Protein Science (1998), 7:1404–1414. Cambridge University Press. Printed in the USA.
Copyright © 1998 The Protein Society

From coiled coils to small globular proteins: Design of a native-like three-helix bundle

JAMES W. BRYSON,[1,4] JOHN R. DESJARLAIS,[2,3,5] TRACY M. HANDEL,[2] AND WILLIAM F. DEGRADO[3]

[1]The DuPont Merck Pharmaceutical Company, P.O. Box 80400, Wilmington, Delaware
[2]Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 94720
[3]The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelpnia, Pennsylvania 19104-6059

(RECEIVED March 2, 1997; ACCEPTED March 4, 1998)

Abstract

A monomolecular native-like three-helix bundle has been designed in an iterative process, beginning with a peptide that noncooperatively assembled into an antiparallel three-helix bundle. Three versions of the protein were designed in which specific interactions were incrementally added. The hydrodynamic and spectroscopic properties of the proteins were examined by size exclusion chromatography, sedimentation equilibrium, fluorescence spectroscopy, and NMR. The thermodynamics of folding were evaluated by monitoring the thermal and guanidine-induced unfolding transitions using far UV circular dichroism spectroscopy. The attainment of a unique, native-like state was achieved through the introduction of: (1) helix capping interactions; (2) electrostatic interactions between partially exposed charged residues; (3) a diverse collection of apolar side chains within the hydrophobic core.

Keywords: coiled coil; genetic algorithm; hydrophobic core packing; protein design; three-helix bundle

The de novo design of proteins (DeGrado et al., 1989; Richardson & Richardson, 1989; Betz et al., 1993; Bryson et al., 1995) is becoming an increasingly feasible goal as understanding of protein structures continues to advance. Studies with designed monomeric α-helices have shown the features required for α-helix formation (Lyu et al., 1990; O'Neil & DeGrado, 1990; Padmanabhan et al., 1990; Scholtz et al., 1991; Armstrong & Baldwin, 1993; Creamer & Rose, 1994; Shalongo & Stellwagen, 1995), and crystallographic studies of designed amphiphilic α-helices have provided detailed pictures of stable helix packing geometries as starting points for designing helical proteins (Hill et al., 1990; O'Shea et al., 1991; Harbury et al., 1994; Schafmeister et al., 1997). While great advances have been made in designing helical proteins that fold into the desired aggregation state and approximate tertiary structure, many designed proteins have not exhibited the two-state thermodynamic behavior expected for native-like folding (Bryson et al., 1995). In fact, the design of proteins that specifically fold into a unique three-dimensional structure remains an important goal of protein design that has only infrequently been demonstrated (Struthers et al., 1996; Dahiyal & Mayo, 1997; Fezoui et al., 1997; Ilyina et al., 1997; Ogihara et al., 1997; Schafmeister et al., 1997).

The conformational specificity of a native protein is apparent at several levels, beginning with its overall fold and topology. Second, within this fold, there are scores of related conformations (Lesk & Chothia, 1980) that differ by rigid body shifts of secondary structural elements. Finally, the side chains in protein interiors are generally well packed and adopt a single rotameric state (McGregor et al., 1987). Specificity reflects the tendency of a protein to achieve a unique state (or set of closely related states) at each of these different levels, and requires a large energetic gap between the native state and the next higher ensemble of conformations (Lazaridis et al., 1995; Mirny & Shakhnovich, 1996; Onuchic et al., 1996). A lack of conformational specificity can have severe physiological consequences as revealed in recent studies of naturally occurring mutations in lysozyme (Booth et al., 1997) and prion proteins (for a recent review, see Mihara & Takahashi, 1997). The origins of conformational specificity are likely to be as wide-ranging as the types of interactions involved in stabilizing protein structures, and include van der Waals packing interactions, hydrogen bonding, salt bridges, and the hydrophobic effect. In a few simple cases, hydrogen bonded interactions between buried side chains have been shown to dictate conformational specificity at the expense of overall thermodynamic stability (Betz et al., 1995; Lumb & Kim, 1995; Woolfson & Alber, 1995; Schneider et al., 1997). The tight packing of amino acid side chains

Reprint requests to: William F. DeGrado, The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104–6059; e-mail: wdegrado@mail.med.upenn.edu.

[4]Present address: Ecogen Inc., 2000 Cabot Blvd. West, Langhorne, Pennsylvania 19047.

[5]Present address: 408 Chandlee Lab, Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802.

in the interior of proteins may also contribute to conformational specificity, although experiments with proteins containing repacked cores suggest that this feature alone cannot account for conformational specificity (Axe et al., 1996; Cordes et al., 1996; Vetter et al., 1996). Thus, conformational specificity reflects a variety of interactions that are broadly distributed throughout the sequence and structure of a protein.

