Circular Dichroism of Model Peptides Emulating the Amphipathic α-Helical Regions of Intermediate Filaments†

N. D. Lazo* and Donald T. Downing

The Marshall Research Laboratories, Department of Dermatology, University of Iowa College of Medicine, Iowa City, Iowa 52242

Received December 13, 1996⊗

ABSTRACT: The *a* and *d* positions of the heptad repeats (*abcdefg*) found in the α-helical sections of intermediate-filament proteins are hydrophobic, and the remaining locations are almost exclusively hydrophilic and often charged. Two synthetic peptides that maximize these features were designed, synthesized, and investigated by circular dichroism for α-helix formation in water and in 50% trifluoroethanol (TFE). A 14-residue peptide, AcNLEELKKKLEELKGNH$_2$ (NLEKG14), had mean residue ellipticities at 222 nm ($[\theta]_{222}$) of $-18\,400 \pm 1000$ and $-37\,200 \pm 1900$ deg cm^2 dmol^{-1}, in water at 2 °C and in 50% TFE at 2 °C, respectively. A longer version of NLEKG14, AcNLEELKKKLEELKQQLEELKKKLEELKQQNH$_2$ (NLEKQ29), had $[\theta]_{222}$ of $-43\,000 \pm 2200$ deg cm^2 dmol^{-1} in water and in 50% TFE at 2 °C. Using $-43\,000$ deg cm^2 dmol^{-1} as $[\theta]_{222}$ for a 100% helix, NLEKG14 in 50% TFE at 25 °C was estimated to be 77% helix. This estimate was confirmed by two-dimensional ^{1}H NMR studies of NLEKG14 in 50% TFE. Comparison with the sequences and conformations found in IF proteins indicates that the α-helical regions in the proteins may be exceptionally stable, but the high values for the ellipticity of α-helices now revealed allow for significant portions of the protein rod regions to be occupied by conformations other than α-helix.

In the classical α-helical conformation of proteins (Crick, 1952, 1953; Pauling & Corey, 1953), a right-handed coil containing 3.6 amino acids per turn is stabilized by hydrogen bonding between peptide carbonyl groups and amide protons in successive turns of the helix. Nonrandom distributions of the amino acids can induce specific properties and higher-order conformations in proteins, most commonly the amphipathic α-helix, in which hydrophobic amino acid side chains are segregated in a curvilinear array along the helix (Cohen & Parry, 1990; Stewart, 1993). Those backbone hydrogen bonds that are thus shielded from water are significantly shortened, resulting in a distortion of the otherwise linear helix. When the amino acids are arranged in a heptadic repeat pattern, where the first and fourth residues are hydrophobic, the distortion produces supercoiling of the helix (Crick, 1952, 1953; Pauling & Corey, 1953; O'Shea et al., 1991). Moreover, the hydrophobic arrays along the helix can induce the lateral association of two, three, or four helices into bundles, which are stabilized by the resulting mutual reduction in exposure of the hydrophobic groups to water (Harbury et al., 1993; Bryson et al., 1995).

The biochemical significance of tertiary and quaternary protein conformations has prompted a wide variety of investigations, including the synthesis of peptides in which specific amphipathic arrays have been evaluated. The design of these peptides has usually been based on fundamental principles (Cohen & Parry, 1990; Stewart, 1993) as well as on native sequences of amino acids, and the efficiencies of α-helix formation and higher-order associations have been evaluated by circular dichroism (CD)[1] and NMR spectroscopies, and by X-ray crystallography. However, instances where synthetic peptides have been designed and synthesized to emulate the sequences found in IF proteins are lacking. Inspection of the sequences of these fibrous proteins indicates that their α-helical regions have hydrophobic residues almost exclusively in the heptadic *a* and *d* locations, while the remaining locations contain predominantly those hydrophilic residues known to stabilize the α-helical conformation. We therefore have evaluated two synthetic peptides that maximize these features: a 14-residue peptide, AcNLEELKKKLEELKGNH$_2$ (NLEKG14), and a 29-residue peptide, AcNLEELKKKLEELKQQLEELKKKLEELKQQNH$_2$ (NLEKQ29). NLEKQ29 in water or in 50% TFE at 2 °C had an $[\theta]_{222}$ of $-43\,000 \pm 2200$ deg cm^2 dmol^{-1}, an ellipticity that is the highest so far reported for an α-helical peptide, and may be useful as a standard against which to evaluate helicities of amphipathic peptides and proteins based on CD spectra.

MATERIALS AND METHODS

Peptide Synthesis and Characterization. Peptide NLEKG14 (100% pure) was synthesized by American Peptide Co. (Sunnyvale, CA) using solid-phase synthesis and *tert*-butyloxycarbonyl (BOC)-benzyl chemistry. Peptide NLEKQ29 (>95% pure) was synthesized by the Protein Structure Facility of the University of Iowa using solid-phase synthesis

†This work has been supported by the National Institutes of Health (AR32374) and by a bequest from the Carl J. Hertzog Foundation.

* To whom correspondence should be addressed. Telephone: 319 335-8077. FAX: 319 335-9559.

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

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; CD, circular dichroism; DQF-COSY, double quantum filtered correlation spectroscopy; IF, intermediate filament; NOESY, nuclear Overhauser enhancement spectroscopy; NMR, nuclear magnetic resonance; TOCSY, total correlation spectroscopy; TFE, 2,2,2-trifluoroethanol; $[\theta]_{208}$, mean residue ellipticity at 208 nm; $[\theta]_{222}$, mean residue ellipticity at 222 nm.

S0006-2960(96)03061-9 CCC: $14.00

and 9-fluorenylmethoxycarbonyl (FMOC) protection and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) activation chemistry. The compositions of both peptides were verified by mass spectrometry and by amino acid analysis.

