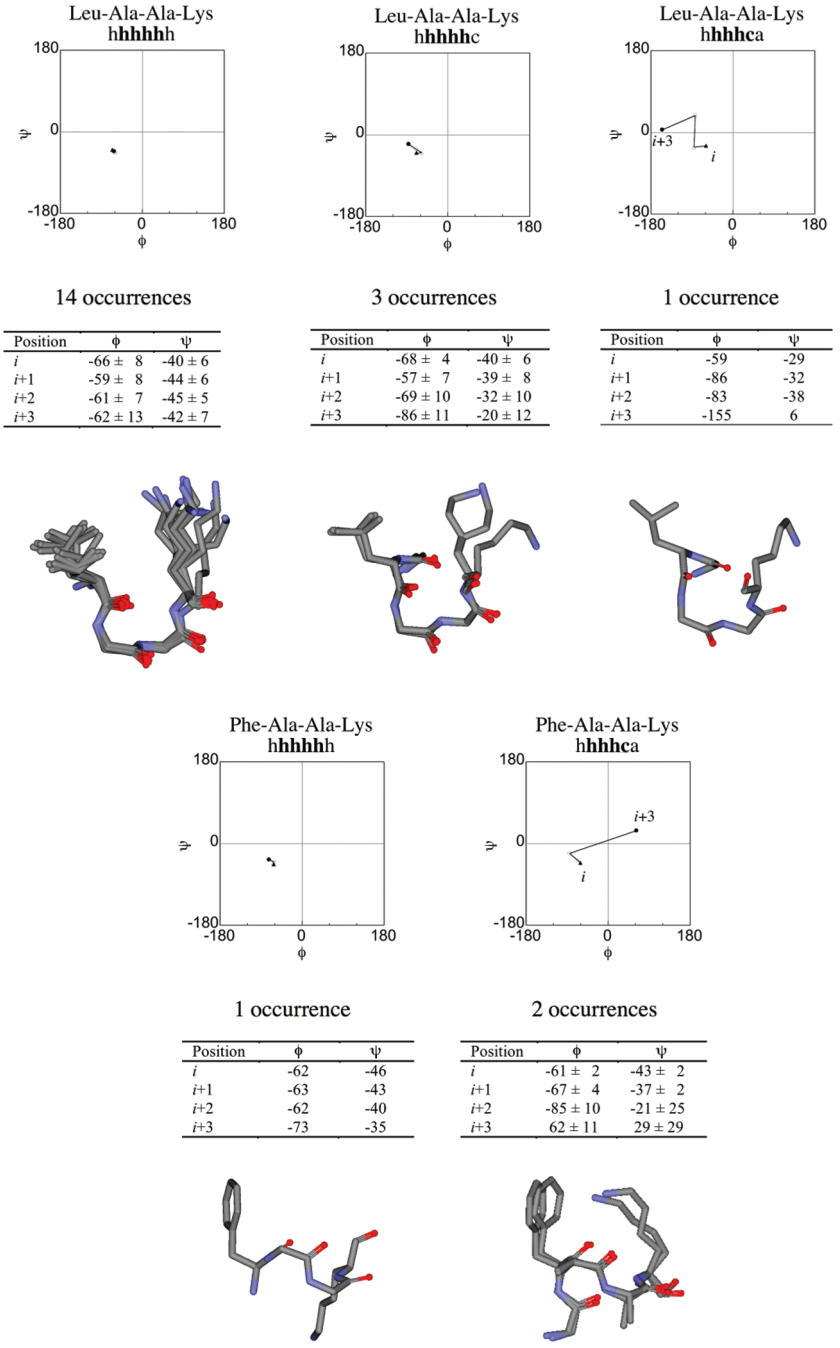


FIGURE 4: Helix-related structural scenarios for Leu-Ala-Ala-Lys (top row) and Phe-Ala-Ala-Lys (bottom row) sequence patterns. There are three structural scenarios for the Leu-Ala-Ala-Lys sequence, whereas there are two structural scenarios for the Phe-Ala-Ala-Lys sequence. The backbone $\phi-\psi$ plots, the number of occurrences, average backbone dihedral angles, and the overlay of the three-dimensional structures for each type are depicted.

intervening Ala residues between Leu and Lys. There were 70 occurrences of the Phe-Lys (i, $i+3$) sequence patterns with Phe as the most C-terminal residue with helical dihedrals (h**h**ca**a**a) (Table 5 and Figure 3). The most prevalent structures (27 of 70) involved a residue in the left-handed helix conformation, and none of the sequences were Phe-Ala-Ala-Lys. In these structures, a hydrophobic residue was frequently present immediately after the left-handed Gly, consistent with the Schellman motif at the C-terminus of natural protein helices (*102*, *111*). Because a left-handed helix conformation was involved for the sequence patterns with Leu_i/Phe_i as the most C-terminal residue with helical dihedrals (h**h**ca**a**a), these sequence structural patterns are not likely to be relevant to this study (involving right-handed intervening Ala residues). In comparison, the Leu-Lys (i, $i+3$) and Phe-Lys (i, $i+3$) sequence patterns with Lys as the most C-terminal residue with helical dihedrals (h**h**hh**h**c) would be more relevant to the experimentally measured high statistical mechanical helix propensity of Leu and Phe. However, no specific side chain interactions between Leu/Phe and Lys were observed in these structures to explain the high helix forming parameters measured experimentally.

These initial survey results considered only the residues at the i and $i+3$ positions, ignoring the intervening sequences. It would be logical to ask how frequently the specific Leu-Ala-Ala-Lys and Phe-Ala-Ala-Lys sequences occur in natural structures. Accordingly, we performed surveys on these two specific sequences

Table 6: Statistical Analysis for Leu-Ala-Ala-Lys and Phe-Ala-Ala-Lys Sequence Patterns in Various Structural Patterns

sequence from i to $i+3$	structure[a]	occurrence[b]	propensity[c]	Z value[d]	P value[e]
Leu-Ala-Ala-Lys	a**aaaa**a	25	1.22 ± 0.27	1.00	3.17×10^{-1}
Leu-Ala-Ala-Lys	h**hhhh**h	14	0.896 ± 0.226	−0.41	6.82×10^{-1}
Leu-Ala-Ala-Lys	h**hhhh**c	3	1.45 ± 1.00	0.65	5.16×10^{-1}
Leu-Ala-Ala-Lys	h**hhhc**a	1	0.423 ± 0.274	−0.89	3.73×10^{-1}
Phe-Ala-Ala-Lys	a**aaaa**a	5	0.530 ± 0.171	−1.45	1.47×10^{-1}
Phe-Ala-Ala-Lys	h**hhhh**h	1	0.180 ± 0.074	−1.97	4.88×10^{-2}
Phe-Ala-Ala-Lys	h**hhhc**a	2	2.39 ± 2.58	1.28	2.01×10^{-1}

[a]The various structural patterns for a six-residue segment from residue $i-1$ to $i+4$; residues i and $i+3$ are underlined. The structure for each amino acid is determined by the backbone dihedrals (ϕ and ψ) to be either helix (h) or nonhelix (c). When the structure of the residue is not restricted, the letter a is used to represent any structure. The ϕ and ψ definition for helix follows a relative inclusive criterion (*102*, *103*). [b]The occurrence of the sequence and structure pattern in the nonredundant protein structure database PDBselect (April 2009, 25% threshold) (*100*, *101*). [c]The occurrence for the Leu-Ala-Ala-Lys or Phe-Ala-Ala-Lys sequences divided by the corresponding expected value. The expected value was obtained by bootstrapping the corresponding individual structural pattern for positions i through $i+3$, thereby removing bias due to amino acid usage in the structure of interest. Because complete bootstrapping was performed 10000 times, this enabled the calculation of a standard deviation for the expected value and thus propensity. [d]The difference between the occurrence and the expected value divided by the standard deviation for the expected value. In other words, the number of standard deviations that separate the occurrence and the expexted value; a positive number means the occurrence is larger than the expected value, whereas a negative number means the opposite. [e]The probability that the occurrence and expected occurrence are the same based on the standard deviation obtained from bootstrapping assuming a Gaussian distribution.