The goal of the present work is to evolve a conformationally specific protein from a rather nonspecific aggregate of amphiphilic α-helical peptides. The insights gained in the process should provide an active understanding of the features required for engineering conformational specificity into protein sequences. Our design began with the three-dimensional structure (Lovejoy et al., 1993) of a trimeric bundle of unlinked α-helices modeled after Hodges' polyheptapeptide models for two stranded coiled coils (Lau et al., 1984). The designed peptide, CoilSer (O'Neil & DeGrado, 1990), has a periodic structure $(Leu_a\text{-}Xxx_b\text{-}Xxx_c\text{-}Leu_d\text{-}Glu_e\text{-}Xxx_f\text{-}Lys_g)_4$ (Fig. 1) and was intended to form a dimer stabilized by hydrophobic interactions between the Leu residues as well as electrostatic interactions between the Glu and Lys residues on neighboring helices. In solution, this peptide exists in a monomer/dimer/trimer equilibrium of low cooperativity.

CoilSer has also been crystallized, and its X-ray structure shows an antiparallel three-stranded coiled coil (Fig. 2A) with an up-down-up topology (Lovejoy et al., 1993). Although the peptide has a very well-packed structure, about one-third of the interior-facing Leu side-chains adopt side-chain conformations that fall outside of the two most populated rotamers for this residue in an α-helical conformation (Schrauber et al., 1993). This finding suggests that the sequence of CoilSer may not represent the best possible sequence for specifying the anti-parallel three-helix bundle structure observed in the crystal structure. Also, in the crystal structure Leu residues at equivalent positions in the individual monomers often adopted different rotamers in the three different helices of the trimer. In solution, the conformations of these side chains may rapidly interconvert, giving rise to a dynamically averaging interior structure. In previous work, we redesigned the sequence of CoilSer to provide a peptide that adopted a uniquely folded parallel three-stranded coiled coil in aqueous solution (Betz et al., 1995; Boice et al., 1996) and the solid state (Ogihara et al., 1997). In the current manuscript we address the more challenging goal of converting CoilSer into a globular, anti-parallel, three-helix bundle; a folding motif that is found in a variety of functionally diverse proteins (Deisenhofer, 1981; Gouda et al., 1992; Shaw et al., 1993; Pascual et al., 1996; Sliz et al., 1997). This is accomplished in an hierarchic approach (Bryson et al., 1995) through the introduction of specific interactions as needed to acheive a uniquely folded protein.

CoilSer:

```
        a b c d e f g  a b c d e f g  a b c d e f g  a b c d e f g
      E W E A L E K K  L A A L E S K  L Q A L E K K  L E A L E H G
```

α_3A:

```
        a b c d e f g  a b c d e f g  a b c d e f  (Loop)
      E W E A L E K K  L N A L E S K  L Q A L E K  G
      N W E A L K K E  L N A L K S E  L Q A L K K  P G
      N W E A L E K K  L N A L E S K  L Q A L E H G
```

α_3B:

```
      E W E A L E K K  L A A L E S K  L Q A L      G G
    N P D E W A A L K K E  L A A L K S E  L Q A L K    G K G
      N P E W E A L E K K  L A A L E S K  L Q A L E H G
```

α_3C:

```
      S W A E E K E E  L A A I K S E  L Q A L      G G
    S E A E  L A A E E K E  I A A E E S E  L Q A Y K    G K G
      N P E  Y E A L E K E  A A A I E S E  L Q A Y E H N
```

Fig. 1. Sequences of CoilSer and the α_3 series of peptides are given in single letter code. The sequences are arranged in heptads, with the coiled coil heptad positions labeled above CoilSer and α_3A. Residues that were altered from the previous design are underlined. The N-termini are free amines; the C-termini are amidated.