Circular Dichroism. CD spectra of NLEKG14 and NLEKQ29 were acquired using an Aviv Associates Model 62DS spectrometer. Stoppered optical cells (Starna Cells, Atascadero, CA) with a path length of 1 mm were used for dilute samples (peptide concentrations of <1 mM). Cells with a path length of 0.1 mm were used for more concentrated samples (peptide concentrations of 1−3 mM). Spectra were obtained by taking readings every 0.5 nm with an averaging time of 4−8 s and a bandwidth of 1 nm. Baseline-corrected spectra were smoothed using a third-order least-squares polynomial. Ellipticity measurements were expressed as mean residue ellipticity, $[\theta]$, in units of degrees centimeter squared per decimole, and calculated from the equation:

$$[\theta] = [\theta]_{obs}(MRW)/10lc \quad (1)$$

in which $[\theta]_{obs}$ is the observed ellipticity in degrees, MRW is the mean residue weight of the peptide (molecular weight divided by the number of amino acid residues), c is the concentration of the peptide in grams per liter, and l is the optical path length of the cell in centimeters. The ellipticity was calibrated with (+)-10-camphorsulfonic acid. Peptide concentrations were determined by quantitative amino acid analysis at the University of Iowa Protein Structure Facility. The accuracy of this laboratory for absolute yield of protein was ±4.4% in a 1994 national survey of the Association of Biomolecular Resource Facilities.

Spectra of poly(L-lysine) hydrochloride (Sigma Chemical Co., St. Louis, MO), which is commonly used as a standard (Yang et al., 1986), were also acquired under conditions that favor α-helix formation (pH = 11 at 25 °C). Spectra with double minima at 222 nm ($[\theta]_{222} = -36\,100 \pm 1800$ deg cm^2 dmol^{-1}) and 208 nm ($[\theta]_{208} = -34\,100 \pm 1700$ deg cm^2 dmol^{-1}) were obtained. The ellipticities are consistent with values reported in the literature ($[\theta]_{222} = -35\,700 \pm 2800$ deg cm^2 dmol^{-1}, $[\theta]_{208} = -32\,600 \pm 4000$ deg cm^2 dmol^{-1}) (Greenfield & Fasman, 1969).

Nuclear Magnetic Resonance. All ^{1}H NMR spectra were collected at 25 °C on a Bruker 600-MHz AMX spectrometer. Spectra were referenced with respect to residual solvent peaks, calibrated externally using 2,2-dimethyl-2-silapentane-5-sulfonate as standard.

In all 2D experiments, low-power presaturation of the water signal was applied during the relaxation delay (2 s), as well as during the mixing time in the case of the NOESY experiments. DQF-COSY and NOESY spectra were recorded in the phase-sensitive detection mode using the hypercomplex method (States et al., 1982). TOCSY spectra were collected with a mixing time of 120 ms, and a MLEV17 pulse sequence was used to drive the coherence transfer. Spectra were collected with 2048 points in t_2, with the exception of DQF-COSY in which 4096 points were acquired. A total of 512 experiments in t_1 with 32−64 scans each were collected. The spectral width in both dimensions was ±2717.39 Hz.

Scalar $^3J_{HNa}$ coupling constants were measured from 1D and DQF-COSY spectra. One-dimensional spectra were resolution-enhanced by apodization with a Lorentzian to Gaussian transformation. Coupling constants from DQF-COSY spectra were calculated from antiphase peak separations (Neuhaus et al., 1985).

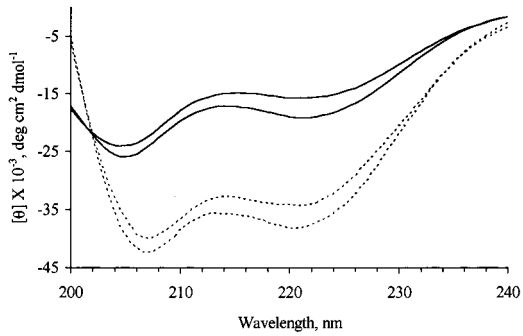

FIGURE 1: CD spectra of NLEKG14 in buffer at 25 °C (upper solid line), buffer at 2 °C (lower solid line), 50% TFE at 25 °C (upper dashed line), and 50% TFE at 2 °C (lower dashed line). Buffer composition is 10 mM sodium phosphate, 1 mM NaN_3, pH 7.

RESULTS AND DISCUSSION

Peptide Design. Peptide NLEKG14 was designed to form an ideal, amphipathic α-helical peptide. The following features were selected to favor the α-helical conformation: (a) an initial asparagine, which is known to favor α-helix initiation (Chakrabartty et al., 1993; Doig & Baldwin, 1995); (b) bulky hydrophobic residues in the *a* and *d* positions of the *abcdefg* heptad repeat unit favor the formation of amphipathic α-helices in which the backbone hydrogen bonds adjacent to the hydrophobic residues are shortened, resulting in twisting of the α-helix to form a coiled-coil (Cohen & Parry, 1990); (c) the like charges (E-E and K-K) that are adjacent in the sequence provide electrostatic repulsion favoring curvature of the chain into a helix; (d) the charged side chains can form ion pairs or hydrogen bonds between successive turns of the helix that are expected to stabilize the conformation (Chakrabartty & Baldwin, 1995); (e) the peptide has no overall charge; and (f) the peptide has no aromatic residues, which can contribute to large CD spectral distortions (Woody & Dunker, 1996).

Peptide NLEKQ29 was designed with the same criteria as for peptide NLEKG14, and in addition, greater chain length was provided to further stabilize the α-helical conformation (Su et al., 1994). In the middle and end of the sequence, pairs of glutamines replaced what might have been pairs of lysines in order to maintain the overall electrical neutrality while preserving the capabilities for side-chain hydrogen bonding between adjacent turns of the helix.

CD of NLEKG14 and NLEKQ29. Figure 1 presents representative CD spectra of NLEKG14. The spectra of the peptide in water exhibit features indicative of the presence of a mixture of α-helix and random coil, i.e., negative bands centered at 221 nm and near 205 nm. The addition of 50% TFE resulted in an increase in the amount of α-helix at the expense of random coil. The intensity of the band at 221 nm increased while the band below 210 nm is now at 207 nm (Figure 1). This observation is consistent with the helix-enhancing effect of TFE on short peptides with α-helical propensity (Nelson & Kallenbach, 1986; Sönnichsen et al., 1992). An isodichroic point near 202 nm was observed when the spectra acquired in water and 50% TFE were superimposed (Figure 1), indicative of a transition between random