(Table 6 and Figure 4). The occurrences were low for both sequences, and extremely low for Phe-Ala-Ala-Lys. Importantly, both sequences seemed to be present in helical structures; 18 of the 25 sequence patterns were involved in helices for Leu-Ala-Ala-Lys (Figure 4). More specifically, Leu-Ala-Ala-Lys occurred mostly within a helix (14 of 25), and somewhat at the C-terminal end of a helix (4 of 25), but not at the N-terminal end of a helix. Two of the five Phe-Ala-Ala-Lys occurrences were at the C-terminal end of a helix with clear Lys−Phe cation−π interactions (*112*) (Figure 4). The Leu/Phe-Ala-Ala-Lys sequence patterns do not occur very frequently in natural helices and are not found at the N-terminus of helices. Despite the low overall occurrences and propensities, both sequences are found at the C-termini of natural helices.

DISCUSSION

Our CD data show that incorporating the same amino acid at different positions results in varying helical contents (Figure 1). Interestingly, there is a general positional trend for the helical content of the peptides: KXaa16 > KXaa6 > KXaa11 (except for Gly). Glycine is peculiar in being capable of adopting various left-handed structures with $\phi > 0$ due to the achiral nature of the residue. Because these conformations are not typically accessible to natural chiral amino acids, there will be no further discussion regarding Gly. Importantly, lower f_{helix} values should be observed upon substitution of Ala with Leu, Phe, or Pff in the middle of the peptide (position 11) compared to substitution at the ends of the peptide (positions 6 and 16) for two reasons. First, all three amino acids have lower helix propensities than the host residue Ala, leading to lower helicities. Second, substituting Ala with a more helix breaking amino acid (than Ala) in the center of the peptide attenuates the helical content more than placing such amino acids near the termini because the termini are inherently weakly helical due to end fraying (*38*, *113*, *114*). The difference between substituting the Ala at positions 16 and 6 is most likely due to the asymmetric geometry of the α-helix. In particular, the side chain Cα−Cβ vectors inherently project toward the N-terminus because of the chirality of the amino acids, backbone dihedrals, and helix handedness. This geometry results in frequent side chain shielding of upstream helix hydrogen bonds, thereby stabilizing the helical conformation. As such, the N-terminus is more stable than the C-terminus in α-helices (*114*−*119*). Although the modified Lifson−Roig theory apparently accounts for this asymmetry by introducing capping parameters n and c (*12*, *26*, *34*, *38*, *81*), the exceptionally high helical content for peptides with non-Ala residues at position 16 cannot be fully rationalized with the capping parameters alone. Therefore, this required us to invoke the exceptionally high helix propensity for the residue at position 16, and the apparent helix propensity for the same amino acid follows the trend $w_{16} > w_6 = w_{11}$. These results are in sharp contrast to those published by Stellwagen and co-workers (*92*), apparently showing positional independence and additivity of amino acid replacements on helix stability at analogous positions. However, Stellwagen and co-workers artificially normalized the CD signals for all peptides, attempting direct comparison between peptides without invoking any statistical mechanical theory. Furthermore, the monomeric helical peptides were assumed to be in a two-state equilibrium, leading to most likely invalid van't Hoff analyses.

Our results show that helix propensity for an amino acid (Leu, Phe, or Pff) is similar at positions 6 and 11, but different at position 16 in a 19-residue Ala-based peptide. In other words, helix propensity for an amino acid is similar near the N-terminus and center of the peptide chain, but different near the C-terminus of the peptide chain. These results appear to somewhat contradict the results from Serrano and co-workers in a different Ala-based system (*94*−*96*). Serrano and co-workers found that helix propensity varies when placed at the first three N-terminal residues of the peptide chain, at the last three C-terminal residues of the peptide chain, and in the middle of the helical peptide (or seven residues from the ends of the peptide chain) (*94*−*96*). However, the fundamental difference between the two systems is the variance in the immediate surrounding residues (±2 amino acids). These neighboring residues change according to the guest position in Serrano's system (*94*−*96*) but remain constant in our Baldwin-based system (*12*). Therefore, the results from the two studies are valid in each of the corresponding contexts and do not contradict one another.

The less than expected occurrence for both Leu-Lys and Phe-Lys (i, $i+3$) sequence patterns in natural protein helices [a**h**hh**h**a (Table 5)] is consistent with the amphiphilic nature of helices in globular proteins (*103*, *120*−*124*). Helices residing on the surface of globular proteins would require a hydrophilic face and a hydrophobic face opposite from one another. Leu and Phe are considered to be hydrophobic, whereas Lys is considered to be hydrophilic. Because the geometry of an ideal helix has 3.6 residues per turn, amino acids three residues apart should have similar characteristics (i.e., both hydrophobic or both hydrophilic). Therefore, it is not surprising that both Leu-Lys and Phe-Lys (i, $i+3$) sequences occur less than expected in natural protein helices. In contrast, both Leu-Lys and Phe-Lys (i, $i+3$) sequences occur more than expected near the C-terminus of natural protein helices involving the structures h**h**hh**h**c and h**h**ca**a**a (Table 5).

However, the h**h**hh**h**c structures involve one residue beyond Lys, and Lys is the last residue in our peptides. Therefore, the relevancy of the h**h**hh**h**c structures may be somewhat debatable but cannot be completely ruled out. Unfortunately, the h**h**ca**a**a structures involve a left-handed conformation ($\phi > 0$) for position $i + 1$ or $i + 2$ (Figure 3), which is extremely rare for the Ala residue in our peptides. Therefore, the h**h**ca**a**a structures are most likely irrelevant to our peptides. In contrast, the specific sequences in our study, Leu-Ala-Ala-Lys and Phe-Ala-Ala-Lys, do not occur frequently in nature (Table 6), suggesting that Ala-based peptides most likely do not represent the situation in natural protein helices. Nonetheless, Ala-based peptides are still good minimalist model systems for fundamental studies.