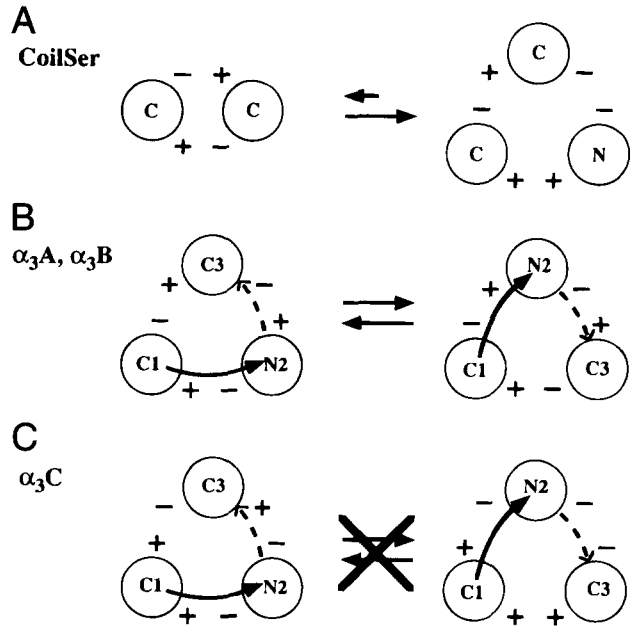

Fig. 2. Cartoon representations of the interhelical electrostatic interactions used to specify the topology of the three-helix bundle. Circles represent the α-helices, viewed end on from either the N- or C-terminal ends (marked with **N** or **C**, respectively). Numbered helices indicate their sequential positions in the single-chain three-helix bundle proteins. "−" signs indicate negatively charged Glu residues in either the "e" or "g" positions, "+" signs indicate positively charged Lys or Arg residues in the same positions. Solid or dashed arrows indicate loops connecting the helices either above or below the plane of the image, respectively. **A:** CoilSer, the interhelical electrostatics, was designed to stabilize a parallel dimer, but an antiparallel trimer is observed in the solid state with presumably unfavorable interactions between like charges on two of the three faces. Topology is irrelevant for three identical unconnected helices. **B:** Two topologies are possible for connected helices, but both are equally stabilized by the interhelical electrostatic interactions as designed in α_3A and α_3B. **C:** The redesigned interhelical electrostatic interactions in α_3C should destabilize the clockwise topology by placing like-charged amino acids along both sides of two faces of the bundle, thereby specifying a counterclockwise topology.

Results and discussion

Design

Round 1

The modeling was initiated by shortening the CoilSer trimer by one heptad to give 21 residue helices whose lengths were more typical of the α-helices in natural three-helical bundles (Speicher & Marchesi, 1984; Uhlén et al., 1984; Parry et al., 1992; Shaw et al., 1993; Sliz et al., 1997). The Trp residues were retained near the N-terminus of the α-helices, and a unique His residue near the

C-terminus of the third helix was also retained. In addition, the positions of Glu and Lys residues in helix two were rearranged to improve the interhelical electrostatic interactions. The Glu and Lys residues at positions "e" and "g" of CoilSer had been arranged to stabilize parallel helix-helix packing; thus the positions of these residues needed to be reversed in the second, antiparallel helix (Fig. 2A,B).

A major challenge in the construction of a three-helical bundle from the crystallographic coordinates of CoilSer was the design of hairpin loops. The loops should direct the formation of an intramolecularly stabilized bundle with a single topology, rather than a multimeric aggregate. Two topologies are possible for an antiparallel three-helix bundle; when viewed from the top of the structure as in Figure 2, helical bundles can adopt a clockwise or a counterclockwise topology. The designed loops should specify a unique topology rather than a mixture of the two possibilities.

In the first round of design, two simple loop sequences consisting of GlyAsn and ProGlyAsn were used, with the asparagine intended as a capping residue (Presta & Rose, 1988; Richardson & Richardson, 1988; Harper & Rose, 1993) at the beginning of the second and third helices. The glycine was included as a flexible helix terminator, and the proline inserted into the second loop was required to bridge the larger gap between helices 2 and 3. This protein, α_3A, associates weakly in a monomer/dimer/trimer equilibrium (vide infra), indicating that the loop was not capable of stabilizing the intramolecularly folded form of the protein. Thus, the loops failed to present the neighboring helices at sufficiently high effective intramolecular concentration to allow internal folding to compete with intermolecular self-association at high micromolar concentrations.

Round 2

A primary goal of the second round of design was to introduce stronger helix stop signals to direct the formation and conformation of the loops. The loop connecting the first two helices features an "N-X-X-E" box, in which the Asn side chain carbonyl "caps" the N-terminus of helix two through a H-bond with a main-chain amide, and the Glu side chain carbonyl makes a reciprocal H-bond with a main-chain amide of the loop (Harper & Rose, 1993; Zhou et al., 1994). The loop on the opposite end of the bundle also features an Asn as an N-cap residue, but does not include a glutamate residue at the N-cap + 3 position as this position is occupied by a Trp residue, the first hydrophobic position of the final helix. Pro residues occupy the N-cap + 1 position as is commonly observed in proteins (Richardson & Richardson, 1988; Dasgupta & Bell, 1993).