Table 1: CD Spectral Features of NLEKG14 and NLEKQ29

	T (°C)	$-[\theta]_{222}$[a]	$-[\theta]_{208}$[a]	R[b]
NLEKG14				
buffer	2	18.4 ± 1.0	22.0 ± 1.1	0.84
buffer	25	14.7 ± 0.7	19.6 ± 1.0	0.75
50% TFE	2	37.2 ± 1.9	41.5 ± 2.1	0.90
50% TFE	25	33.3 ± 1.7	38.5 ± 1.9	0.86
NLEKQ29				
buffer	2	43.0 ± 2.2	39.2 ± 2.0	1.10
buffer	25	40.3 ± 2.0	38.5 ± 1.9	1.05
buffer	75	33.4 ± 1.7	35.2 ± 1.8	0.95
50% TFE	2	43.5 ± 2.2	44.3 ± 2.2	0.98
50% TFE	25	40.0 ± 2.0	42.5 ± 2.1	0.94
50% TFE	75	32.9 ± 1.6	38.4 ± 1.9	0.86

[a] $\times 10^{-3}$ deg cm^2 dmol^{-1}; values are ±5%. [b] $[\theta]_{222}/[\theta]_{208}$.

coil and α-helix (Merutka et al., 1990). Table 1 shows $[\theta]_{222}$, $[\theta]_{208}$, and ellipticity ratios R ($[\theta]_{222}/[\theta]_{208}$) for NLEKG14. All of the ratios R are less than 1, which suggests that the peptide does not form a coiled-coil (Cooper & Woody, 1990). Spectra of NLEKG14 in water or 50% TFE at 25 °C were independent of concentration in the range of 40 μM to 3 mM (spectra not shown). The $-[\theta]_{222}$ of NLEKG14 in water or 50% TFE increased when the sample temperature was lowered from 25 to 2 °C. The increases were 25% and 12% in water and in 50% TFE, respectively.

The amount of α-helix present can be determined from $[\theta]_{222}$ by taking the ratio of $[\theta]_{222}$(obs) to $[\theta]_{222}$(100% helix). Determination of $[\theta]_{222}$(100% helix) is complicated by the fact that CD for an α-helix is dependent on the number of residues. To account for this length dependence, the following formula has been used to determine $[\theta]_{222}$(100% helix) (Chen et al., 1974):

$$X_H^n = X_H^\infty\left(1 - \frac{k}{n}\right) \quad (2)$$

where X_H^∞ is the molar ellipticity for an infinite helix, k is a wavelength-dependent factor (equal to 2.5 at 222 nm), and n is the number of residues in the helix. A range of values has been used for X_H^∞: −37 400 (Chen et al., 1974) to −40 000 (Scholtz et al., 1991) deg cm^2 dmol^{-1}. Using the higher limit, $[\theta]_{222}$(100% α-helix) for a 14-residue peptide is approximately equal to −33 000 deg cm^2 dmol^{-1}. Using this standard, the helicity estimates for NLEKG14 in 50% TFE at 2 and 25 °C are >100%!

To obtain better helicity estimates, CD studies of a longer version of NLEKG14 were performed. Figure 2 presents representative CD spectra of NLEKQ29 in water and in 50% TFE. The spectra show the presence of a predominantly α-helical peptide as indicated by intense bands at 222 nm (n−π* transitions of the α-helix) and at 208 nm ($\pi-\pi$* transitions of the α-helix). Table 1 shows $[\theta]_{222}$, $[\theta]_{208}$, and ellipticity ratios R determined for the peptide at various conditions. The values of R for the peptide in buffer at 2 and 25 °C are greater than 1, indicating the presence of a coiled-coil (Cooper & Woody, 1990; Lau et al., 1984; Su et al., 1994). Additional evidence for this was obtained from the concentration dependence of the spectra at 25 °C (spectra not shown), indicative of a monomer−oligomer equilibrium (Su et al., 1994). At 75 °C, R is less than 1, indicating the conversion of the coiled-coil to a single-stranded α-helical structure. The change in the value of R, i.e., >1 at 25 and 2 °C to <1 at 75 °C, suggests that the formation of a coiled-coil from NLEKQ29 is due to a hydrophobic effect.

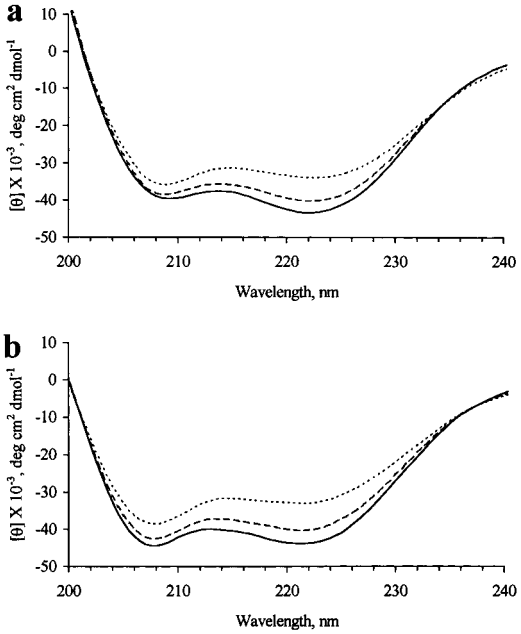

FIGURE 2: CD spectra of NLEKQ29 (a) in buffer at 2 °C (solid line), at 25 °C (dashed line), and at 75 °C (dotted line); and (b) in 50% TFE at 2 °C (solid line), at 25 °C (dashed line), and at 75 °C (dotted line).

All of the values of R for NLEKQ29 in 50% TFE were less than 1 (Table 1), indicating the presence of a single-stranded α-helical peptide. Furthermore, spectra acquired in the presence of 50% TFE were independent of concentration in the range of 6−60 μM, suggesting a maximally stabilized monomer. TFE is known to disrupt hydrophobic interactions (Thomas & Dill, 1993) which contribute to the stability of coiled-coils (Greenfield & Hitchcock-DeGregori, 1995).

Table 2 lists a number of amphipathic peptides that have been synthesized and examined, by previous investigators, using CD spectroscopy. The 14-residue peptide NLEKG14 (peptide 1 in Table 2) that was prepared for the present study was found to have a $[\theta]_{222}$ in water similar to that of previously reported peptides (peptides 4−5, 7−8, and 11 in Table 2) of similar chain length. This observation supported the prediction that limiting hydrophobic groups to the *a* and *d* positions, as in NLEKG14, would efficiently promote α-helix formation. Conversely, additional hydrophobic residues in other heptadal positions do not detract from the helicity, as evidenced by peptides 3−11 in Table 2.

