Controlling Topology and Native-like Behavior of *de Novo*-Designed Peptides: Design and Characterization of Antiparallel Four-Stranded Coiled Coils

Stephen F. Betz[‡] and William F. DeGrado*[,‡,§]

The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059, and Chemical and Physical Sciences Department, The DuPont Merck Pharmaceutical Company, P.O. Box 80328, Wilmington, Delaware 19880-0328

Received January 16, 1996; Revised Manuscript Received March 27, 1996[⊗]

ABSTRACT: The *de novo* design of peptides and proteins has emerged as an attractive approach for investigating protein structure and function. Here, the design, synthesis, and characterization of a new series of α-helical peptides intended to form antiparallel four-stranded coiled coils is described. Computer models were generated without the use of extant protein structures and were used to refine the sequence. The peptides are of the general formula Ncap-$(X_aZ_bZ_cL_dZ_eZ_fZ_g)_3$-Ccap, where X is either Ala, Val, Thr, or Leu, and Ncap and Ccap are sequences designed to satisfy the helices unpaired amide nitrogens and carbonyl oxygens, respectively. The hydrophobic residues (at positions **a** and **d**) were chosen so that geometric packing of large and small hydrophobes would favor an antiparallel arrangement. Special attention was also given to residues at the helix−helix interfaces. These residues were chosen to balance potential attractive and repulsive electrostatic forces so that the desired topology was favored while other possible folds were destabilized. Two of the four peptides associate under neutral conditions into the desired tetramers. One of the complexes (**a** = Val) behaves like a native-like protein as judged by NMR, thermodynamics, and apolar dye (ANS) binding. The other tetrameric complex (**a** = Leu) exhibits broader NMR resonances, diminished values of ΔH and ΔC_p, and tight binding of the hydrophobic dye ANS, similar to early designed proteins. These results reinforce the importance of optimizing van der Waals packing interactions in protein design but demonstrate that hydrophobic packing must be balanced with hydrogen-bonding and electrostatic interactions to produce novel native-like proteins.

In recent years, considerable progress has been made in understanding the interactions and geometric features that determine the structures of proteins. For example, numerous naturally occurring α-helical bundles have been structurally characterized, leading to the promise that helical proteins with predetermined geometries might be rationally designed from scratch (Mutter & Vuillevmier, 1989; Richardson & Richardson, 1989; Bryson et al., 1995). However, attempts to design such structures have not always met with success. Designed α-helical bundles have often adopted structures with fluctuating conformations, in which the interior hydrophobic side chains are less well-defined than in natural proteins (Betz et al., 1993; Mihara et al., 1993; Sasaki & Lieberman, 1993). Further, in cases where the structures have been crystallographically characterized, the results have not always matched the design. For instance, a peptide designed to assemble into an α-helical dimer has been shown to adopt a trimer in the solid state (Lovejoy et al., 1993), and a fragment of a peptide designed to form a four-helix bundle crystallizes as an α-helical hexamer at low pH (Hill et al., 1990). Furthermore, this same fragment adopts a membrane-like bilayer at neutral pH (Prive et al., 1995) [as does an 18-residue peptide designed to assemble into a hexameric bundle (Taylor et al., 1996)], while a detergent-like peptide intended to solubilize membrane proteins forms four-helix bundles in the solid state (Schafmeister et al., 1993). Finally, subtle changes to the hydrophobic residues of the two-stranded coiled coil from GCN4 cause this structure to adopt two-, three-, or four-stranded structures (Harbury et al., 1993, 1994).

However, while the results of designs have not always matched the expectations of their designers, they have nevertheless provided a deeper understanding of the principles stabilizing proteins (Betz et al., 1995b). Iterative design of four-helix bundles from minimal models and detailed solution measurements of these proteins have helped determine the features required to define a unique conformation in aqueous solution (Raleigh & DeGrado, 1992; Handel et al., 1993; Olofsson et al., 1995; Raleigh et al., 1995; Struthers et al., 1996). Also, the structures of designed helical bundles have provided a growing data base to mine new insights into the principles of helix−helix association (Hill et al., 1990; Lovejoy et al., 1993; Schafmeister et al., 1993; Harbury et al., 1993, 1994; Taylor et al., 1996). In the present paper, we analyze the structures of parallel and antiparallel two-, three-, and four-helix bundles and extract a comprehensive set of principles for the *de novo* construction of these folds that includes the use of hydrophobic and electrostatic components to define a particular topology. We then put these principles to the test by designing a series of novel antiparallel four-stranded coiled coils. The solution properties of these peptides are critically examined and compared to those of other earlier designs of coiled coils and four-helix bundles.

* To whom correspondence should be addressed. Please address correspondence to the University of Pennsylvania address. Fax: 215-898-4217. E-mail: wdegrado@mail.med.upenn.edu.

[‡] The DuPont Merck Pharmaceutical Co.

[§] University of Pennsylvania.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1996.

S0006-2960(96)00095-5 CCC: $12.00

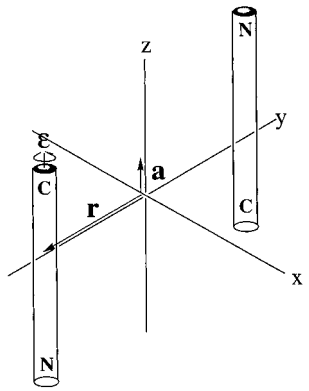

FIGURE 1: Schematic representation of the parameters used to mathematically generate α-helical bundles.