There are several possible explanations for the apparent positional dependence for the helix propensities in our studies. First, Lys is preferred at the C-cap and C1 positions in protein helices, whereas Leu is preferred at the C3 and C4 positions (*111*, *125*−*127*). These general preferences are consistent with potential intrahelical Xaa−Lys (i, $i + 3$) interactions in our peptide, and intrahelical (i, $i + 3$) interactions are well-documented (*21*, *22*, *25*, *40*−*66*). Second, the C-terminal end of a helix may adopt a 3_{10}-helix, because 3_{10}-helices are observed at the C-terminal end of protein α-helices (*2*, *128*). Furthermore, Ala-based peptides have been shown to adopt a mixture of α- and 3_{10}-helices (*129*−*131*), with the C-terminal end adopting a 3_{10}-helix for a 21-residue peptide (*130*). Interestingly, the three-residue per turn geometry of the 3_{10}-helix could be further stabilized by an intrahelical (i, $i + 3$) interaction as proposed by Doig (*132*). Analogous intrahelical Xaa−Lys (i, $i + 3$) interactions have been implied by Doig (*60*) and invoked by Serrano (*22*) for α-helices. Third, residues near the terminus of the peptide chain could exhibit different helix propensities. This proximity to the terminus of the peptide chain is different from being positioned at the ends of helices (or capping effects). On the basis of the statistical mechanical models that describe the multistate nature for the equilibrium of monomeric helical peptides, conformational states with helices terminating in the center of the peptide do exist, and capping effects can account for them (*12*, *26*, *34*, *38*, *81*). However, these capping effects alone cannot account for the high helical content for peptides with non-Ala residues incorporated near the C-terminus of the peptide chain in this study. Therefore, the high C-terminal helix propensity observed in this study may be due to the proximity of position 16 to the C-terminus of the peptide chain (the fourth residue from the C-terminus of the peptide chain). In contrast, position 6 near the N-terminus is farther from the N-terminus of the peptide chain, and thereby unaffected. This appears to be consistent with molecular dynamics simulations on polymers, revealing high main chain translational mobility near the termini based on mean-square-displacement analysis (*133*, *134*). Also, this high main chain translational mobility is no longer present by the fifth monomer unit from the terminus (*133*).

Helices are rarely present at the ends of the protein chain in natural proteins. Furthermore, the residues at the termini of natural proteins are generally structurally disordered and unresolved in atomic-resolution structures. Nonetheless, the upstream and downstream protein chain can create different structural contexts compared to the N- and C-termini of the peptide chain, respectively. In particular, the steric restrictions would be higher and the conformational entropy should be lower, in the presence of the upstream or downstream protein chain compared to their absence (or near the termini of the peptide chain). This would account for the molecular mechanics results showing a difference in solvent exposure, the number of van der Waals interactions, and configurational entropy for residues near the N-terminus compared to internal residues (*94*). Therefore, the helix propensity should be different on the basis of the proximity to the ends of the peptide chain as shown in previous studies (*94*−*96*) and the C-terminus (position 16) in this study. The lack of this chain terminal effect near the N-terminus (position 6) in this study suggests that the effect is diminished at the sixth residue from the end of the peptide chain, consistent with simulation results on a polymer (*133*). The length dependence of helix propensity in Ala-based peptides (*85*, *91*) may also be due to this chain terminal effect, because the proportion of residues in the proximity of the ends of the peptide chain changes with chain length. Because the chain terminal effect would not exist at internal positions of a protein, no position dependence in helix propensity was observed in an internal helix of a protein (*93*). Overall, the chain terminal effect appears to explain the high w_{16} compared to $w_{6,11}$ in this study, the positional dependence of helix propensity in previous studies of Ala-based peptides (*94*−*96*), the length dependence of helix propensity (*85*, *91*), and the position independence in a protein helix (*93*).

CONCLUSIONS

Circular dichroism spectroscopy coupled with calculations based on modified Lifson−Roig theory (*12*, *26*, *34*, *38*, *81*) has been used to determine the helix propensities for Leu, Phe, and Pff at different positions of the same helical peptide. Incorporating Leu, Phe, or Pff at various positions of the same peptide resulted in varying helical contents. The helicity of the substituted host peptide follows the general trend KXaa16 > KXaa6 > KXaa11. Position dependence of the helix propensities was observed for Leu, Phe, and Pff, with exceptionally high helix propensity at position 16 (close to the C-terminus) compared to positions 6 and 11. The unique behavior for these non-Ala residues at position 16 may be due to specific C-cap interactions, adoption of a 3_{10}-helix near the C-terminus perhaps with Xaa−Lys (i, $i + 3$) interactions, or the chain terminal effect. Nonetheless, it appears that helix propensity is generally position-independent except in the presence of alternative structures or in the proximity of either terminus of the peptide chain. These results should facilitate the design of helical peptides, proteins, and non-natural foldamers.

ACKNOWLEDGMENT

We thank the Computer and Information Networking Center at National Taiwan University for the support of the high-performance computing facilities. We thank National Science Foundation (REU, CHE0453206) for supporting Nathan S. Hutson, who participated in this project. We thank Professor Cheu-Pyeng Cheng (Department of Chemistry, National Tsing Hua University, Hsinchu, Taiwan, China) and Professor Alison McCurdy (Department of Chemistry and Biochemistry, California State University, Los Angeles, CA) for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Details of the synthesis and characterization of the peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Chou, P. Y., and Fasman, G. D. (1974) Conformational parameters for amino-acids in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry 13*, 211–222.