As demonstrated by previous investigators, the helicity of synthetic peptides is greatly influenced by the number of amino acids in the chain, as with peptide 17 in Table 2, which has 35 amino acids and a $[\theta]_{222}$ of −32 900 deg cm^2 dmol^{-1} in water at 20 °C. However, the 29-residue peptide NLEKQ29 in the present study (peptide 2, Table 2) had a considerably higher $[\theta]_{222}$ in water at 25 °C (−40 300 ± 2000 deg cm^2 dmol^{-1}), perhaps because peptide 17 (Table 2) contained 5 glycine residues, which are thought to destabilize α-helices. The $[\theta]_{222}$ of peptide 17 was not enhanced when two chains were joined through cystine cross-links (peptide 18, Table 2), nor was $[\theta]_{222}$ increased in other peptides of similar chain length by cystine cross-linking (peptides 14−16, Table 2).

Table 2: Primary Sequences[a] and $-[\theta]_{222}$ ($\times 10^{-3}$ deg cm^2 dmol^{-1}) of Some Synthetic Amphipathic Peptides

	Sequence		T°C	$-[\theta]_{222}$ buffer	50% TFE
1.	AcN<u>L</u>EE<u>L</u>KKK<u>L</u>EE<u>L</u>KGNH$_2$[b]		2	18.4±1.0	37.2±1.9
			25	14.7±0.7	33.3±1.7
2.	AcN<u>L</u>EE<u>L</u>KKK<u>L</u>EE<u>L</u>KQQ<u>L</u>EE<u>L</u>KKK<u>L</u>EE<u>L</u>KQQNH$_2$[b]		2	43.0±2.2	43.5±2.2
			25	40.3±2.0	40.0±2.0
3.	AcK<u>L</u>L<u>K</u>L<u>A</u>AKA<u>L</u>LK<u>L</u>LK<u>L</u>AANH$_2$[c]		5		34.6
4.	AcE<u>I</u>EA<u>L</u>KKE<u>I</u>EA<u>L</u>KKNH$_2$[d]	linear	20	18.6	30.0
		lactam-bridged	20	30.4	32.1
5.	AcE<u>A</u>EA<u>L</u>KKE<u>A</u>EA<u>L</u>KKNH$_2$[d]	linear	20	11.5	28.7
		lactam-bridged	20	29.9	31.5
6.	AcGN<u>A</u>DE<u>L</u>YRML<u>D</u>A<u>L</u>REHL<u>Q</u>S<u>L</u>RRK<u>L</u>RSGNH$_2$[e]		25	22.0	
7.	AcGE<u>L</u>EE<u>L</u>LKK<u>L</u>KE<u>L</u>LKGNH$_2$[f]		25	20.0	
8.	FMOC<u>L</u>KK<u>L</u>LK<u>L</u>LKK<u>L</u>LK<u>L</u>[g]		25	20.0	
9.	AcE<u>A</u>EK<u>A</u>AKE<u>A</u>EK<u>A</u>AKE<u>A</u>EKNH$_2$[h]		20	10.0	
10.	AcE<u>L</u>EK<u>L</u>LKE<u>L</u>EK<u>L</u>LKE<u>L</u>EKNH$_2$[i]		20		26.6
11.	AcGE<u>L</u>KK<u>L</u>LEE<u>L</u>EK<u>L</u>LKGNH$_2$[j]	monomer	22	17.3	
		tetramer	22	26.4	
12.	AcD<u>A</u>QA<u>A</u>AAQA<u>A</u>AAQA<u>A</u>YNH$_2$[k]		0	17.4	
13.	Ac<u>A</u>A<u>K</u>A<u>A</u>AAK<u>A</u>AAK<u>A</u>AA<u>K</u>A<u>A</u>AK<u>A</u>ANH$_2$[l]		0	34.3	
14.	AcY<u>K</u>CK<u>S</u><u>L</u>ESK<u>V</u>K<u>S</u><u>L</u>ESK<u>A</u>K<u>S</u><u>L</u>ESK<u>V</u>K<u>S</u><u>L</u>ESK<u>V</u>K<u>S</u><u>L</u>ESNH$_2$		5	35.8	
	|				
	AcY<u>K</u>CK<u>S</u><u>L</u>ESK<u>V</u>K<u>S</u><u>L</u>ESK<u>A</u>K<u>S</u><u>L</u>ESK<u>V</u>K<u>S</u><u>L</u>ESK<u>V</u>K<u>S</u><u>L</u>ESNH$_2$[m]				
15.	Ac<u>K</u>CAE<u>L</u>EGK<u>L</u>EA<u>L</u>EGK<u>A</u>EA<u>L</u>EGK<u>L</u>EA<u>L</u>EGK<u>L</u>EA<u>L</u>EGNH$_2$		20	30.4	27.1
	|				
	Ac<u>K</u>CAE<u>L</u>EGK<u>L</u>EA<u>L</u>EGK<u>A</u>EA<u>L</u>EGK<u>L</u>EA<u>L</u>EGK<u>L</u>EA<u>L</u>EGNH$_2$[n]				
16.	Ac<u>K</u>CAE<u>L</u>EGK<u>L</u>EA<u>L</u>EGK<u>A</u>EA<u>L</u>EGK<u>L</u>EA<u>L</u>EGK<u>L</u>EA<u>L</u>EGNH$_2$		20	31.8	31.0
	|				
	H$_2$NGE<u>C</u>AE<u>L</u>KGE<u>L</u>AE<u>L</u>KGE<u>L</u>AE<u>A</u>KGE<u>L</u>AE<u>L</u>KGE<u>L</u>EA<u>L</u>KAc[n]				
17.	Ac<u>Q</u>CGA<u>L</u>QKQ<u>V</u>GA<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QKNH$_2$[o]		20	32.9	32.6
18.	Ac<u>Q</u>CGA<u>L</u>QKQ<u>V</u>GA<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QKNH$_2$		20	32.1	30.6
	|				
	Ac<u>Q</u>CGA<u>L</u>QKQ<u>V</u>GA<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QK<u>Q</u>V<u>G</u>A<u>L</u>QKNH$_2$[o]				
19.	Ac<u>K</u>AE<u>I</u>EA<u>L</u>KAE<u>I</u>EA<u>L</u>KAE<u>I</u>EA<u>L</u>KAE<u>I</u>EA<u>C</u>KANH$_2$[p]		20	32.3	32.6