MATERIALS AND METHODS

Mathematical Generation of α*-Helical Bundles*. Helical bundles were created by applying symmetry operators to polyalanine helices using well-established approaches (Phillips, 1992; Seo & Cohen, 1993). For the sake of simplicity, we will illustrate our general approach using an antiparallel two-helix bundle as an example. A polyalanine helix (created using $\phi = -57°$ and $\psi = -47°$) is converted from an α-helix with 3.6 residues per turn to one with 3.5 residues per turn by aligning C_α carbons of the helix along the *z*-axis of the Cartesian coordinate system and applying the matrix

$$\begin{bmatrix} x\text{coil}(i) \\ y\text{coil}(i) \\ z\text{coil}(i) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(i) \\ y(i) \\ z(i) \end{bmatrix}$$

$$\text{where} \qquad \theta = -2\pi[z_{(i)} - z_{\min}]/p$$

where $x(i)$, $y(i)$, and $z(i)$ are the original Cartesian coordinates of the α-helix and $x\text{coil}(i)$, $y\text{coil}(i)$, and $z\text{coil}(i)$ are the new coordinates for the twisted helix. $z_{\min}$ is the minimum of the *z*-coordinates for all the C_α carbons, $z_{(i)}$ is the *z*-coordinate of the C_α carbon of residue *i*, and *p* is the pitch of the coiled coil, in this case 189 Å. This transformation positions each **a** and **d** residue [**a**−**g** refer to traditional coiled coil heptad nomenclature (Cohen & Parry, 1990)] such that they line up along the helical axis rather than spiraling around the helix as in a canonical α-helix. The resulting helix is then built into a two-helix bundle using the program PSSHOW by varying the following parameters: ϵ, the rotation of the helix about its own axis; *r*, the displacement along the *y*-axis; and *a*, the displacement of the helix along the *z*-axis of the coordinate system (Figure 1). The symmetry-related helix is generated by rotating the first helix about the *x*-axis, creating the two-helix bundle. Other helical bundles, including antiparallel four-helix bundles, are generated using appropriate symmetry operators.

The values of ϵ, *r*, and *a* are chosen to position the **a** and **d** residues toward the center of the bundle such that they lie in layers. The helical bundle is then given a left-handed twist using the above transformation matrix, but with the signs of the elements reversed to result in rotation in the opposite direction. The above transformations do not preserve bond lengths and angles, but because the difference in the *z*-value is small for atoms that are bonded to each other, internal geometries are readily restored upon energy minimization. Initially, hydrophobic core residues (i.e. the **a** and **d** positions) are created with the side chains in low energy rotamer conformations (McGregor et al., 1987; Schrauber et al., 1993; Dunback & Karplus, 1994), and the packing is qualitatively assessed. After a desired interhelical geometry is decided upon, the remaining residues of the polyalanine helix are then converted to the appropriate sequence, again with side chain rotamers in low-energy conformations (McGregor et al., 1987; Schrauber et al., 1993; Dunback & Karplus, 1994). The structure is then subjected to energy minimization using the default force field in Discover (Biosym), typically employing 1000 cycles of steepest descents and with the position of the C_α carbon atoms fixed and a cutoff distance of 10 Å for electrostatic interactions. The C_α carbon atom restraints are removed, and another 1000 cycles of steepest descents followed by 3000 cycles of conjugate gradients are performed. The resulting structures are then examined for potentially poor contacts, the appearance of which leads to sequence refinement, followed by further rounds of modeling and energy minimization.

Peptide Synthesis and Purification. Peptides were synthesized by standard automated methods on a Milligen 9050 peptide synthesizer using FMOC-protected amino acids and purified by reverse phase HPLC (Choma et al., 1994). The peptides were determined to be homogeneous by analytical HPLC and electrospray ionization mass spectrometry.

Analytical Ultracentrifugation. Sedimentation equilibrium analysis was performed with a Beckman XLA analytical ultracentrifuge (Harding et al., 1992). Initial peptide concentrations ranged from 5 to 800 μM in 20 mM MOPS at pH 6.9. The samples were centrifuged at 35 000, 42 000, and 48 000 rpm. Equilibrium was determined when successive radial scans at the same speed were indistinguishable.

The behavior of a single species at equilibrium can be described by the following equation:

$$M_b = M_w(1 - \bar{v}\rho)$$

where M_b is the measured buoyant molecular weight, M_w is the molecular weight in daltons, $\bar{v}$ is the partial specific volume of the solute, and ρ is the density of the sample solution. The partial specific volume was calculated from the weighted average of the amino acid content using the method of Cohn and Edsall (1943).

The aggregation state was determined by fitting the data to cooperative monomer−*n*mer equilibria, as well as to a single species using the program Igor Pro (WaveMetrics). Radial absorbance spectra were recorded at several appropriate wavelengths and were shown to yield indistinguishable results.

NMR Spectroscopy. ^{1}H NMR spectra were recorded with a Bruker AMX-600 spectrometer at 298 K using presaturation of the water resonance. One-dimensional spectra were recorded with 8K points. The raw data were transformed and phased using Felix 1.1 software (Hare/Biosym).

Fluorescence Measurements. Fluorescence spectra of ANS−peptide complexes were measured using a Spex Fluorolog fluorimeter at 298 K. ANS−peptide complexes were monitored by titrating peptide into a constant amount (10 μM) of ANS in a 1 cm cuvette. Buffer conditions were 20 mM MOPS at pH 6.9. The excitation wavelength was 370 nm, with emission spectra recorded from 420 to 500

Table 1: Structures of Parallel and Antiparallel α-Helical Bundles

structure	topology	number of helices	reference
GCN4	parallel	two	O'Shea et al., 1991
seryl-tRNA synthetase	antiparallel	two	Cusack et al., 1990
GCN4 variant	parallel	three	Harbury et al., 1994
coil-Ser	antiparallel	three	Lovejoy et al., 1993
GCN4 variant	parallel	four	Harbury et al., 1993
Rop	antiparallel	four	Banner et al., 1987
peptitergent	antiparallel	four	Schafmeister et al., 1993

nm. Binding curves were derived by plotting peptide concentration versus emission at 460 nm.