The helix capping interactions serve two functions. The first is to provide a strong helix start/stop signal to ensure the protein does not form a long helix capable of aggregation. The second is to dictate the trajectory of the loop as it enters the helix, in effect determining the relative disposition of the two helices. Observing the helix vertically down its axis, the position of the helix capping box determines the trajectory of the loop as it enters the helix, thus influencing whether the neighboring helix will pack on the right or left (Efimov, 1991, 1993). The helix capping boxes can therefore be used to help define the topology of the bundle.

The second protein, α_3B, is indeed monomeric (vide infra), indicating that the incorporation of helix capping interactions prevents aggregation. In many ways α_3B approaches the desired native-like behavior of natural proteins, but some characteristics of the "molten globule" state remain.

Round 3

The goal of the third round of design was to eliminate any remaining nonnative characteristics observed for α_3B. One possible source of nonnative behavior of α_3B is coexistence of the two topologies possible for a three helix bundle: clockwise turning or counterclockwise turning (see Fig. 2B). With equivalent hydrophobic core residues contributed by each helix and equivalent electrostatic interactions between the Lys and Glu residues at the "e" and "g" positions, the interhelical interactions should be identical in either topology. Only the position and length of the loops would favor one topology over the other. The other possible cause of dynamic behavior in α_3B is the redundancy in the hydrophobic core. Side-chain repacking calculations (Desjarlais & Handel, 1995) indicate that the 15 Leu residues in α_3B are able to pack in many different combinations of conformations with roughly equivalent overall energies, suggesting that a different combination of hydrophobic residues could pack with greater specificity. Continuing with the hierarchic approach, the third round of design therefore focused on repacking of the hydrophobic core, rearrangement of the interhelical electrostatic interactions, and helix capping interactions.

Helix capping interactions

A helix capping box was incorporated into the first helix by exchanging the positions of amino acids three and four to move a Glu residue to the N-cap + 3 position. The N-cap residue for the first two helices was changed to Ser, as this residue often offers greater stabilization than Asn in such capping boxes (Lyu et al., 1992).

Redesign of the interhelical electrostatic interactions

In α_3B, Lys and Glu residues placed at the "e" and "g" positions were meant to favor the antiparallel arrangement of the three helices (O'Shea et al., 1993; Lumb & Kim, 1996). The arrangement of these residues, however, leads to equivalent electrostatic interactions for either of the two possible antiparallel topologies (see Fig. 2A). In the design of α_3C, the charged residues were rearranged to preferentially favor the counterclockwise topology; positively charged residues were placed at both the "e" and "g" positions of helix one, negatively charged residues were placed at both the "e" and "g" positions of helix three, and negatively charged residues were placed at the "e" positions and positively charged residues at the "g" positions of the antiparallel helix two. As illustrated in Figure 2B, this arrangement should destabilize the clockwise arrangement of helices and stabilize the desired, counterclockwise topology. In addition, Arg was used in α_3C in place of some of the Lys residues. This replacement reduces the sequence redundancy and may provide added stability as Arg-Glu interactions are energetically more favorable than Lys-Glu interactions in coiled coils (Krylov et al., 1994).

Redesign of the hydrophobic core

The third and perhaps most significant refinement was the complete redesign of the hydrophobic core, using a genetic repacking algorithm (repacking of core (ROC), Desjarlais & Handel, 1995). A number of constraints were imposed on the sequence: a unique Trp residue was retained at the first hydrophobic position of helix 1 as a fluorescent probe (although its conformation was determined

by the algorithm), and all of the remaining core positions (17 residues) were allowed to be Ala, Val, Ile, Leu, or Phe. Thirty runs of the program generated 30 different sequences, and amino acid residues that occurred with very high frequency were fixed in a second round of optimization with ROC (Table 1).

Eight positions were fixed in the second round of repacking, although their conformations were determined by the algorithm. Thirty runs of ROC in the second round produced 18 different sequences. Models were built for several of the predicted structures, and on the basis of energy minimizations, solvent accessible surface areas, and the sequence consensus from the second round, four additional positions were fixed (Phe29, Phe36, Val51, Phe68) based on favorable packing interactions and rotamers. In addition, position 58 was released for re-optimization due to possible clashes between Ile58 and Phe36.

In the third round of optimization, 11 residues were fixed to a particular amino acid, and 30 runs of ROC produced only six different sequences. From the consensus of these sequences, an additional six residues were fixed, all to Leu. In the fourth and final round of optimization, only one residue, position 58, was allowed to vary. ROC predicted three possibilities for position 58; Ala, Ile, and Leu. Although Ile occurred most frequently, the conformation predicted for it was not particularly favorable. Ala was selected for this position rather than Leu in order to maximize the diversity of the amino acid content of the hydrophobic core. The model for α_3C generated from the predictions of ROC was further modified after energy minimization and solvent-accessible surface area evaluations. Two of the Phe residues (positions 43 and 68) were changed to Tyr residues because of their partial exposure to the surface.