[a] Underlined letters indicate *a* or *d* residues of *abcdefg* heptad repeat. [b] This study. [c] Rothemund et al. (1995). [d] Houston et al. (1996). [e] Betz & DeGrado (1996). [f] Ho & DeGrado (1987). [g] DeGrado & Lear (1985). [h] Zhou et al. (1993). [i] Zhou et al. (1992). [j] Chmielewski & Lipton (1994). [k] Huyghues-Despointes et al. (1993). [l] Chakrabartty et al. (1993). [m] Yu et al. (1996). [n] Monera et al. (1993). [o] Kohn et al. (1995). [p] Su et al. (1994).

The values of $[\theta]_{222}$ of NLEKQ29 and peptides 15−19 (Table 2) were not enhanced in 50% TFE, presumably because the helicity of these peptides was maximized by coiled-coil formation in water. On the other hand, $[\theta]_{222}$ of NLEKG14 was dramatically increased (>100%) by 50% TFE. This is consistent with the lack of coiled-coil formation from NLEKG14. It has been suggested that a minimum of 28 amino acid residues is required for the formation of coiled-coils (Lau et al., 1984).

The $[\theta]_{222}$ of NLEKQ29 both in water and in 50% TFE at 2 °C was $-43\,000 \pm 2200$ deg cm^2 dmol^{-1} (Table 1). Using $[\theta]_{222}$ of NLEKQ29 at 2 °C as the value for $[\theta]_{222}$-(100% α-helix), NLEKG14 in 50% TFE at 25°C ($[\theta]_{222} = -33\,300 \pm 1700$ deg cm^2 dmol^{-1}) was estimated to be 77% α-helix.

NMR of NLEKG14. To obtain an independent estimate of the amount of α-helix present in NLEKG14 in 50% TFE at 25 °C, studies were conducted using NMR which provides more specific structural information than CD. Furthermore, several NMR parameters can be used to estimate the amount of helical structure present.

Figure 3 shows the effect of 50% TFE on the NH region of the 1D spectrum of NLEKG14. In the absence of TFE, only 3 of the 14 backbone amide resonances were resolved. In the presence of 50% TFE, the NH region was well dispersed, resulting in the resolution of 11 of the 14 backbone amide resonances. A well-dispersed NH region is typical of the ^{1}H spectra of predominantly α-helical peptides, presumably due in part to the helix dipole (Wishart et al., 1991). From Figure 3b, it was possible to measure 10 NH−CαH spin−spin coupling constants. Several coupling constants had values <5.0 Hz (see below). The spectra shown

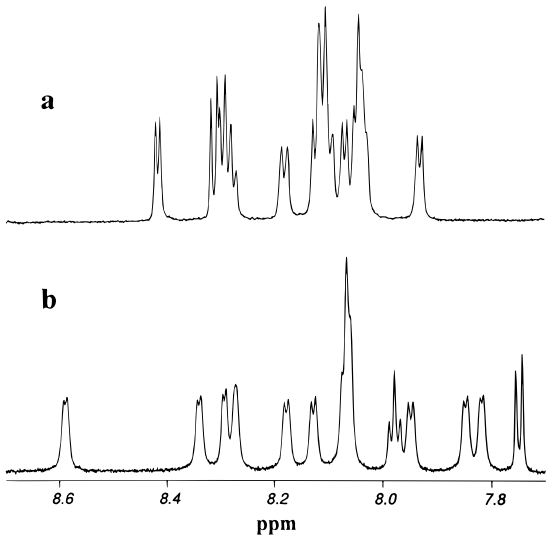

FIGURE 3: Backbone amide regions of the 1D spectra of NLEKG14 in the (a) absence and (b) presence of 50% TFE. Peptide concentration in both samples was 3 mM.

Table 3: ^{1}H Chemical Shifts of NLEKG14 in 50% TFE and 50% Buffer (10 mM Phosphate, 1 mM NaN_3, pH 5.2) at 25 °C

residue	NH	CαH	CβH	CγH	CδH	CεH
N 1	7.74	4.77	3.00, 2.87			
L 2	8.29	4.18	1.73	1.73	1.01, 0.94	
E 3	8.06	4.10	2.22, 2.16	2.54		
E 4	7.94	4.10	2.16	2.56, 2.49		
L 5	7.84	4.14	1.95	1.66	0.94	
K 6	8.27	3.90	1.95	1.42, 1.62	1.71	2.96
K 7	7.81	4.06	1.99	1.63	1.76	3.02
K 8	8.06	4.10	2.09	1.54, 1.74	2.05	3.01
L 9	8.58	4.13	1.99, 1.51	1.89	0.88	
E 10	8.34	4.03	2.36, 2.19	2.74, 2.51		
E 11	8.12	4.13	2.40, 2.24	2.79, 2.54		
L 12	8.17	4.22	1.96, 1.67	1.89	0.94	
K 13	8.06	4.26	1.96	1.60	1.74	3.01
G 14	7.97	3.96				

in Figure 3 suggest that NLEKG14 in 50% TFE is more α-helical than in the absence of 50% TFE, consistent with the interpretation of the CD spectra shown in Figure 1.

All ^{1}H NMR resonances of peptide NLEKG14 in 50% TFE were assigned using TOCSY and DQF-COSY spectra. Sequential assignments were made by locating CαH−NH-(i, i+1) and NH−NH(i, i+1) connectivities in the NOESY spectrum. Table 3 presents ^{1}H chemical shifts of NLEKG14 in 50% TFE.

Residues in α-helices have CαH resonances upfield from their random-coil values (Wishart et al., 1991). Plots of the differences between the chemical shifts of the CαH protons of NLEKG14 in 50% TFE and random-coil values determined by Bundi and Wüthrich (1979) and Merutka et al. (1995) are shown in Figure 4. The chemical shifts of the CαH protons of residues 2 through 13 are shifted upfield from random-coil values, indicating a predominantly α-helical structure.