Circular Dichroism. The circular dichroism of peptide solutions was monitored with an AVIV 62DS spectropolarimeter. Buffer conditions were 20 mM MOPS at pH 6.9. The peptide concentration ranged from 5 μM to 3 mM. The concentrations of peptide stock solutions were determined by the absorbance at 275 nm in 6 M guanidinium chloride, using an extinction coefficient of 1420 cm^{-1} M^{-1} (there is one tyrosine per peptide) (Gill & von Hippel, 1989). Molar ellipticities were calculated using the equation

$$[\theta] = \theta_{obsd}/10lc$$

where θ_{obsd} is the ellipticity measured in millidegrees, l is the length of the cell in centimeters, and c is the peptide concentration in moles per liter.

Mean residue ellipticity was fit as a function of peptide concentration (DeGrado & Lear, 1985) using the program Igor Pro. The equilibrium constant K_{diss} is defined as $[M]^n/[T]$, in which [T] is the concentration of the associated species, [M] is the concentration of monomer, and n is the association state. Thus, $K_{diss} = n(\alpha^n)(C^{n-1})/(1 - \alpha)$, where α is the fraction of monomer $[(\theta_{obsd} - \theta_{assoc})/(\theta_{mon} - \theta_{assoc})]$ and C is the total peptide concentration. The values of θ_{assoc} and θ_{mon} are fit, and the value of n is determined experimentally from a high-concentration sedimentation equilibrium.

RESULTS

Analysis of α-Helical Bundles and Coiled Coils. The successful design of a protein requires not only a sequence that is consistent with the desired topology but also one that destabilizes alternate topologies. Otherwise, structures with fluctuating conformations are observed (Betz et al., 1993). Therefore, it is important to understand the scope of and the features stabilizing probable alternative conformations for a given desired fold. To accomplish this, we began by examining a series of parallel and antiparallel helical bundles consisting of one to four helices (Table 1).

The geometries of parallel α-helical bundles are shown in Figure 2A. The helices are represented as having an exact seven-residue repeat to simulate the effect of the tilting or supercoiling of helical bundles, and the positions in the repeat are labeled **a**−**g**, following traditional coiled coil nomenclature (Cohen & Parry, 1990). Two-, three-, and four-stranded parallel coiled coils are fundamentally similar; each structure is stabilized by burying apolar residues at the **a** and **d** positions. In each heptad, all the **a** residues segregate into a single layer, and the **d** positions segregate into a second, nonequivalent layer. Although the **a** and **d** positions

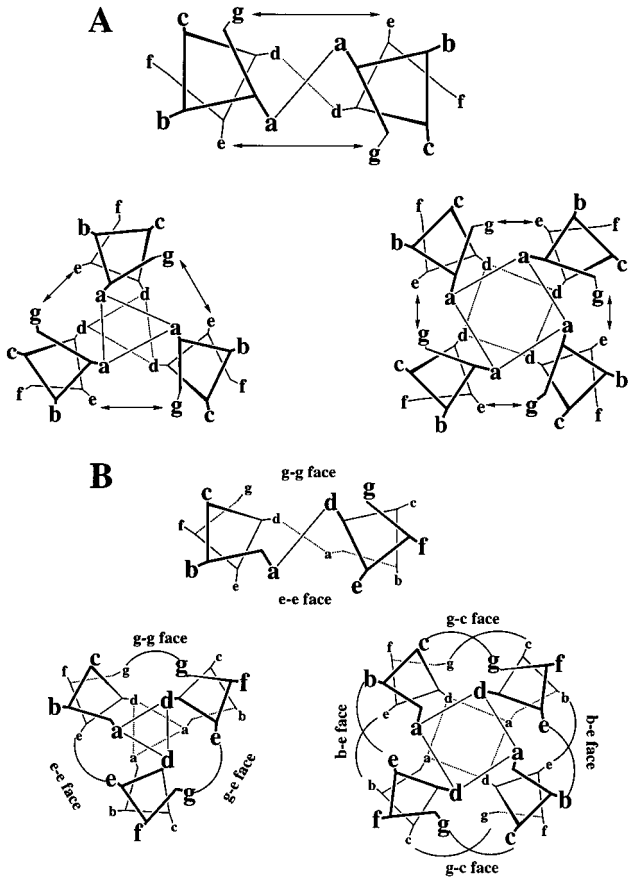

FIGURE 2: (A) Heptad arrangements in parallel α-helical bundles as observed in wild-type GCN4 (top) (O'Shea et al., 1991), pII-GCN4 (lower left) (Harbury et al., 1994), and pLI-GCN4 (lower right) (Harbury et al., 1993). The arrows indicate e−g interfacial interactions. (B) Heptad arrangements in antiparallel α-helical bundles as observed in seryl-tRNA synthetase (top) (Cusack et al., 1990), coil-Ser (lower left) (Lovejoy et al., 1993), and Rop (lower right) (Banner et al., 1987). Note the differences in the interhelical interfaces in each packing.

appear similar, they in fact point into the helical interface at different angles, leading to different sequence preferences for different aggregation states (Harbury et al., 1993, 1994; Betz et al., 1995b). The residues at the **e** and **g** positions help sequester the internal hydrophobic residues from solvent. They are often held in place via salt bridges between a Lys or Arg at **e** and an acidic group at a preceding **g** of the neighboring helix. In the tetramer, in addition to **e** and **g** pairing, **b** and **c** positions can potentially interact.