The sequence of α_3C is shown in Figure 1, and the model of the hydrophobic core is shown in Figure 3. Eleven of the 18 core positions have changed; seven Leu residues are present, six of which were present in α_3B. The addition of several β-branched amino acids (Val, Ile) may decrease the conformational flexibility of the core because, within an α-helix, these residues have only one low-energy torsion angle (McGregor et al., 1987; Schrauber et al., 1993). A Val residue at position 51 packs against the indole group of Trp2. Also three Ile residues at positions 12, 33, and 61 from each of the three helices pack near the center of the bundle in a layer that lies roughly orthogonal to the central axis of the bundle. The hydrophobic residues above and below this triad of Ile side chains also pack in layers and show geometric complementarity in their size and shape. For instance Leu9, Phe36, and Ala58 pack above the Ile triad. Three other layers contain one aromatic residue and two Leu residues.

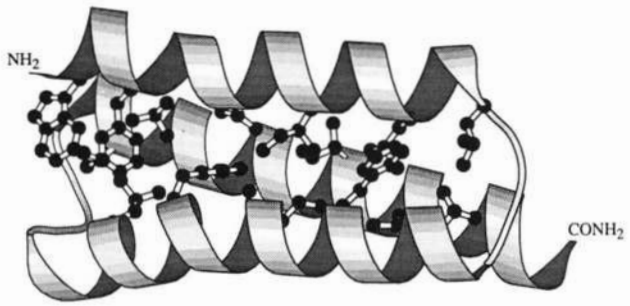

Fig. 3. MOLSCRIPT diagram (Kraulis, 1991) showing the molecular model of α_3C. Side chains of the amino acids of the repacked hydrophobic core consisting of all "a" and "d" positions are shown in ball-and-stick representation.

Characterization

Sedimentation equilibrium

The hydrodynamic radius of α_3A and α_3B in solution was initially investigated by size exclusion chromatography using a Superdex75 (Pharmacia) column eluted with 20 mM HEPES (pH 7.0), 750 mM NaCl. Both peptides eluted with apparent molecular weights approximately 1.5-fold higher than the calculated molecular weights, consistent with the asymmetric geometry of the model. Thus, this technique suggests that both proteins were monomeric. However, sedimentation equilibrium at 50–250 μM peptide concentration detected a weak association of α_3A, which was best fit (see Materials and methods) to a monomer/dimer/trimer equilibrium. Figure 4A shows the data for α_3A at 250 μM peptide in 50 mM MOPS, 750 mM NaCl, pH 7.0 along with the best fits calculated for pure monomer, dimer, and trimer. When fit to a single species, a molecular weight (MW) of 13,100 Da is predicted (actual molecular weight of α_3A = 7,762 Da). The best fit is obtained for a mixed monomer/dimer/trimer equilibrium (not shown). Similar results are observed in 50 mM MOPS, 100 mM NaCl, pH 7.0 (data not shown). Due to this aggregation behavior, α_3A was not characterized further.

In contrast, α_3B and α_3C were found to be completely monomeric. Sedimentation equilibrium data for α_3B and α_3C (Fig. 4B,C) are well described as monomolecular species of molecular mass within 4% of that predicted from the amino acid sequence. The expected and observed MW for α_3B were 7,762 and 7,715, respectively; the corresponding values for α_3C were 7,878

Table 1. *Amino acids constrained at the start of each round of ROC*[a]

Position:	2	5	9	12	16	19	26	29	33	36	40	43	51	54	58	61	65	68
Round 1	W																	
Round 2	W	F	L	I					I			F			I	I		
Round 3	W	F	L	I				F	I	F		F	V			I		F
Round 4	W	F	L	I	L	L	L	F	I	F	L	F	V	L		I	L	F
α_3C	W	F	L	I	L	L	L	F	I	F	L	F	V	L	A	I	L	F

[a]The positions refer to the residue numbers in the designed protein, as opposed to the structure of CoilSer. Helix 1 spans residues 2–19, helix 2 spans 23–43, and helix 3 spans residues 49–69.