2. Barlow, D. J., and Thornton, J. M. (1988) Helix geometry in proteins. *J. Mol. Biol. 201*, 601–619.
3. We recently conducted survey on a nonredundant protein structure database PDBselect {December 2003, 25% threshold [(1994) *Protein Sci. 3*, 522−524]}. Using hydrogen bonding to define the α-helix conformation as described by Kabsch and Sander [(1983) *Biopolymers 22*, 2577−2637], we found 101210 α-helical residues of 341444 total residues; 29.6% of all residues were α-helical.
4. Gellman, S. H. (1998) Foldamers: A manifesto. *Acc. Chem. Res. 31*, 173–180.
5. Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S., and Moore, J. S. (2001) A field guide to foldamers. *Chem. Rev. 101*, 3893–4011.
6. Sanford, A. R., and Gong, B. (2003) Evolution of helical foldamers. *Curr. Org. Chem. 7*, 1649–1659.
7. Cheng, R. P. (2004) Beyond de novo protein design: De novo design of non-natural folded oligomers. *Curr. Opin. Struct. Biol. 14*, 512–520.
8. Hecht, S., and Huc, I., Eds. (2007) Foldamers. Structure Properties, and Applications, Wiley-VCH, Weinheim, Germany.
9. Platzer, K. E. B., Ananthanarayanan, V. S., Andreatta, R. H., and Scheraga, H. A. (1972) Helix-coil stability constants for the naturally occurring amino acids in water. IV. Alanine parameters from random poly(hydroxypropylglutamine-*co*-L-alanine). *Macromolecules 5*, 177–187.
10. Sueki, M., Lee, S., Powers, S. P., Denton, J. B., Konishi, Y., and Scheraga, H. A. (1984) Helix-coil stability constants for the naturally occurring amino acids in water. 22. Histidine parameters from random poly[(hydroxybutyl)glutamine-co-L-histidine]. *Macromolecules 17*, 148–155.
11. Marqusee, S., Robbins, V. H., and Baldwin, R. L. (1989) Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U.S.A. 86*, 5286–5290.
12. Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci. 3*, 843–852.
13. O'Neil, K. T., and DeGrado, W. F. (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*, 646–651.
14. Lyu, P. C., Liff, M. I., Marky, L. A., and Kallenbach, N. R. (1990) Side chain contributions to the stability of α-helical structure in peptides. *Science 250*, 669–673.
15. Spek, E. J., Olson, C. A., Shi, Z., and Kallenbach, N. R. (1999) Alanine is an intrinsic α-helix stabilizing amino acid. *J. Am. Chem. Soc. 121*, 5571–5572.
16. Kemp, D. S., Boyd, J. G., and Muendel, C. C. (1991) The helical *s* constant for alanine in water derived from template-nucleated helices. *Nature 352*, 451–454.
17. Kemp, D. S., Oslick, S. L., and Allen, T. J. (1996) The structure and energetics of helix formation by short templated peptides in aqueous solution. 3. Calculation of the helical propogation constant *s* from the template stability constants t/c for Ac-Hel$_1$-Ala$_n$-OH, $n = 1-6$. *J. Am. Chem. Soc. 118*, 4249–4255.
18. Serrano, L., Sancho, J., Hirshberg, M., and Fersht, A. R. (1992) α-Helix stability in proteins I. Empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. *J. Mol. Biol. 227*, 544–559.
19. Horovitz, A., Matthews, J. M., and Fersht, A. R. (1992) α-Helix stability in proteins II. Factors that influence stability at an internal position. *J. Mol. Biol. 227*, 560–568.
20. Blaber, M., Zhang, X.-j., and Matthews, B. W. (1993) Structural basis of amino acid α helix propensity. *Science 260*, 1637–1640.
21. Park, S.-H., Shalongo, W., and Stellwagen, E. (1993) Residue helix parameters obtained from dichroic analysis of peptides of defined sequence. *Biochemistry 32*, 7048–7053.
22. Muñoz, V., and Serrano, L. (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol. 1*, 399–409.
23. Zhou, N. E., Kay, C. M., Sykes, B. D., and Hodges, R. S. (1993) A single-stranded amphipathic α-helix in aqueous solution: Design, structural characterization, and its application for determining α-helical propensities of amino acids. *Biochemistry 32*, 6190–6197.
24. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1997) A direct comparison of helix propensity in proteins and peptides. *Proc. Natl. Acad. Sci. U.S.A. 94*, 2833–2837.
25. Butterfield, S. M., Patel, P. R., and Waters, M. L. (2002) Contribution of aromatic interactions to α-helix stability. *J. Am. Chem. Soc. 124*, 9751–9755.
26. Chiu, H.-P., Suzuki, Y., Gullickson, D., Ahmad, R., Kokona, B., Fairman, R., and Cheng, R. P. (2006) Helix propensity of highly fluorinated amino acids. *J. Am. Chem. Soc. 128*, 15556–15557.
27. Chiu, H.-P., and Cheng, R. P. (2007) Chemoenzymatic synthesis of (*S*)-hexafluoroleucine and (*S*)-tetrafluoroleucine. *Org. Lett. 9*, 5517–5520.
28. Serrano, L., and Fersht, A. R. (1989) Capping and α-helix stability. *Nature 342*, 296–299.
29. Lyu, P. C., Zhou, H. X., Jelveh, N., Wemmer, D. E., and Kallenbach, N. R. (1992) Position-dependent stabilizing effects in α-helices: N-Terminal capping in synthetic model peptides. *J. Am. Chem. Soc. 114*, 6560–6562.
30. Lyu, P. C., Wemmer, D. E., Zhou, H. X., Pinker, R. J., and Kallenbach, N. R. (1993) Capping interactions in isolated α helices: Position-dependent substitution effects and structure of a serine-capped peptide helix. *Biochemistry 32*, 421–425.
31. Zhou, H. X., Lyu, P. C., Wemmer, D. E., and Kallenbach, N. R. (1994) Structure of a C-terminal α-helix cap in a synthetic peptide. *J. Am. Chem. Soc. 116*, 1139–1140.
32. Bell, J. A., Becktel, W. J., Sauer, U., Baase, W. A., and Matthews, B. W. (1992) Dissection of helix capping in T4 lysozyme by structural and thermodynamic analysis of six amino acid substitutions at Thr 59. *Biochemistry 31*, 3590–3596.
33. Chakrabartty, A., Doig, A. J., and Baldwin, R. L. (1993) Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci. U.S.A. 90*, 11332–11336.
34. Doig, A. J., Chakrabartty, A., Klingler, T. M., and Baldwin, R. L. (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.
35. Doig, A. J., and Baldwin, R. L. (1995) N- and C-capping preferences of all 20 amino acids in α-helical peptides. *Protein Sci. 4*, 1325–1336.
36. Forood, B., Feliciano, E. J., and Nambiar, K. P. (1993) Stabilization of α-helical structures in short peptides via end capping. *Proc. Natl. Acad. Sci. U.S.A. 90*, 838–842.
37. Zhukovsky, E. A., Mulkerrin, M. G., and Presta, L. G. (1994) Contribution to global protein stabilization of the N-capping box in human growth hormone. *Biochemistry 33*, 9856–9864.
38. Andersen, N. H., and Tong, H. (1997) Empirical parameterization of a model for predicting peptide helix/coil equilibrium populations. *Protein Sci. 6*, 1920–1936.
39. Schneider, J. P., and DeGrado, W. F. (1998) The design of efficient α-helical C-capping auxiliaries. *J. Am. Chem. Soc. 120*, 2764–2767.
40. Marqusee, S., and Baldwin, R. L. (1987) Helix stabilization by Glu^-...Lys^+ salt bridges in short peptides of *de novo* design. *Proc. Natl. Acad. Sci. U.S.A. 84*, 8898–8902.
41. Scholtz, J. M., Qian, H., Robbins, V. H., and Baldwin, R. L. (1993) The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry 32*, 9668–9676.
42. Huyghues-Despointes, B. M. P., Scholtz, J. M., and Baldwin, R. L. (1993) Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci. 2*, 80–85.
43. Padmanabhan, S., and Baldwin, R. L. (1994) Helix-stabilizing interaction between tyrosine and leucine or valine when the spacing in *i*, *i*+*4*. *J. Mol. Biol. 241*, 706–713.
44. Padmanabhan, S., and Baldwin, R. L. (1994) Tests for helix-stabilizing interations between various nonpolar side chains in alanine-based peptides. *Protein Sci. 3*, 1992–1997.
45. Stapley, B. J., Rohl, C. A., and Doig, A. J. (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.
46. Luo, P., and Baldwin, R. L. (2002) Origin of the different strengths of the (i,i+4) and (i,i+3) leucine pair interactions in helices. *Biophys. Chem. 96*, 103–108.
47. Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W., and Kallenbach, N. R. (1991) The helix-coil transition in heterogeneous peptides with specific side-chain interactions: Theory and comparison with CD spectral data. *Biopolymers 31*, 1605–1614.
48. Lyu, P. C., Gans, P. J., and Kallenbach, N. R. (1992) Energetic contribution of solvent-exposed ion pairs to alpha-helix structure. *J. Mol. Biol. 223*, 343–350.
49. Mayne, L., Englander, S. W., Qiu, R., Yang, J., Gong, Y., Spek, E. J., and Kallenbach, N. R. (1998) Stabilizing effect of a multiple salt bridge in a prenucleated peptide. *J. Am. Chem. Soc. 120*, 10643–10645.
50. Olson, C. A., Shi, Z., and Kallenbach, N. R. (2001) Polar interactions with aromatic side chains in α-helical peptides: CH···O H-bonding and cation-π interactions. *J. Am. Chem. Soc. 123*, 6451–6452.