Antiparallel helices appear deceptively similar to parallel coiled coils, but their interaction patterns are very different. Figure 2B illustrates packing diagrams taken from the antiparallel dimer in seryl-tRNA synthetase (Cusack et al., 1990), the antiparallel trimer observed in the synthetic, "coil-Ser" peptide (Lovejoy et al., 1993), and the antiparallel tetramer from Rop (Banner et al., 1987). In the dimer and the tetramer, the side chains pack into pseudo-equivalent layers consisting of equal numbers of **a** and **d** residues. By contrast, the trimer shows alternating, nonequivalent layers consisting of a layer with one **a** and two **d** residues, followed by a layer with one **d** and two **a** residues. Unlike the parallel structures, in which the interactions of the interfacial residues were similar for each aggregation state, the antiparallel dimer, trimer, and tetramer show different patterns. In the antiparallel dimer, the interfacial **e** and **g** positions segregate on

Table 2: List of Interactions in Parallel and Antiparallel α-Helical Coiled Coils

topology	stabilizing	destabilizing
parallel two-stranded	none	two Arg-Lys two Arg-His segregation of large and small layers in the hydrophobic core[a] nearly full exposure of Met and Tyr to solvent
anti-parallel two-stranded	geometric packing within the hydrophobic core[a]	two His-Lys two Arg-Arg nearly full exposure of Met and Tyr to solvent
parallel three-stranded	none	three Arg-Lys three Arg-His segregation of large and small layers in the hydrophobic core[a]
antiparallel three-stranded	geometric packing within the hydrophobic core[a]	two His-Lys two Arg-Arg one Arg-His one Arg-Lys
parallel four-stranded	none	four Arg-Lys four Arg-His segregation of large and small layers in the hydrophobic core[a]
antiparallel four-stranded	four Asp-Arg four Gln-Arg four Arg-Tyr geometric packing within the hydrophobic core[a]	none

[a] Does not apply to coil-LL.

opposite sides of the structure (labeled the **e**−**e** and **g**−**g** faces in Figure 2B). The antiparallel trimer has one parallel interface, virtually identical to the interface in the parallel three-stranded coiled coil, and two antiparallel helix−helix interfaces. Thus, it has one **e**−**e** face, one **g**−**g** face, and one **g**−**e** face (Figure 2B). Finally, in the tetramer, the **b** and **e** positions interact; correspondingly, the **g** and **c** positions also interact. Thus, the tetramer shows two **g**−**c** faces and two **b**−**e** faces.

The differences in packing interactions between antiparallel dimers, trimers, and tetramers are fortunate for the protein designer because they provide many opportunities for engineering specificity for a given association state. This contrasts with the parallel coiled coils, where changes as seemingly minor as an Ile for Leu substitution at the **a** or **d** position can cause a change in the association state (Harbury et al., 1993, 1994; Betz et al., 1995b). This sensitivity may reflect the interfacial equivalence of the parallel coiled coils (see Figure 2A). Thus, the above helical wheels are invaluable road maps in the early stages of sequence analysis for a designed protein. However, it is essential to build full atomic models to assess the stereochemistry of the proposed interactions as described in Materials and Methods.

Design of an Antiparallel Four-Stranded Coiled Coil. In keeping with our *de novo* approach (DeGrado, 1988; DeGrado et al., 1989), the design of antiparallel coiled coils did not begin with either the sequence or the atomic coordinates of naturally occurring proteins, as in protein-engineering work from other laboratories (Munson et al., 1994). Instead, we began with an isolated polyalanine helix and mathematically generated a family of bundles with varying interhelical spacings, rotations of the helices, and interhelical offsets as described in Materials and Methods. The bundle was given a slight left-handed twist to simulate the curving of the helices observed in this and related long coiled coils (Barlow & Thornton, 1988; Cohen & Parry, 1990). For consistency with earlier designs, and in keeping with a minimalist philosophy, we began with Leu at each **a** and **d** position, and these were evaluated in each of their low-energy conformations (McGregor et al., 1987; Schrauber et al., 1993; Dunback & Karplus, 1994). After evaluation of the models, optimal packing was obtained with an interhelical spacing of 10.5 Å (between adjacent helices, i.e. 14.8 Å diagonally across the bundle). Side chains were next placed at the **c**−**g** and **b**−**e** interfaces. The identities of these residues were chosen to (i) shield the **a** and **d** positions from solvent, (ii) destabilize parallel bundles via unfavorable electrostatic interactions, and (iii) form energetically reasonable polar interactions. In particular, several Arg−Tyr and Arg−Gln hydrogen bonds and Arg−Asp salt bridges were introduced which should form only if the helices associate into an antiparallel four-helix bundle. A list of interhelical interactions for the desired as well as the undesired topologies is given in Table 2.

Efficient van der Waals packing of apolar residues within the core of a protein has long been recognized as a feature necessary for adopting a native structure (Ponder & Richards, 1987), and recent work with designed proteins has reinforced this concept (Raleigh & DeGrado, 1992; Munson et al., 1994; Raleigh et al., 1995; Desjarlais & Handel, 1995). We therefore explored the effects of geometric packing by modeling structures in which the **a** positions were small, aliphatic residues (Ala or Val) similar to Rop, while keeping the Leu residues constant at **d**. Our modeling suggested that coil-VL and -LL would pack well as tetramers. [The two-letter nomenclature refers to the residues at the **a** and **d** positions, respectively.] Similarly, coil-TL could pack well, although this requires burying the hydroxyl groups within

Table 3: Thermodynamic Parameters of Coil-VL and Coil-LL

	method	T (K)[a]	ΔG (kcal/mol)	$\Delta\Delta G$ per L/V (kcal/mol)	ΔH_m (kcal/mol)	$\Delta\Delta H_m$ per L/V (kcal/mol)	$T\Delta S_m$ (kcal/mol)	$\Delta(T\Delta S_m)$ per L/V (kcal/mol)
coil-LL	sedimentation equilibrium	298	−22					
coil-VL	sedimentation equilibrium	298	−18	0.33				
coil-LL	thermal	298	−20.6					
coil-VL	thermal	298	−16.8	0.32				
coil-LL	thermal	330	−20.1		−33.2[b]		13.1	
coil-VL	thermal	330	−14.1	0.50	−60.6[c]	−2.3	47.5	2.9