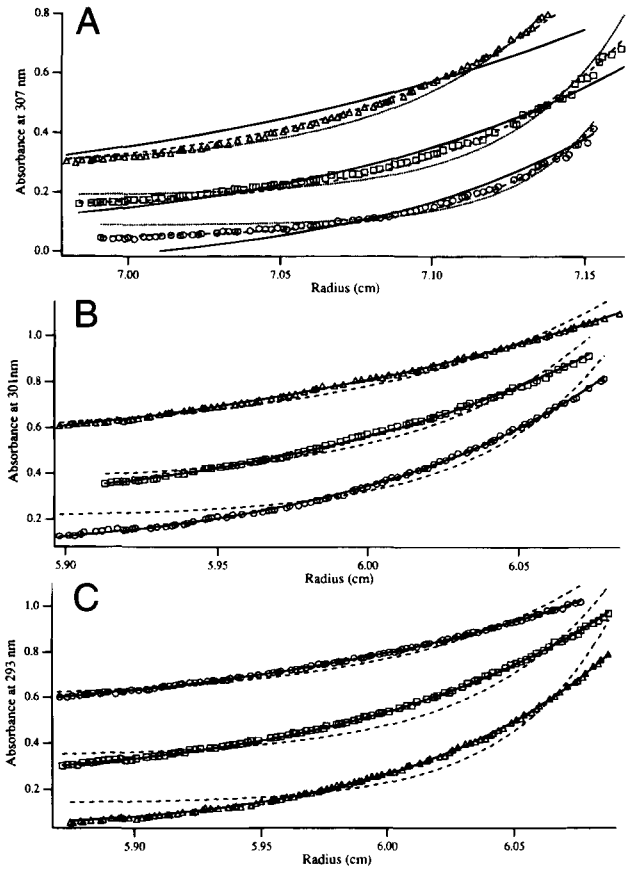

Fig. 4. Sedimentation equilibrium. Analytical ultracentrifugation data showing absorbance versus radius for raw data collected at 35,000 rpm (△), 42,000 rpm (□), and 48,000 rpm (○), and theoretical fits to monomer (solid line), dimer (dashed line), and trimer (dotted line, α_3A only). **A:** Data and fits showing α_3A self-associates. **B:** Data and fits showing α_3B is monomeric. **C:** Data and fits showing α_3C is monomeric.

and 7,560. Attempts to fit the data to other aggregation states, such as a dimer (Fig. 4B), resulted in considerably poorer fits to the experimental data.

Chemical denaturation

α_3B is predominantly helical as assessed by far UV CD spectroscopy (mean residue ellipticities of $[\theta_{222}] = -22{,}800$, $[\theta_{208}] = -18{,}600$, $[\theta_{192}] = 32{,}500$ deg cm^2 dmol^{-1} observed at pH 7.0, 25 °C). Further, the helical content was lost in a single cooperative transition when the protein was titrated with GdnHCl. The curve is well described by Equation 1:

$$\Delta G_{obs} = \Delta G_D + m[\text{GdnHCl}] \quad (1)$$

in which ΔG_{obs} is the free energy for the two-state unfolding equilibrium observed at a given concentration of GdnHCl, ΔG_D is the free energy of denaturation extrapolated to zero [GdnHCl], [GdnHCl] is the concentration of guanidinium hydrochloride, and m is a constant that provides a measure of the cooperativity of the process. The values of ΔG_D, m, and the midpoint of the transition are 7.2 kcal mol^{-1}, 1.7 kcal mol^{-1} M^{-1}, and 4.2 M, respectively. α_3C is also predominantly helical, and shows a cooperative unfolding transition with $\Delta G_D = 5.5$ kcal mol^{-1}, $m = 2.1$ kcal mol^{-1} M^{-1}, and $[\text{GdnHCl}]_{1/2} = 2.6$ M. For comparison the respective values for the Z domain of staphylococcal Protein A are 6.6 kcal mol^{-1}, 1.7 kcal mol^{-1} M^{-1}, and 3.9 M, respectively (Cedergren et al., 1993). Thus, the stability and cooperativity of the transitions for α_3B and α_3C are close to those observed for a natural three-helix bundle protein.

The slight decrease in stability of α_3C relative to α_3B may reflect a large number of factors. Interestingly, differences in the helical propensities are able to account for the entire (1.7 kcal mol^{-1}) stability difference between the two proteins. The helix propensity scale developed using CoilSer predicts that α_3C should be less stable than α_3B by 1.9 kcal mol^{-1}, in close agreement with the experimental value (O'Neil & DeGrado, 1990). Clearly, there are also many other differences in the van der Waals packing and hydrophobic interactions between the two structures. Evidently, these factors are compensatory in nature and result in little experimental difference in stability. Thus, increases in stability arising from improvements in the packing of the core of α_3C may be offset by a loss of conformational entropy in α_3B.