^{1}H−^{1}H NOE connectivities observed in the NOESY spectrum of NLEKG14 in 50% TFE provided further evidence for the presence of a significant amount of α-helical conformers. Intense short-range NH−NH(i, i+1) NOE connectivities from residues 2 to 14 and medium-range [CαH−NH(i, i+3) and CαH−NH(i, i+4)] NOE connectivi-

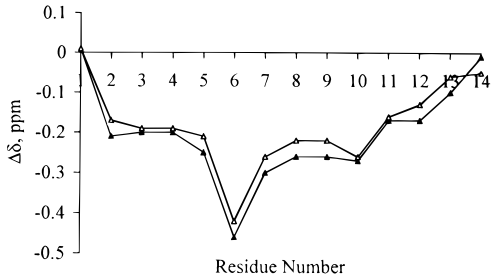

FIGURE 4: Plots of the differences between the CαH chemical shifts of NLEKG14 in 50% TFE and random coil values determined by Bundi and Wüthrich (1979) (filled triangles) and Merutka et al. (1995) (open triangles) versus residue number.

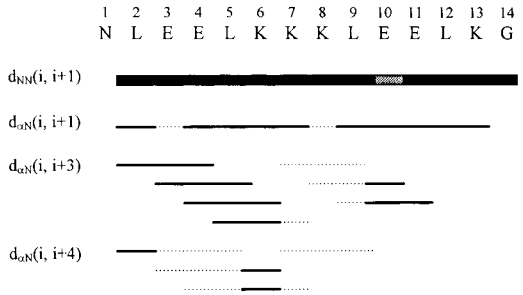

FIGURE 5: Summary of $^1H-^1H$ NOE connectivities for NLEKG14 in 50% TFE. Dotted box and broken lines indicate connectivities which could not be resolved due to overlap.

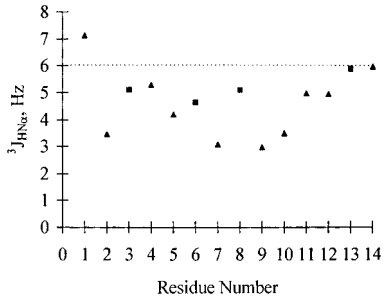

FIGURE 6: Plot of $^3J_{HN\alpha}$ coupling constants of NLEKG14 in 50% TFE versus residue number. Triangles denote values determined from a 1D spectrum. Squares denote values determined from a DQF-COSY spectrum.

ties along the entire peptide backbone were observed (Figure 5).

Another parameter that can be used to estimate secondary structure is the three-bond, spin−spin coupling constant between the NH and CαH protons ($^3J_{HN\alpha}$) (Wüthrich, 1986). Figure 6 shows $^3J_{HN\alpha}$ coupling constants determined for NLEKG14 in 50% TFE. It has been suggested that three or more consecutive residues with $^3J_{HN\alpha} < 6.0$ Hz indicate the presence of an α-helix (Wüthrich, 1986). Based on this criterion, peptide NLEKG14 in 50% TFE is predominantly α-helical from residues 2 through 12.

Quantitation of the amount of helix present in peptides using $^3J_{HN\alpha}$ coupling constants has been suggested by Bradley et al. (1990). This analysis is based on a two-state model; i.e., each amino acid residue is in dynamic equilibrium between an α-helical and a random-coil (rc) state:

$$\%\ \mathrm{H} = \frac{\sum_{i=1}^{n}\left(\frac{J_i - J_{rc}}{J_{hix} - J_{rc}} \times 100\right)}{n} \qquad (3)$$

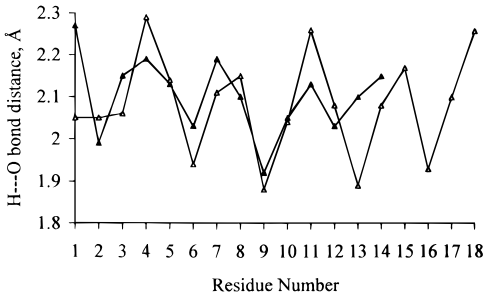

FIGURE 7: Plots of the hydrogen bond lengths (distance from the backbone amide proton to the carbonyl oxygen as calculated from eq 4) versus residue number for NLEKG14 (filled triangles) and LL9 (open triangles) (Zhou et al., 1992).

where J_i is the measured coupling constant for residue i, J_{rc} is the ideal random-coil coupling constant (7.4 Hz), J_{hix} is the ideal α-helix coupling constant (3.9 Hz), and n is the number of residues. Applying this method to the $^3J_{HN\alpha}$ data of NLEKG14 results in a time-averaged helical population of 76%. This is in good agreement with the estimate from CD using $[\theta]_{222}$ of NLEKQ29 at 2 °C as the value for $[\theta]_{222}$ (100% α-helix).

Implications for the Estimation of α-Helix Content in Amphipathic Peptides and Proteins by CD. The high $[\theta]_{222}$ of NLEKG14 and NLEKQ29 obtained at 2 °C indicates higher values for $[\theta]_{222}$ (100% helix) than eq 2 predicts. It has been suggested that the CD of an α-helix is sensitive to solvent-induced distortions (Blundell et al., 1983). Theoretical calculations have shown that hydrophilic α-helices in which the peptide carbonyl groups are tilted outward will have lower mean residue ellipticities (Manning et al., 1988). The high mean residue ellipticities observed for NLEKG14 suggest the formation of a tight amphipathic helix having minimal distortion of the backbone hydrogen bonds (see below). This can also be inferred from the $^3J_{HN\alpha}$ coupling constants of NLEKG14 (Figure 6). Several of the amino acid residues of NLEKG14 have coupling constants between 3 and 3.5 Hz, corresponding to values of the torsion angle ϕ of −49 to −54° (Wüthrich, 1986). These values of ϕ correspond to α-helices in which the carbonyls are internally hydrogen bonded; i.e., they are not hydrogen bonded to side chains or solvent.