51. Shi, Z., Olson, C. A., Bell, A. J., Jr., and Kallenbach, N. R. (2001) Stabilization of α-helix structure by polar side-chain interactions: Complex salt bridges, cation-π interactions, and C-H···O H-bonds. *Biopolymers 60*, 366–380.
52. Olson, C. A., Spek, E. J., Shi, Z., Vologodskii, A., and Kallenbach, N. R. (2001) Cooperative helix stabilization by complex Arg-Glu salt bridges. *Proteins: Struct., Funct., Genet. 44*, 123–132.
53. Shi, Z., Olson, C. A., and Kallenbach, N. R. (2002) Cation−π interaction in model α-helical peptides. *J. Am. Chem. Soc. 124*, 3284–3291.
54. Shi, Z., Olson, C. A., Bell, A. J., Jr., and Kallenbach, N. R. (2002) Non-classical helix stabilizing interactions: C−H···O H-bonding between Phe and Glu side chains in α-helical peptides. *Biophys. Chem. 101−102*, 267–279.
55. Merutka, G., Shalongo, W., and Stellwagen, E. (1991) A model peptide with enhanced helicity. *Biochemistry 30*, 4245–4248.
56. Merutka, G., and Stellwagen, E. (1991) Effect of amino acid ion pairs on peptide helicity. *Biochemistry 30*, 1591–1594.
57. Stellwagen, E., Park, S. H., Shalongo, W., and Jain, A. (1992) The contribution of residue ion-pairs to the helical stability of a model peptide. *Biopolymers 32*, 1193–1200.
58. Fisinger, S., Serrano, L., and Lacroix, E. (2001) Computational estimation of specific side chain interaction energies in α helices. *Protein Sci. 10*, 809–818.
59. Smith, J. S., and Scholtz, J. M. (1998) Energetics of polar side-chain interactions in helical peptides: Salt effects on ion pairs and hydrogen bonds. *Biochemistry 37*, 33–40.
60. Andrew, C. D., Penel, S., Jones, G. R., and Doig, A. J. (2001) Stabilizing nonpolar/polar side-chain interactions in the α-helix. *Proteins: Struct., Funct., Genet. 45*, 449–455.
61. Andrew, C. D., Bhattacharjee, S., Kokkoni, N., Hirst, J. D., Jones, G. R., and Doig, A. J. (2002) Stabilizing interactions between aromatic and basic side chains in α-helical peptides and proteins. Tyrosine effects on helix circular dichroism. *J. Am. Chem. Soc. 124*, 12706–12714.
62. Iqbalsyah, T. M., and Doig, A. J. (2005) Pairwise coupling in an Arg-Phe-Met triplet stabilizes α-helical peptide via shared rotamer preferences. *J. Am. Chem. Soc. 127*, 5002–5003.
63. Iqbalsyah, T. M., and Doig, A. J. (2005) Anticooperativity in a Glu-Lys-Glu salt bridge triplet in an isolated α-helical peptide. *Biochemistry 44*, 10449–10456.
64. Errington, N., and Doig, A. J. (2005) A phosphoserine-lysine salt bridge within an α-helical peptide, the strongest α-helix side-chain interaction measured to date. *Biochemistry 44*, 7553–7558.
65. Tsou, L. K., Tatko, C. D., and Waters, M. L. (2002) Simple cation-π interaction between a phenyl ring and a protonated amine stabilizes an α-helix in water. *J. Am. Chem. Soc. 124*, 14917–14921.
66. Cheng, R. P., Girinath, P., and Ahmad, R. (2007) Effect of lysine side chain length on intra-helical glutamate-lysine ion pairing interactions. *Biochemistry 46*, 10528–10537.
67. Kirshenbaum, K., Zuckermann, R. N., and Dill, K. A. (1999) Designing polymers that mimic biomolecules. *Curr. Opin. Struct. Biol. 9*, 530–535.
68. Seebach, D., and Matthews, J. L. (1997) β-Peptides: A surprise at every turn. *Chem. Commun.,* 2015–2022.
69. Cheng, R. P., Gellman, S. H., and DeGrado, W. F. (2001) β-Peptides: From structure to function. *Chem. Rev. 101*, 3219–3232.
70. Moore, J. S. (1997) Shape-persistent molecular architectures of nanoscale dimension. *Acc. Chem. Res. 30*, 402–413.
71. Sanford, A. R., Yamato, K., Yang, X., Yuan, L., Han, Y., and Gong, B. (2004) Well-defined secondary structures. Information-storing molecualr duplexes and helical foldamers based on unnatural peptide backbones. *Eur. J. Biochem. 271*, 1416–1425.
72. Huc, I. (2004) Aromatic oligoamide foldamers. *Eur. J. Org. Chem.,* 17–29.
73. Arvidsson, P. I., Rueping, M., and Seebach, D. (2001) Design, machine synthesis, and NMR-solution structure of a β-heptapeptide forming a salt-bridge stabilised 3_{14}-helix in methanol and in water. *Chem. Commun.,* 649–650.
74. Rueping, M., Mahajan, Y. R., Jaun, B., and Seebach, D. (2004) Design, synthesis and structural investigations of a β-peptide forming a 3_{14}-helix stabilized by electrostatic interactions. *Chem.—Eur. J. 10*, 1607–1615.
75. Cheng, R. P., and DeGrado, W. F. (2001) De novo design of a monomeric helical β-peptide stabilized by electrostatic interactions. *J. Am. Chem. Soc. 123*, 5162–5163.
76. Hart, S. A., Bahadoor, A. B. F., Matthews, E. E., Qiu, X. J., and Schepartz, A. (2003) Helix macrodipole control of β^3-peptide 14-helix stability in water. *J. Am. Chem. Soc. 125*, 4022–4023.
77. Raguse, T. L., Lai, J. R., and Gellman, S. H. (2003) Environment-independent 14-helix formation in short β-peptides: Striking a balance between shape control and functional diversity. *J. Am. Chem. Soc. 125*, 5592–5593.
78. Raguse, T. L., Lai, J. R., and Gellman, S. H. (2002) Evidence that the β-peptide 14-helix is stabilized by β^3-residues with side-chain branching adjacent to the β-carbon atom. *Helv. Chim. Acta 85*, 4154–4164.
79. Kritzer, J. A., Tirado-Rives, J., Hart, S. A., Lear, J. D., Jorgensen, W. L., and Schepartz, A. (2005) Relationship between side chain structure and 14-helix stability of β^3-peptides in water. *J. Am. Chem. Soc. 127*, 167–178.
80. Zimm, B. H., and Bragg, J. K. (1959) Theory of the phase transitions between helix and random coil in polypeptide chains. *J. Chem. Phys. 31*, 526–535.
81. Lifson, S., and Roig, A. (1961) On the theory of helix-coil transition in polypeptides. *J. Chem. Phys. 34*, 1963–1974.
82. Muñoz, V., and Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. 2. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol. 245*, 275–296.
83. Muñoz, V., and Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. 3. Temperature and pH-dependence. *J. Mol. Biol. 245*, 297–308.
84. Renold, P., Tsang, K. Y., Shimizu, L. S., and Kemp, D. S. (1996) For short alanine-lysine peptides the helical propensities of lysine residues (*s* values) are strongly temperature dependent. *J. Am. Chem. Soc. 118*, 12234–12235.
85. Job, G. E., Kennedy, R. J., Heitmann, B., Miller, J. S., Walker, S. M., and Kemp, D. S. (2006) Temperature- and length-dependent energetics of formation of polyalanine helices in water: Assignment of $w_{Ala}(n, T)$ and temperature-dependent CD ellipticity standards. *J. Am. Chem. Soc. 128*, 8227–8233.
86. Moreau, R. J., Schubert, C. R., Nasr, K. A., Török, M., Miller, J. S., Kennedy, R. J., and Kemp, D. S. (2009) Context-independent, temperature-dependent helical propensities for amino acid residues. *J. Am. Chem. Soc. 131*, 13107–13116.
87. Luo, P., and Baldwin Robert, L. (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluorethanol/water mixtures back to water. *Biochemistry 36*, 8413–8421.
88. Luo, P., and Baldwin, R. L. (1999) Interaction between water and polar groups of the helix backbone: An important determinant of helix propensities. *Proc. Natl. Acad. Sci. U.S.A. 96*, 4930–4935.
89. Scheraga, H. A., Vila, J. A., and Ripoll, D. R. (2002) Helix−coil transitions re-visited. *Biophys. Chem. 101−102*, 255–265.
90. Rohl, C. A., Chakrabartty, A., and Baldwin, R. L. (1996) Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci. 5*, 2623–2637.
91. Kennedy, R. J., Tsang, K.-Y., and Kemp, D. S. (2002) Consistent helicities from CD and template t/c data for N-templated polyalanines: Progress towards resolution of the alanine helicity problem. *J. Am. Chem. Soc. 124*, 934–944.
92. Merutka, G., and Stellwagen, E. (1990) Positional independence and additivity of amino acid replacements on helix stability in monomeric peptides. *Biochemistry 29*, 894–898.
93. Ermolenko, D. N., Richardson, J. M., and Makhatadze, G. I. (2003) Noncharged amino acid residues at the solvent-exposed positions in the middle and at the C terminus of the α-helix have the same helical propensity. *Protein Sci. 12*, 1169–1176.
94. Petukhov, M., Muñoz, V., Yumoto, N., Yoshikawa, S., and Serrano, L. (1998) Position dependence of non-polar amino acid intrinsic helical propensities. *J. Mol. Biol. 278*, 279–289.
95. Petukhov, M., Uegaki, K., Yumoto, N., Yoshikawa, S., and Serrano, L. (1999) Position dependence of amino acid intrinsic helical propensities II: Non-charged polar residues: Ser, Thr, Asn, and Gln. *Protein Sci. 8*, 2144–2150.
96. Petukhov, M., Uegaki, K., Yumoto, N., and Serrano, L. (2002) Amino acid intrinsic α-helical propensities III: Positional dependence at several positions of C terminus. *Protein Sci. 11*, 766–777.
97. Fields, G. B., and Noble, R. L. (1990) Solid-phase peptide-synthesis utilizing 9-fluorenylmethoxycarbonyl amino-acids. *Int. J. Pept. Protein Res. 35*, 161–214.
98. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci. 4*, 2411–2423.
99. Robert, C. H. (1990) A hierarchical nesting approach to describe the stability of α-helices with side-chain interactions. *Biopolymers 30*, 335–347.