[a] Temperature at which thermodynamic parameters were evaluated. [b] Evaluated at 700 μM. [c] Evaluated at 40 μM.

heptad position	a	b	c	d	e	f	g	
G N	A	D	E	L	Y	R	M	
	X	D	A	L	R	E	H	
	X	Q	S	L	R	R	K	
	X	R	S	G				
	X =	Val		Coil-VL				
		Ala		Coil-AL				
		Thr		Coil-TL				
		Leu		Coil-LL				

FIGURE 3: Sequence of the coil-XL series of peptides. The sequence is arranged by heptads, with the altered **a** position indicated by an *X*. The N-terminus is acetylated, and the C-terminus is amidated.

the protein interior. Nevertheless, hydrogen-bonded interactions with a carbonyl at the $i - 4$ position (Baker & Hubbard, 1984) can partially compensate for the unfavorable free energy of dehydration of the hydroxyl group. Finally, with Ala at the **a** position (coil-AL), optimal packing was only achieved with more closely packed helices (9.4 Å between the centers of adjacent helices). Additionally, Ncap and Ccap groups were added (Richardson & Richardson, 1988; Seale et al., 1994), providing the final target sequences (Figure 3).

Sedimentation Equilibrium. The association of the peptides was examined using sedimentation equilibrium centrifugation with initial loading concentrations ranging from 5 to 800 μM. At all concentrations, coil-AL and coil-TL were predominantly monomeric. In contrast, coil-VL and coil-LL cooperatively assemble into tetramers in a concentration dependent manner. These peptides sediment as single homogeneous species with molecular weights consistent with tetramer formation at the loading concentrations of 400 μM or greater (Figure 4). At lower loading concentrations, the sedimentation curves are described well by a cooperative monomer−tetramer equilibrium with an overall K_{diss} of $\sim 10^{-13}$ M^3 for coil-VL, corresponding to a free energy of assembly of −18 kcal mol^{-1} (1 M standard state), and a K_{diss} of $\sim 10^{-16}$ M^3 for coil-LL ($\Delta G = -22$ kcal mol^{-1}).

During the purification of a derivative of coil-LL, a side product was isolated in which Met-9 was deleted. This peptide, ΔM-coil-LL, was analyzed by sedimentation equilibrium and shown to be monomeric at concentrations up to 2 mM (data not shown). Because the deletion of Met-9 disrupts the hydrophobic repeat of the peptide, the loss of the ability to associate qualitatively supports the packing within the model. Additionally, the presence of Met-9 in the first heptad also supports the model being antiparallel because, in parallel models, the two remaining heptads with the same hydrophobic repeat could drive association (especially at high concentrations), but none is seen.

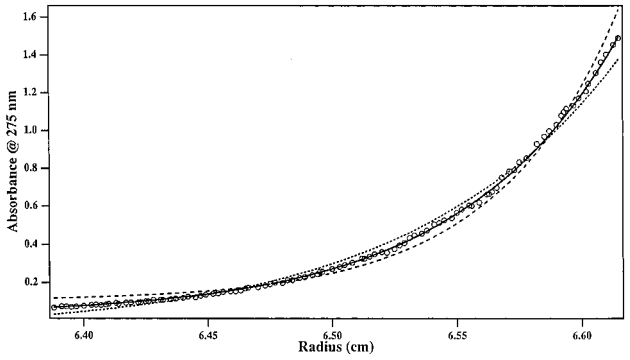

FIGURE 4: Sedimentation equilibrium. Representative analytical ultracentrifugation run for 400 μM coil-LL in 20 mM MOPS at pH 6.9 and 298 K showing absorbance at 275 nm versus radius for raw data (circles) and different theoretical monomer−*n*mer fits: solid line, fit to tetramer; dashed line, fit to pentamer; and dotted line, fit to trimer.

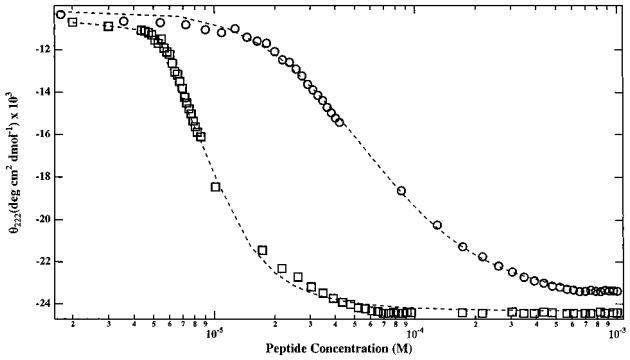

FIGURE 5: Concentration dependent circular dichroism. Mean residue ellipticity at 222 nm as a function of peptide concentration for coil-VL (circles) and coil-LL (squares) in 20 mM MOPS at pH 6.9 and 298 K. The dashed lines show the theoretical fits to tetramer. Data for each complex were acquired in two experiments using a dual-syringe titrator (AVIV Associates) and were collected to emphasize the pre- and post-transitions where the cooperativity of the association would be most easily measured.