Correlation of NH Chemical Shifts of NLEKG14 in 50% TFE with α-Helical Peptide Conformation. In the usual interpretation of amphipathic α-helix conformation, an otherwise linear α-helix is distorted into a supercoil by shortening of specific hydrogen bonds between amide protons and peptide carbonyl groups in *i, i−4* relationships, respectively. This bond-shortening is considered to be produced by shielding of the hydrogen bonds belonging to the *a* and *d* hydrophobic residues (Cohen & Parry, 1990; Stewart, 1993). It would therefore be expected that the shortest hydrogen bonds in NLEKG14 would be those belonging to the leucine residues, but the data expressed in Figure 7 showed that this was not so.

Figure 7 shows a plot of hydrogen bond length (NH···O═C) against residue number for NLEKG14. The hydrogen bond lengths were calculated using the formula (Wagner et al., 1983):

$$\Delta\delta_{NH} = 19.2d^{-3} - 2.3 \qquad (4)$$

where $\Delta\delta_{NH}$ is the difference between the NH chemical shifts

and random-coil values and *d* is the bond distance. Random-coil values determined by Wishart and Sykes (1994) were used. Figure 7 shows that the shortest hydrogen bonds were those in the *a* and *e* positions (residues 2, 6, and 9), which were occupied by leucine, lysine, and leucine, respectively. The same *a* and *e* periodicity in shortness of hydrogen bonds was evident in the data for the leucine zipper peptide (Kuntz et al., 1991) and for peptide LL9 of Zhou et al. (1992) (peptide 10, Table 2, and Figure 7). This phenomenon was unremarkable for peptide LL9 because its *a*, *d*, and *e* positions were all occupied by leucine residues. An explanation for the short amide−hydrogen bond of lysine-6 in NLEKG14 was sought by the construction of CPK space-filling models, which showed that the NH bond of residue 6 was shielded from solvent by three hydrophobic leucine side chains (in positions 2, 5, and 9) as well as by the 4-methylene chain of lysine-6. Likewise, the short hydrogen bond of leucine-2 was shielded by two leucine side chains (positions 2 and 5) and by the methylenes of asparagine-1, while the shortest hydrogen bond in NLEKG14 (leucine-9) was shielded by three leucines (5, 9, 12) and by the lysine-8 methylene chain. It may therefore be deduced that the length of the backbone NH···O═C bond in position *i* is governed by the nature of the four side chains in the surrounding *i*, *i*−1, *i*−4, and *i*+3 positions.

All of the remaining backbone hydrogen bonds in the heptad repeat of NLEKG14 were longer than positions *a* and *e*, but were equal to or slightly longer than the average bond length for the random-coil conformation (Zhou et al., 1992). This is in contrast to the data for peptide LL9 (Figure 7), for which the backbone hydrogen bond of the residue in the *c* position was much longer than for average random-coil values. Inspection of the CPK model indicated that the heptadal *c* position of LL9 was unusually exposed to solvent by having the adjacent side chains held aside through lysine−glutamate salt bridges in *i, i+3* and *i, i+4* relationships. The resulting outward distortion of the peptide carbonyl group belonging to the residue in the *c* position is expected to decrease the CD ellipticity of the peptide (Manning et al., 1988). As also judged from CPK models, side-chain interactions did not preferentially expose backbone hydrogen bonds in any of the positions in the hydrophilic face of the peptide NLEKG14, explaining why those bonds were not lengthened to the same extent as position *c* of peptide LL9, and the overall helicity of NLEKG14 (as measured by CD ellipticity) was less affected by solvent-induced distortion.

Comparison of Synthetic Peptide Properties with IF Proteins. Definition of sequence-conformation correlations among synthetic peptides provides an opportunity to further evaluate the connections between structure and function in filamentous proteins. The wide variety of cytoskeletal IF proteins conform to a common structural pattern consisting of a central rod-shaped section between two regions of random coil. The rod section has been shown to consist predominantly of coiled-coil α-helix, interrupted by short sections of random coil (Quinlan et al., 1994). Additional sections of the rod region have recently been proposed as favoring the *β*-strand conformation, which may form tight *β*-helices that promote dimer formation (Downing 1995, 1996a,b). When these non-α-helical sections are excluded, α-helical sections remain in which the heptadal *a* and *d* positions are heavily hydrophobic and the remaining locations are almost exclusively hydrophilic and often charged.

Table 4: Distribution of Amino Acid Types among the Heptadal Positions of IF Protein α-Helical Rod Regions

	position						
	a	b	c	d	e	f	g
human keratin 1							
hydrophobic	24	1	3	26	2	3	4
hydrophilic	2	24	25	1	26	24	21
alanine	1	2	0	4	0	1	3
human keratin 10							
hydrophobic	22	2	2	25	6	5	2
hydrophilic	5	23	24	2	21	22	24
alanine	1	2	2	1	1	0	1
human vimentin							
hydrophobic	26	4	4	27	3	2	1
hydrophilic	5	22	27	2	25	25	25
alanine	1	6	1	3	2	3	4

Summaries of the distribution of amino acid residues in the α-helical regions of human vimentin and epidermal keratins are given in Table 4. The peptides NLEKG14 and NLEKQ29 that were evaluated in the present study were designed according to these amino acid distributions, and the results may be applied to the understanding of fibrous proteins. From the present results, it is clear that the location of hydrophobic groups in heptadal *a* and *d* positions alone is sufficient to maximize the helicity of the amino-acid chains, especially in longer peptide sequences. Inspection of models of the fibrous proteins (Downing 1995, 1996a, 1996b) shows that the α-helices range from 8 to 61 residues in length, averaging 17, 20, and 24 residues per course for human keratins 1 and 10 and human vimentin, respectively. By comparison with the synthetic peptides NLEKG14 and NLEKQ29, it is clear that the helix lengths in the IF proteins are sufficient to ensure high degrees of helicity. However, the situation may be even more favorable for helix stability in IF proteins in view of the possibility that the courses of random coil within the rod regions may sprout from the heptadal *b*−*c* or *e*−*f* positions without disrupting the continuity of the α-helices (Downing 1996b). In addition, several sequences recognized as conferring a *β* conformation may form tight *β*-helices that could greatly stabilize adjacent α-helices (Downing, 1995, 1996a,b). The overall effect of these structural motifs would be to produce relatively rigid dimeric or tetrameric structures using the entire rod region of each IF protein. Assessment of the proportions of α-helix in IF proteins has been made by previous investigators using CD data, but these results may need to be recalculated on the basis of the high ellipticities of the synthetic peptides obtained in the present study.