100. Hobohm, U., and Sander, C. (1994) Enlarged representative set of protein structures. *Protein Sci. 3*, 522–524.
101. Griep, S., and Hobohm, U. (2010) PDBselect 1992−2009 and PDBfilter-select. *Nucleic Acids Res. 38*, D318–D319.
102. Gunasekaran, K., Nagarajaram, H. A., Ramakrishnan, C., and Balaram, P. (1998) Stereochemical punctuation marks in protein structures: Glycine and proline containing helix stop signals. *J. Mol. Biol. 275*, 917–932.
103. Engel, D. E., and DeGrado, W. F. (2004) Amino acid propensities are position-dependent throughout the length of α-helices. *J. Mol. Biol. 337*, 1195–1205.
104. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577–2637.
105. Efron, B., and Gong, G. (1983) A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am. Stat. 37*, 36–48.
106. Kuebler, R. R., and Smith, H. (1976) in Statistics: A beginning, pp 302, John Wiley & Sons, Inc., New York.
107. Klugh, H. E. (1970) in Statistics: The essentials for research, pp 350, John Wiley & Sons, Inc., New York.
108. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*, 1948–1954.
109. Chakrabartty, A., Kortemme, T., Padmanabhan, S., and Baldwin, R. L. (1993) Aromatic side-chain contribution to far-ultraviolet circular-dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry 32*, 5560–5565.
110. Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., and Baldwin, R. L. (1990) Relative helix-forming tendencies of nonpolar amino acids. *Nature 344*, 268–270.
111. Aurora, R., and Rose, G. D. (1998) Helix capping. *Protein Sci. 7*, 21–38.
112. Gallivan, J. P., and Dougherty, D. A. (1999) Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A. 96*, 9459–9464.
113. Charkrabartty, A., Schellman, J. A., and Baldwin, R. L. (1991) Large differences in the helix propensities of alanine and glycine. *Nature 351*, 586–588.
114. Shalongo, W., Dugad, L., and Stellwagen, E. (1994) Distribution of helicity within the model peptide acetyl(AAQAA)$_3$amide. *J. Am. Chem. Soc. 116*, 8288–8293.
115. Tirado-Rives, J., and Jorgensen, W. L. (1991) Molecular dynamics simulations of the unfolding of an α-helical analogue of ribonuclease A S-peptide in water. *Biochemistry 30*, 3864–3871.
116. Soman, K. V., Karimi, A., and Case, D. A. (1991) Unfolding of an α-helix in water. *Biopolymers 31*, 1351–1361.
117. Brooks, C. L., III, and Case, D. A. (1993) Simulations of peptide conformational dynamics and thermodynamics. *Chem. Rev. 93*, 2487–2502.
118. Hirst, J. D., and Brooks, C. L., III (1995) Molecular-dynamics simulations of isolated helices of myoglobin. *Biochemistry 34*, 7614–7621.
119. Young, W. S., and Brooks, C. L., III (1996) A microscopic view of helix propagation: N- and C-terminal helix growth in alanine helices. *J. Mol. Biol. 259*, 560–572.
120. Kanehisa, M. I., and Tsong, T. Y. (1980) Local hydrophobicity stabilizes secondary structures in proteins. *Biopolymers 19*, 1617–1628.
121. Cid, H., Bunster, M., Arriagada, E., and Campos, M. (1982) Prediction of secondary structure of proteins by means of hydrophobicity profiles. *FEBS Lett. 150*, 247–254.
122. Wuilmart, C., and Urbain, J. (1984) α secondary structures generate weak but recurrent periodicity in proteins. *Eur. J. Biochem. 139*, 35–40.
123. Horne, D. S. (1988) Prediction of protein helix content from an autocorrelation analysis of sequence hydrophobicities. *Biopolymers 27*, 451–477.
124. Reyes, V. E., Phillips, L., Humphreys, R. E., and Lew, R. A. (1989) Prediction of protein helices with a derivative of the strip-of-helix hydrophobicity algorithm. *J. Biol. Chem. 264*, 12854–12858.
125. Richardson, J. S., and Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of α helices. *Science 240*, 1648–1652.
126. Dasgupta, S., and Bell, J. A. (1993) Design of helix ends. Amino acid preferences, hydrogen bonding, and electrostatic interactions. *Int. J. Pept. Protein Res. 41*, 499–511.
127. Kumar, S., and Bansal, M. (1998) Dissecting α-helices: Position-specific analysis of α-helices in globular proteins. *Proteins 31*, 460–476.
128. Baker, E. N., and Hubbard, R. E. (1984) Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol. 44*, 97–179.
129. Miick, S. M., Martinez, G. V., Fiori, W. R., Todd, A. P., and Millhauser, G. L. (1992) Short alanine-based peptides may form 3_{10}-helices and not α-helices in aqueous solution. *Nature 359*, 653–655.
130. Fiori, W. R., Miick, S. M., and Millhauser, G. L. (1993) Increasing sequence length favors α-helix over 3_{10}-helix in alanine-based peptides: Evidence for a length-dependent structural transition. *Biochemistry 32*, 11957–11962.
131. Bolin, K. A., and Millhauser, G. L. (1999) α and 3_{10}: The split personality of polypeptide helices. *Acc. Chem. Res. 32*, 1027–1033.
132. Sun, J. K., and Doig, A. J. (1998) Addition of side-chain interactions to 3_{10}-helix/coil and α-helix/3_{10}-helix/coil theory. *Protein Sci. 7*, 2374–2383.
133. Lyulin, A. V., and Michels, M. A. J. (2002) Molecular dynamics simulation of bulk atactic polystyrene in the vicinity of T_g. *Macromolecules 35*, 1463–1472.
134. Eslami, H., and Müller-Plathe, F. (2009) Structure and mobility of poly(ethylene terephthalate): A molecular dynamics simulation study. *Macromolecules 42*, 8241–8250.