Far-UV CD Spectroscopy. The far-UV CD spectra of coil-AL and coil-TL indicated that these monomeric peptides were partially α-helical (data not shown). The coil-AL peptide was more helical (45%) than coil-TL (<30%), consistent with the higher helical potential of Ala relative to that of Thr (O'Neil & DeGrado, 1990; Muñoz & Serrano, 1994; Bryson et al., 1995). In contrast, coil-VL and coil-LL showed concentration dependent CD spectra (Figure 5), consistent with the monomer−tetramer equilibria determined by sedimentation equilibrium. At high concentrations, the spectra are consistent with nearly complete helix formation, whereas at lower concentrations, the spectra are consistent with a mixture of α-helix and random coil. Analysis of the concentration dependence of the spectra (DeGrado & Lear,

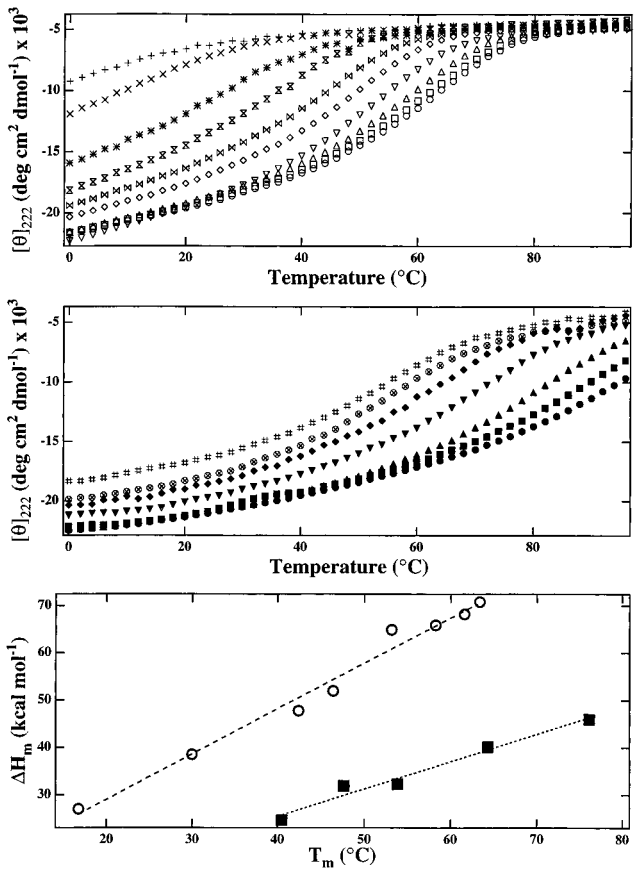

FIGURE 6: (Top panel) Mean residue ellipticity at 222 nm versus temperature for coil-VL in 20 mM MOPS at pH 6.9 as a function of peptide concentration. Sample concentrations were as follows: 1.25 mM (circles), 1.0 mM (squares), 750 μM (triangles), 400 μM (inverted triangles), 300 μM (diamonds), 200 μM (bow ties), 100 μM (hourglasses), 50 μM (*), 20 μM ($\times$), and 10 μM (+). (Middle panel) Mean residue ellipticity at 222 nm versus temperature for coil-LL in 20 mM MOPS at pH 6.9 as a function of concentration. Sample concentrations were as follows: 1.0 mM (solid circles), 500 μM (solid squares), 200 μM (solid triangles), 75 μM (solid inverted triangles), 50 μM (solid diamonds), 35 μM ($\otimes$), and 20 μM (#). (Lower panel) The van't Hoff enthalpy (ΔH_m) versus T_m for coil-VL (open circles) and coil-LL (filled squares). The slopes of the lines are equal to ΔC_p: for coil-VL, 950 cal mol^{-1} K^{-1}; and for coil-LL, 575 cal mol^{-1} K^{-1}.

1985; Betz et al., 1995a) gives values of K_{diss} of 9.9×10^{-14} and 3.4×10^{-16} M^3 for coil-VL and coil-LL, respectively. These values are in good agreement with the more approximate values estimated by sedimentation equilibrium centrifugation.

Thermal Denaturation. The far-UV CD spectra of coil-VL and coil-LL were measured as a function of concentration and temperature (Figure 6). Coil-VL exhibited thermal-unfolding curves similar to those of native proteins. As expected for a monomer−tetramer equilibrium, the midpoint shifts to higher temperatures as the concentration increases. van't Hoff analysis of the data gives a ΔH_m of 62 kcal (mol of tetramer)$^{-1}$ (measured at 400 μM peptide concentration) and an overall ΔC_p of 8.8 cal mol^{-1} K^{-1} res^{-1}, values which are similar to those of native proteins this size (Pace et al., 1989). By contrast, coil-LL yields a broad thermal-unfolding curve at high peptide concentrations. The transition remained quite broad at the midpoint of its concentration dependent association curve (Figure 5), where temperature dependent changes in its equilibrium are readily measured. Analysis yields a ΔH_m of 44 kcal (mol of tetramer)$^{-1}$ (measured at

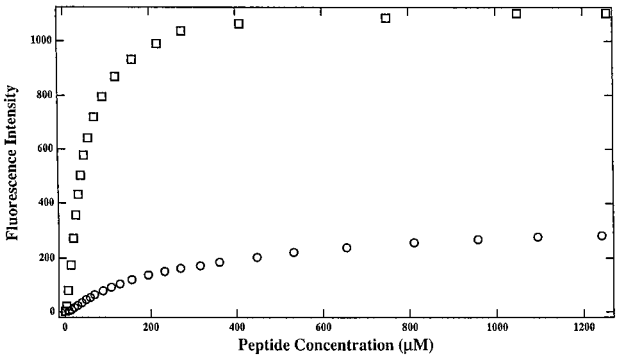

FIGURE 7: Emission fluorescence at 460 nm versus peptide concentration in the presence of 5 μM 8-anilino-1-naphthalenesulfonic acid (ANS) for coil-VL (circles) and coil-LL (squares). The excitation wavelength is 370 nm.

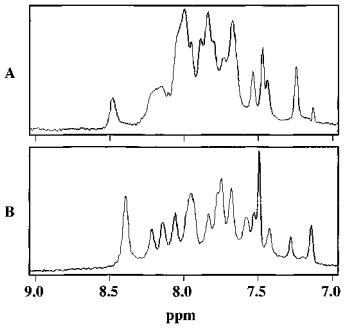

FIGURE 8: Downfield region of the ^{1}H NMR spectrum of coil-LL (A) and coil-VL (B) at 298 K in 10 mM potassium phosphate and 10% D_2O at pH 7.0. Spectra were recorded on a Bruker AMX-600 MHz spectrometer with weak irradiation to saturate residual HOD. Chemical shifts are given in parts per million from DSS.