ACKNOWLEDGMENT

We thank Dr. Earle Stellwagen for helpful discussions on the CD sections of this paper, Dr. Andrew D. Robertson and his students Gregory DeKoster and William Forsyth for discussions and technical assistance on CD experiments, John Snyder of the University of Iowa High-Field NMR Facility for technical assistance on NMR experiments, and Ty Oturu and Dr. Elena Rus of the University of Iowa Protein Structure Facility for synthesis of NLEKQ29 and quantitative amino acid analyses of CD samples, respectively.

REFERENCES

Betz, S. F., & DeGrado, W. F. (1996) *Biochemistry 35*, 6955−6962.

Blundell, T., Barlow, D., Borkakoti, N., & Thornton, J. (1983) *Nature 306*, 281−283.
Bradley, E. K., Thomason, J. F., Cohen, F. E., Kosen, P. A., & Kuntz, I. D. (1990) *J. Mol. Biol. 215*, 607−622.
Bryson, J. W., Betz, S. F., Lu, H. S., Suich, D. J., Zhou, H. X., O'Neil, K. T., & DeGrado, W. F. (1995) *Science 270*, 935−941.
Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285−298.
Chakrabartty, A., & Baldwin, R. L. (1995) *Adv. Protein Chem. 46*, 141−176.
Chakrabartty, A., Doig, A. J., & Baldwin, R. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11332−11336.
Chen, Y.-H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
Chmielewski, J., & Lipton, M. (1994) *Int. J. Pept. Protein Res. 44*, 152−157.
Cohen, C., & Parry, D. A. D. (1990) *Proteins: Struct., Funct., Genet. 7*, 1−15.
Cooper, T. M., & Woody, R. W. (1990) *Biopolymers 30*, 657−676.
Crick, F. H. C. (1952) *Nature 170*, 882−883.
Crick, F. H. C. (1953) *Acta Crystallogr. 6*, 689−697.
DeGrado, W. F., & Lear, J. D. (1985) *J. Am. Chem. Soc. 107*, 7684−7689.
Doig, A. J., & Baldwin, R. L. (1995) *Protein Sci. 4*, 1325−1336.
Downing, D. T. (1995) *Proteins: Struct., Funct., Genet. 23*, 204−217.
Downing, D. T. (1996a) *Proteins: Struct., Funct., Genet. 25*, 215−224.
Downing, D. T. (1996b) *Proteins: Struct., Funct., Genet. 26*, 472−478.
Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 10*, 4108−4116.
Greenfield, N. J., & Hitchcock-DeGregori, S. E. (1995) *Biochemistry 34*, 16797−16805.
Harbury, P. B., Zhang, T., Kim, P. S., & Alber, T. (1993) *Science 262*, 1401−1407.
Ho, S. P., & DeGrado, W. F. (1987) *J. Am. Chem. Soc. 109*, 6751−6758.
Houston, M. E., Jr., Campbell, A. P., Lix, B., Kay, C. M., Sykes, B. D., & Hodges, R. S. (1996) *Biochemistry 35*, 10041−10050.
Huyghues-Despointes, B. M. P., Scholtz, J. M., & Baldwin, R. L. (1993) *Protein Sci. 2*, 1604−1611.
Kohn, W. D., Monera, O. D., Kay, C. M., Hodges, R. S. (1995) *J. Biol. Chem. 270*, 25495−25506.
Kuntz, I. D., Kosen, P. A., & Craig, E. C. (1991) *J. Am. Chem. Soc. 113*, 1406−1408.
Lau, S. Y. M., Teneja, A. K., & Hodges, R. S. (1984) *J. Biol. Chem. 259*, 13253−13261.
Manning, M. C., Illangasekare, M., & Woody, R. W. (1988) *Biophys. Chem. 31*, 77−86.
Merutka, G., Lipton, W., Shalongo, W., Park, S.-H., & Stellwagen, E. (1990) *Biochemistry 29*, 7511−7515.
Merutka, G., Dyson, H. J., & Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.
Monera, O. D., Zhou, N. E., Kay, C. M., & Hodges, R. S. (1993) *J. Biol. Chem. 268*, 19218−19227.
Nelson, J. W., & Kallenbach, N. R. (1986) *Proteins: Struct., Funct., Genet. 1*, 211−217.
Neuhaus, D., Wagner, G., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257−273.
O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539−544.
Pauling, L., & Corey, R. B. (1953) *Nature 171*, 59−61.
Quinlan, R., Hutchison, C., & Lane, B. (1994) *Protein Profiles 1*, 779−911.
Rothemund, S., Beyermann, M., Krause, E., Krause, G., Bienert, M., Hodges, R. S., Sykes, B. D., & Sönnichsen, F. D. (1995) *Biochemistry 43*, 12954−12962.
Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M., & Baldwin, R. L. (1991) *Biopolymers 31*, 1463−1470.
Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S., & Sykes, B. D. (1992) *Biochemistry 31*, 8790−8798.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
Stewart, J. M. (1993) in *The Amphipathic Helix* (Epand, R. M., Ed.) pp 21−37, CRC Press, Boca Raton, FL.
Su, J. Y., Hodges, R. S., & Kay, C. M. (1994) *Biochemistry 33*, 15501−15510.
Thomas, P. D., & Dill, K. A. (1993) *Protein Sci. 2*, 2050−2065.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948−5949.
Wishart, D. S., & Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
Wishart, D. S., Richards, F. M., & Sykes, B. D. (1991) *J. Mol. Biol. 222*, 311−333.
Woody, R. W., & Dunker, A. D. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.) pp 109−157, Plenum Press, New York.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp 166−169, J. Wiley & Sons, New York.
Yang, J. T., Wu, C.-S. C., & Martinez, H. (1986) *Methods Enzymol. 130*, 208−269.
Yu, Y., Monera, O. D., Hodges, R. S., & Privalov, P. L. (1996) *J. Mol. Biol. 255*, 367−372.
Zhou, N. E., Zhu, B.-Y., Sykes, B. D., & Hodges, R. S. (1992) *J. Am. Chem. Soc. 114*, 4320−4326.
Zhou, N. E., Kay, C. M., Sykes, B. D., & Hodges, R. S. (1993) *Biochemistry 32*, 6190−6197.

BI963061B