The value of m obtained from an analysis of GdnHCl-induced denaturation curves provides a measure of the difference in the accessibility of the native versus the unfolded states of a protein to the denaturant. The value of m is larger for α_3C than for α_3B, indicating that the former protein experiences a larger change in accessibility to GdnHCl upon unfolding. Given that these proteins unfold at relatively high concentrations of GdnHCl, it is most likely that their unfolded states are similar and resemble fully denatured random coils. Thus, the difference in m may reflect an improvement in the packing of the core of α_3C, which results in a decreased exposure of the native state to guanidinium ions.

Thermal denaturation

Thermal melting curves of α_3B and α_3C at various concentrations of [GdnHCl] are illustrated in Figures 5A and 6A and are typical of small proteins, showing evidence for both low- and high-temperature denaturation (Becktel & Schellman, 1987; Chen & Schellman, 1989). The data were fit to the Gibbs-Helmholtz Equation 2:

$$\Delta G_T = \Delta H_{T_m}(1 - T/T_m) + (T - T_m)\Delta C_p - T\Delta C_p \ln(T/T_m) \quad (2)$$

where T_m is the midpoint of the transition, ΔH_{T_m} is the van't Hoff enthalpy at the midpoint, and ΔC_p is the change in the heat capacity averaged over the temperature range of the experiment. The values of ΔH_{T_m}, ΔC_p, and T_m could be obtained with reasonable confidence between 3 and 4 M GdnHCl for α_3B and between 1.5 and 3 M GdnHCl for α_3C, where both phases of the denaturation curve were resolved. Thermodynamic parameters derived from fitting the baseline corrected thermal melts displayed in Figures 5B and 6B are listed in Table 2.

The value of ΔC_p was also determined by evaluating the van't Hoff enthalpy in the vicinity of T_m at each concentration of [GdnHCl]. A plot of $\Delta H_{(\text{van't Hoff})}$ versus T_m is linear over a wide range of [GdnHCl] and temperatures, ranging from 31 to 89 °C. The value of ΔC_p determined from the slope of this plot is 589 cal mol^{-1} K^{-1}, in excellent agreement with the average of the values of ΔC_p

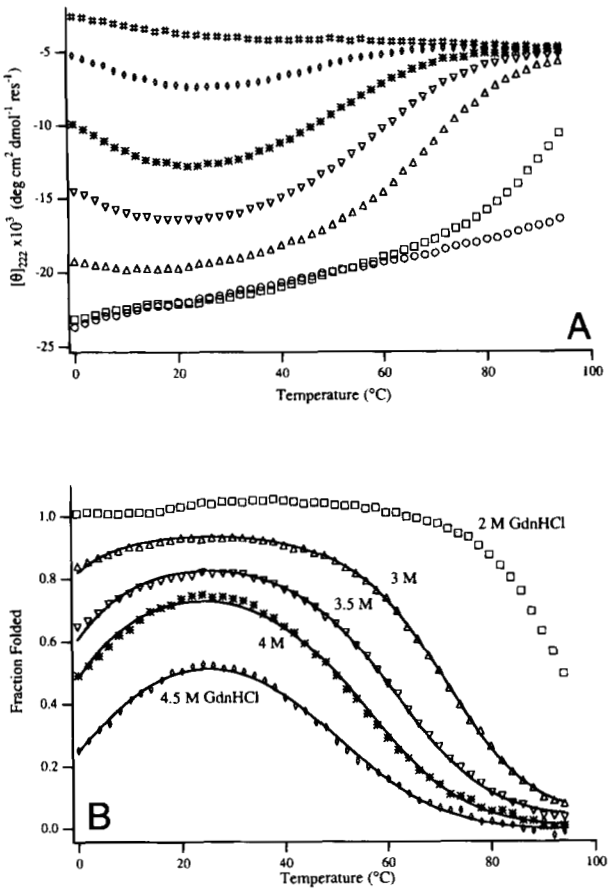

Fig. 5. Thermal denaturation of α_3B. Circular dichroism data monitoring the loss of helical signal at 222 nm with increasing temperatures at varying concentrations of GdnHCl: 0 M (○), 2 M (□), 3 M (△), 3.5 M (▽), 4 M (✳), 4.5 M (◇), and 5 M (#). **A:** Uncorrected CD data. **B:** Baseline corrected data converted to fraction folded along with fits (solid lines) to the Gibbs-Helmholtz expression (see text).