400 μM peptide concentration) and an overall ΔC_p of 5.3 cal mol^{-1} K^{-1} res^{-1}. These values of ΔH_m and ΔC_p are small compared to naturally occurring proteins of this size. This behavior is reminiscent of the diminished values of ΔH_m and ΔC_p observed for the folding of molten globules (Haynie & Friere, 1993; Hagihara et al., 1994).

Binding of 8-Anilinonaphthalenesulfonic Acid (ANS). Molten globules generally bind the fluorescent dye ANS, with dissociation constants in the micromolar range, and exhibit a large increase in the quantum yield and a blue shift in the emission maximum in the bound state (Semisotnov et al., 1991; Ptitsyn, 1992). Figure 7 illustrates the fluorescence emission at 460 nm of 10 μM ANS in the presence of increasing concentrations of coil-VL and coil-LL. Both show saturable binding isotherms, although the association was tighter for coil-LL (midpoint = 40 μM) than for coil-VL (midpoint = 220 μM). At saturating coil-LL concentrations, the emission maximum for ANS shifts to 464 nm and a large increase in intensity is observed, typical of the behavior observed for molten globules. In contrast, at saturating concentrations of coil-VL, the emission maximum only shifts to 475 nm and the increase in intensity is considerably smaller, consistent with ANS being in a more polar environment. In summary, coil-LL binds ANS in a manner reminiscent of the molten globule, whereas coil-VL binds it very weakly, similar to the nonspecific association observed for some native proteins (Semisotnov et al., 1991; Ptitsyn, 1992).

NMR Spectroscopy. Figure 8 shows the amide region of the ^{1}H NMR spectra of coil-VL and coil-LL (peptide concentration of ~1 mM). The resonances of the coil-LL

are broad and poorly defined and suggestive of internal motions on the millisecond time scale. The resonances of coil-VL by comparison are better resolved. Both peptides show good dispersion of chemical shifts for an α-helical structure with only one aromatic residue (Tyr-7).

DISCUSSION

The design of the coil-XL family of proteins represents an evolution from traditional, minimalist protein design (DeGrado et al., 1989) and differs from the earlier α_1 peptides in three fundamental ways, each of which should favor a more uniquely folded structure. First, the helices are longer; helices that are the length of those in the α_1 series (12−16 residues) can pack in a variety of geometries (Murzin & Finkelstein, 1988), but lengthier helices can pack optimally only in elongated bundles. Second, the α_1 series of peptides contains three apolar residues per heptad, two of which are buried in the apolar interior, while the third lies at the interhelical interface. A potential problem with that design is that a sequence with three Leu residues per heptad can adopt one of two energetically similar packings in which the interfacial Leu residues can lie at either the **b**−**e** or the **c**−**g** interface (Raleigh et al., 1995). In the coil-XL design, we used more specific hydrogen-bonded and electrostatic interactions at the interfacial positions to stabilize the structure while also favoring antiparallel versus parallel packing. Finally, we systematically varied the Leu at the **a** position which has several low-energy rotamers in an α-helical conformation (McGregor et al., 1987; Schrauber et al., 1993; Dunback & Karplus, 1994) to Val, Thr, or Ala which all have just one predominant rotomeric state in this conformation. We felt that, if the conformations of the **a** positions were fixed, then the more flexible Leu residues at the **d** positions would adopt single conformations that favor optimal packing within the core. Indeed, in all respects examined, the coil-VL tetramer behaved more like a native protein than coil-LL.

It is interesting to note that the free energies of tetramerization of coil-VL and coil-LL are similar. The free energy difference expressed per Leu/Val replacement (12 sites per tetramer) is approximately 0.3 kcal mol^{-1} at 300 K, which is close to the difference in helix propensity between the two amino acids [see review in Bryson et al. (1995)]. However, the driving forces for tetramerization as inferred from the values of ΔH_m, ΔC_p, and ΔS_m for the transitions are different. The folding and tetramerization of coil-VL is a more enthalpically driven process ($\Delta\Delta H_m$ is −2.3 kcal mol^{-1} per Leu/Val mutation), consistent with the hypothesis that this structure has better van der Waals packing than coil-LL. In contrast, coil-LL shows a greater degree of entropic stabilization; $T\Delta S_m$ is 2.9 kcal mol^{-1} per Leu/Val mutation. Further, the lower value of ΔC_p observed for coil-LL would suggest that its apolar side chains are not as efficiently desolvated in the folded state, again consistent with a more loosely folded structure that might permit the percolation of water molecules into the core.

It is also interesting to note that coil-AL and coil-TL failed to associate even at concentrations up to 1 mM. It is likely that the dehydration of the smaller methyl side chains of Ala did not provide a sufficient driving force to cause folding. Similarly, the dehydration of three Thr side chains per monomer is likely too energetically costly to allow folding and association.

Early designed proteins incrementally tested fundamental principles of protein structure and stability (hydrophobic balance and periodicity, interior packing, and specificity elements) and have yielded important insights about what is required to design a native-like protein. Here, rather than proceed through similar small steps, the design of the coil-XL peptides sought to incorporate each of these concepts into one molecule. This hierarchic approach emphasizes incorporation of geometric packing within the hydrophobic core with the use of interfacial interactions to produce a stably folded form of the desired topology while simultaneously destabilizing alternate undesired folds. The success of these designs suggests that the protein-folding code is understood enough to allow design of stable, native-like molecules from scratch.