Supporting Information

Positional Effects on Helical Ala-Based Peptides

Richard P. Cheng,* Prashant Girinath, Yuta Suzuki, Hsiou-Ting Kuo, Hao-Chun Hsu, Wei-Ren Wang, Po-An Yang, Donald Gullickson, Cheng-Hsun Wu, Marc J. Koyack, Hsien-Po Chiu, Yi-Jen Weng, Pier Hart, Bashkim Kokona, Robert Fairman, Tzu-En Lin, Olivia Barrett

KAla (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 733.9 mg (0.150 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 966.1 mg of resin (72.3% yield). The cleavage yielded 289.4 mg of crude peptide (99.4% yield, 58.1% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.5% purity. Retention time on analytical RP-HPLC was 23.8 min. Retention time on GPC was 32.2 min (monomeric). The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{75}H_{128}N_{24}O_{21}$ [MH^+]: 1701.976; observed: 1701.969. Apparent molecular weight from sedimentation equilibrium analysis: 1450±100.

KLeu6 (Ac-Tyr-Gly-Gly-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 213.4 mg (0.051 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 279.7 mg of resin (71.1% yield). The cleavage yielded 160.3 mg of crude peptide (>99% yield, 43.1% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.2% purity. Retention time on analytical RP-HPLC was 26.3 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{78}H_{134}N_{24}O_{21}$ [MH^+]: 1744.023; observed: 1744.001. Apparent molecular weight from sedimentation equilibrium analysis: 2120±230.

KLeu11 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 489.9 mg (0.100 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 648.8 mg of resin (72.8% yield). The cleavage yielded 178.6 mg of crude peptide (99.5% yield, 59.7% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.7% purity. Retention time on analytical RP-HPLC was 26.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{78}H_{134}N_{24}O_{21}$ [MH^+]: 1744.023; observed: 1744.168. Apparent molecular weight from sedimentation equilibrium analysis: 1800±180.

KLeu16 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-NH_2). The peptide was synthesized using 224.1 mg (0.052 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 252.6 mg of resin (35.5% yield). The cleavage yielded 71.3 mg of crude peptide (>99% yield, 51.3% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.6% purity. Retention time on analytical RP-HPLC was 26.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{78}H_{134}N_{24}O_{21}$ [MH^+]: 1744.023; observed: 1743.935. Apparent molecular weight from sedimentation equilibrium analysis: 3100±450. Although the apparent molecular weight exceeds a monomer, model analysis revealed that this peptide is a monomer in solution, because the data was not fit better using dimers or other larger species.

KLeu611 (Ac-Tyr-Gly-Gly-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 278.5 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 341.3 mg of resin (64.5% yield). The cleavage yielded 102.0 mg of crude peptide (>99% yield, 58.3% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.9% purity. Retention time on analytical RP-HPLC was 29.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{140}N_{24}O_{21}$ [MH^+]: 1786.070; observed: 1785.062. Apparent molecular weight from sedimentation equilibrium analysis: 2010±210.

KLeu616 (Ac-Tyr-Gly-Gly-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-NH_2). The peptide was synthesized using 278.9 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 339.7 mg of resin (62.3% yield). The cleavage yielded 104.5 mg of crude peptide (>99% yield, 54.9% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.9% purity. Retention time on analytical RP-HPLC was 30.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{140}N_{24}O_{21}$ [MH^+]: 1786.070; observed: 1785.546. Apparent molecular weight from sedimentation equilibrium analysis: 1770±210.

KLeu1116 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-NH_2). The peptide was synthesized using 279.1 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 371.0 mg of resin (74.3% yield). The cleavage yielded 111.1 mg of crude peptide (>99% yield, 59.2% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.9% purity. Retention time on analytical RP-HPLC was 31.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{140}N_{24}O_{21}$ [MH^+]: 1786.070; observed: 1785.253. Apparent molecular weight from sedimentation equilibrium analysis: 2190±190.

KLeu61116 (Ac-Tyr-Gly-Gly-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-Ala-Leu-Ala-Ala-Lys-NH_2). The peptide was synthesized using 279.8 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 309.1 mg of resin (28.7% yield). The cleavage yielded 79.8 mg of crude peptide (>99% yield, 59.2% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.4% purity. Retention time on analytical RP-HPLC was 33.4 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{84}H_{146}N_{24}O_{21}$ [MH^+]: 1828.117; observed: 1827.109. Apparent molecular weight from sedimentation equilibrium analysis: 1450±170.

KPhe6 (Ac-Tyr-Gly-Gly-Lys-Ala-Phe-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 218.2 mg (0.050 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 248.4 mg of resin (37.3% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.8% purity. Retention time on analytical RP-HPLC was 26.6 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{132}N_{24}O_{21}$ [MH^+]: 1778.008; observed: 1778.139. Apparent molecular weight from sedimentation equilibrium analysis: 1680±110.

KPhe11 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Phe-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 220.4 mg (0.050 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 244.4 mg of resin (25.3% yield). The cleavage yielded 82.4 mg of crude peptide (>99% yield, 31.3% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 98.9% purity. Retention time on analytical RP-HPLC was 27.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{132}N_{24}O_{21}$ [MH^+]: 1778.008; observed: 1778.139. Apparent molecular weight from sedimentation equilibrium analysis: 1500±100.