ACKNOWLEDGMENT

We thank Dan Raleigh, Jim Bryson, and Blake Hill for helpful discussions. We also thank Gregg Dieckmann for initial work on computer modeling of coiled coils. More detailed procedures for modeling multistranded coiled coils can be found in his Ph.D. Dissertation (1995), University of Michigan, Ann Arbor, MI.

REFERENCES

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.

Banner, D. W., Kokkinidis, M., & Tsernoglou, D. (1987) *J. Mol. Biol. 196*, 657−675.

Barlow, D. J., & Thornton, J. M. (1988) *J. Mol. Biol. 201*, 601−619.

Betz, S. F., Raleigh, D. P., & DeGrado, W. F. (1993) *Curr. Opin. Struct. Biol. 3*, 601−610.

Betz, S., Fairman, R., O'Neil, K., Lear, J., & DeGrado, W. (1995a) *Philos. Trans. R. Soc. London, Ser. B 348*, 81−88.

Betz, S. F., Bryson, J. W., & DeGrado, W. F. (1995b) *Curr. Opin. Struct. Biol. 5*, 457−463.

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.

Choma, C. T., Lear, J. D., Nelson, M. J., Dutton, P. L., Robertson, D. E., & DeGrado, W. F. (1994) *J. Am. Chem. Soc. 116*, 856−865.

Cohen, C., & Parry, D. A. D. (1990) *Proteins: Struct., Funct., Genet. 7*, 1−15.

Cohn, E. J., & Edsall, J. T. (1943) *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, pp 370−377, Reinhold Publishing Corp., New York.

Cusack, S., Berthetcolominas, C., Hartlein, M., Nassar, N., & Leberman, R. (1990) *Nature 347*, 249−255.

DeGrado, W. F. (1988) *Adv. Protein Chem. 39*, 51−124.

DeGrado, W. F., & Lear, J. D. (1985) *J. Am. Chem. Soc. 107*, 7684.

DeGrado, W. F., Wasserman, Z. R., & Lear, J. D. (1989) *Science 243*, 622−628.

Desjarlais, J. R., & Handel, T. M. (1995) *Protein Sci. 4*, 2006−2018.

Dunback, R. L., & Karplus, M. (1994) *Nat. Struct. Biol. 1*, 334−339.

Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.

Hagihara, Y., Tan, Y., & Goto, Y. (1994) *J. Mol. Biol. 237*, 336−348.

Handel, T. M., Williams, S. A., & DeGrado, W. F. (1993) *Science 261*, 879−885.

Harbury, P. B., Zhang, T., Kim, P. S., & Alber, T. (1993) *Science 262*, 1401−1407.

Harbury, P. B., Kim, P. S., & Alber, T. (1994) *Nature 371*, 80−83.

Harding, S. E., Rowe, A. J., & Horton, J. C. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, The Royal Society of Chemistry, Cambridge.
Haynie, D. T., & Friere, E. (1993) *Proteins: Struct., Funct., Genet. 16*, 115−140.
Hill, C. P., Anderson, D. H., Wesson, L., DeGrado, W. F., & Eisenberg, D. (1990) *Science 249*, 543−546.
Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., & Eisenberg, D. (1993) *Science 259*, 1288−1293.
McGregor, M. J., Islam, S. A., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 198*, 295−310.
Mihara, H., Tomizaki, K., Nishino, N., & Fujimoto, T. (1993) *Chem. Lett.*, 1533−1536.
Muñoz, V., & Serrano, L. (1994) *Proteins: Struct., Funct., Genet. 20*, 301−311.
Munson, M., O'Brien, R., Sturtevant, J. M., & Regan, L. (1994) *Protein Sci. 3*, 2015−2022.
Murzin, A. G., & Finkelstein, A. V. (1988) *J. Mol. Biol. 204*, 749−769.
Mutter, M., & Vuillevmier, S. (1989) *Angew. Chem. Int. Ed. Engl. 28*, 535−554.
Olofsson, S., Johansson, G., & Baltzer, L. (1995) *J. Chem. Soc., Perkin Trans. 2*, 1−10.
O'Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646−651.
O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539−544.
Pace, C. N., Shirley, B. A., & Thomson, J. A. (1989) in *Protein Structure: a Practical Approach*, pp 287−310, IRL Press, Oxford.
Phillips, G. N. (1992) *Proteins: Struct., Funct., Genet. 14*, 425−429.
Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
Prive, G., Ogihara, N., Wesson, L., Cascio, D., & Eisenberg, D. (1995) American Crystallography Association Annual Meeting.
Ptitsyn, O. B. (1992) *Protein Folding* pp 243−299, Freeman, New York.
Raleigh, D. P., & DeGrado, W. F. (1992) *J. Am. Chem. Soc. 114*, 10079−10081.
Raleigh, D. P., Betz, S. F., & DeGrado, W. F. (1995) *J. Am. Chem. Soc. 117*, 7558−7559.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648−1652.
Richardson, J. S., & Richardson, D. C. (1989) *Trends Biochem. Sci. 14*, 304−309.
Sasaki, T., & Lieberman, M. (1993) *Tetrahedron 49*, 3677−3689.
Schafmeister, C. E., Miercke, L. J. W., & Stroud, R. M. (1993) *Science 262*, 734−738.
Schrauber, H., Eisenhaber, F., & Argos, P. (1993) *J. Mol. Biol. 230*, 592−612.
Seale, J. W., Srinivasan, R., & Rose, G. D. (1994) *Protein Sci. 3*, 1741−1745.
Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., & Gilmanshin, R. I. (1991) *Biopolymers 31*, 119−128.
Seo, J., & Cohen, C. (1993) *Proteins: Struct., Funct., Genet. 15*, 223−234.
Struthers, M. D., Cheng, R. P., & Imperiali, B. (1996) *Science 271*, 392−395.
Taylor, K. S., Lou, M. Z., Chin, T. M., Yang, N. C., & Garavito, R. M. (1996) *Protein Sci. 6*, 414−421.

BI960095A