KPhe16 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Phe-Ala-Ala-Lys-NH_2). The peptide was synthesized using 219.9 mg (0.050 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 262.8 mg of resin (47.9% yield). The cleavage yielded 101.3 mg of crude peptide (>99% yield, 29.2% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 99.3% purity. Retention time on analytical RP-HPLC was 27.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{132}N_{24}O_{21}$ [MH^+]: 1778.008; observed: 1778.691. Apparent molecular weight from sedimentation equilibrium analysis: 1740±120.

KPff6 (Ac-Tyr-Gly-Gly-Lys-Ala-Pff-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 218.4 mg (0.050 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 273.0 mg of resin (53.8% yield). The cleavage yielded 59.7 mg of crude peptide (>99% yield, 37.0% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 98.8% purity. Retention time on analytical RP-HPLC was 28.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{127}N_{24}O_{21}F_5$ [MH^+]: 1867.961; observed: 1866.952. Apparent molecular weight from sedimentation equilibrium analysis: 1950±130.

KPff11 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Pff-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 220.4 mg (0.050 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 278.5 mg of resin (50.6% yield). The cleavage yielded 66.0 mg of crude peptide (98.9% yield, 17.0% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 98.8% purity. Retention time on analytical RP-HPLC was 29.4 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{127}N_{24}O_{21}F_5$ [MH^+]: 1867.961; observed: 1867.936. Apparent molecular weight from sedimentation equilibrium analysis: 2700±250. Although the apparent molecular weight exceeds a monomer, model analysis shows approximately 5% of the peptide may be a higher order aggregate in solution, so the major species in solution is a monomer.

KPff16 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Pff-Ala-Ala-Lys-NH_2). The peptide was synthesized using 216.8 mg (0.050 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 276.3 mg of resin (49.2% yield). The cleavage yielded 165.6 mg of crude peptide (>99% yield, 19.2% purity). The peptide was purified by preparative RP-HPLC using a C18 column to 98.7% purity. Retention time on analytical RP-HPLC was 30.2 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{81}H_{127}N_{24}O_{21}F_5$ [MH^+]: 1867.961; observed: 1867.031. Apparent molecular weight from sedimentation equilibrium analysis: 1810±160.

KGly6 (Ac-Tyr-Gly-Gly-Lys-Ala-Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 300.0 mg (0.06 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 336.4 mg of resin (32.5% yield). The cleavage yielded 104.9 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.4% purity. Retention time on analytical RP-HPLC was 21.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{74}H_{126}N_{24}O_{21}$ [MH^+]: 1687.961; observed: 1687.792. Apparent molecular weight from sedimentation equilibrium analysis: 1780±150.

KGly11 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala- Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-NH_2). The peptide was synthesized using 513.3 mg (0.10 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 630.7 mg of resin (61.6% yield). The cleavage yielded 179.2 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.6% purity. Retention time on analytical RP-HPLC was 21.7 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{74}H_{126}N_{24}O_{21}$ [MH^+]: 1687.961; observed: 1687.340.

KGly16 (Ac-Tyr-Gly-Gly-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Gly-Ala-Ala-Lys-NH_2). The peptide was synthesized using 255 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 231 mg of resin. The cleavage yielded 82.5 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.4% purity. Retention time on analytical RP-HPLC was 20.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{74}H_{126}N_{24}O_{21}$ [MH^+]: 1687.961; observed: 1687.838. Apparent molecular weight from sedimentation equilibrium analysis: 1830±150.

NCapAla (Ac-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Gly-Gly-Tyr-NH_2). The peptide was synthesized using 248.3 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 301.0 mg of resin (59.2 % yield). The cleavage yielded 69.8 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.4% purity. Retention time on analytical RP-HPLC was 23.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{69}H_{116}N_{22}O_{20}$ [MH^+]: 1573.881; observed: 1573.708.

NCapLeu (Ac-Leu-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Gly-Gly-Tyr-NH_2). The peptide was synthesized using 238.1 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 295.0 mg of resin (64.9 % yield). The cleavage yielded 61.1 mg of crude peptide (96.0 % yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.6% purity. Retention time on analytical RP-HPLC was 27.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{72}H_{122}N_{22}O_{20}$ [MH^+]: 1615.928; observed: 1615.821.

NCapPhe (Ac-Phe-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Gly-Gly-Tyr-NH_2). The peptide was synthesized using 248.6 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 282.1 mg of resin (36.1 % yield). The cleavage yielded 72.6 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to

99.7% purity. Retention time on analytical RP-HPLC was 27.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{75}H_{120}N_{22}O_{20}$ [MH^+]: 1649.913; observed: 1649.825.

NCapPff (Ac-Pff-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Gly-Gly-Tyr-NH_2). The peptide was synthesized using 123.3 mg (0.026 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 142.4 mg of resin (36.0 % yield). The cleavage yielded 36.1 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 98.7% purity. Retention time on analytical RP-HPLC was 30.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{75}H_{115}F_5N_{22}O_{20}$ [MH^+]: 1739.886; observed: 1739.908

NCapGly (Ac-Gly-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Gly-Gly-Tyr-NH_2). The peptide was synthesized using 238.1 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 290.8 mg of resin (62.3 % yield). The cleavage yielded 69.8 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.4% purity. Retention time on analytical RP-HPLC was 24.3 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{68}H_{114}N_{22}O_{20}$ [MH^+]: 1559.866; observed: 1559.514.

CCapAla (Ac-Tyr-Gly-Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-NH_2). The peptide was synthesized using 245.5 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 289.6 mg of resin (50.2 % yield). The cleavage yielded 64.4 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 99.4% purity. Retention time on analytical RP-HPLC was 24.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{69}H_{116}N_{22}O_{20}$ [MH^+]: 1573.881; observed: 1573.668.

CCapLeu (Ac-Tyr-Gly-Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Leu-NH_2). The peptide was synthesized using 245.8 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 634.1 mg of resin (>99 % yield). The cleavage yielded 50.7 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C18 column to 98.5% purity. Retention time on analytical RP-HPLC was 28.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{72}H_{122}N_{22}O_{20}$ [MH^+]: 1615.928; observed: 1615.637.

CCapPhe (Ac-Tyr-Gly-Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Phe-NH_2). The peptide was synthesized using 241.3 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 368.3 mg of resin (>99 % yield). The cleavage yielded 39.0 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C18 column to 99.0% purity. Retention time on analytical RP-HPLC was 29.0 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{75}H_{120}N_{22}O_{20}$ [MH^+]: 1649.913; observed: 1649.360.

CCapPff (Ac-Tyr-Gly-Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Pff-NH_2). The peptide was synthesized using 247.7 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin.

The synthesis gave 258.5 mg of resin (11.1 % yield). The cleavage yielded 38.3 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 98.7% purity. Retention time on analytical RP-HPLC was 31.7 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{69}H_{116}N_{22}O_{20}$ [MH^+]: 1739.886; observed: 1738.973.

CCapGly (Ac-Tyr-Gly-Gly-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Ala-Ala-Ala-Lys-Ala-Gly-NH_2). The peptide was synthesized using 244.2 mg (0.05 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 171.0 mg of resin. The cleavage yielded 17.2 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C18 column to 98.5% purity. Retention time on analytical RP-HPLC was 24.2 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for $C_{68}H_{114}N_{22}O_{20}$ [MH^+]: 1559.866; observed: 1559.150.