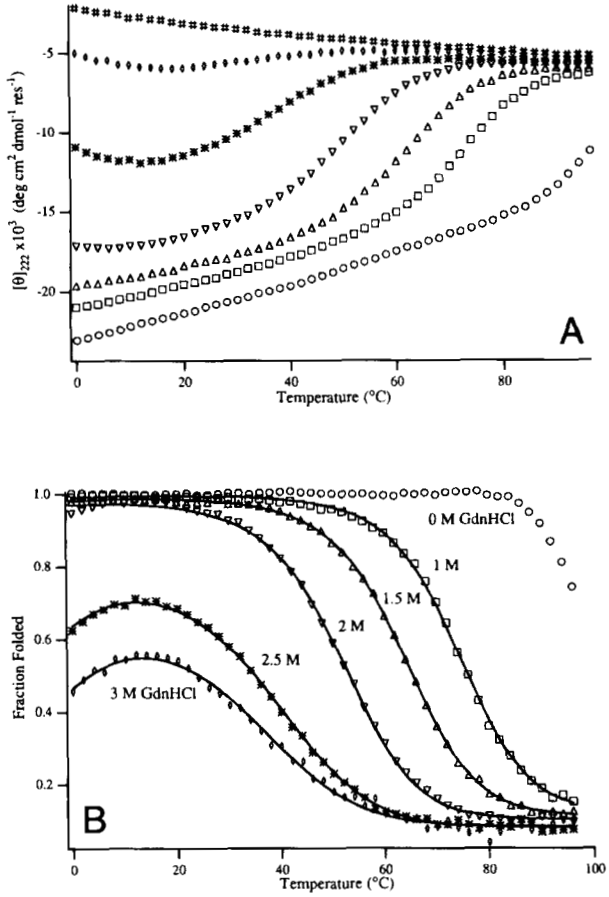

Fig. 6. Thermal denaturation of α_3C. Circular dichroism data monitoring the loss of helical signal at 222 nm with increasing temperatures at varying concentrations of GdnHCl: 0 M (○), 1 M (□), 1.5 M (△), 2 M (▽), 2.5 M (✳), 3 M (◇), and 4 M (#). **A:** Uncorrected CD data. **B:** Baseline corrected data converted to fraction folded along with fits (solid lines) to the Gibbs-Helmholtz expression (see text).

derived from fitting the fraction unfolded to the Gibbs-Helmholtz expression for α_3B, which is 574 ± 42 cal mol^{-1} K^{-1}. The value of ΔC_p on a per residue basis (8.1 cal mol^{-1} K^{-1} res^{-1}) is within the range of theoretical values calculated from the model of α_3B based on the solvent accessibility of its polar and nonpolar groups (Makhatadze & Privalov, 1990; Privalov & Makhatadze, 1990; Murphy & Freire, 1992; Spolar et al., 1992). While this value is somewhat lower than the value found for native proteins (10–14 cal mol^{-1} K^{-1} per residue) (Kuwajima, 1989; Ptitsyn et al., 1990), the difference can be explained by the elongated structure and high surface area to volume ratio of the protein. It should also be noted that the value of ΔC_p for α_3B is greater than that observed for both the unfolding of the dimeric coiled-coil GCN4 (3.8 cal mol^{-1} K^{-1} for GCN4) (Thompson et al., 1993), and for most molten globule proteins (Freire, 1995).

Thermal melting curves for α_3C are displayed in Figure 5A, and the baseline-corrected curves at 1.5–3 M GdnHCl are displayed along with the best fits to the Gibbs-Helmholtz equation in Figure 6B. Thermodynamic parameters derived from those fits are listed in Table 2. ΔC_p values from these four fits range from 593 to 663 cal mol^{-1} K^{-1}, with an average of 627 ± 32 cal mol^{-1} K^{-1}. ΔC_p was also analyzed by calculating the van't Hoff enthalpies versus $T_m(T_{ref})$ between 0 and 3.0 M GdnHCl; the resulting plot is linear over the entire range examined. The computed value of ΔC_p (688 cal mol^{-1} K^{-1}) is in very good agreement with that determined using the Gibbs-Helmholtz method, particularly when one considers the large range in temperature (17 to 93 °C) over which the van't Hoff enthalpies were measured. Taking the mean of the two measures of the change in heat capacity gives ΔC_p = 658 cal mol^{-1} K^{-1} or 9.2 cal mol^{-1} K^{-1} res^{-1} for α_3C. This value is similar to that observed for the trimeric coiled-coil peptide Coil-V_aL_d (ΔC_p = 8.6 cal mol^{-1} K^{-1} per residue) (Boice et al., 1996), which has a cooperatively folded three-helix structure with a well-packed interior.

The change in heat capacity for α_2C is approximately 15% greater than that observed for α_3B. The value of ΔC_p associated with protein folding is often interpreted as a measure of the difference in accessibility of the native versus the unfolded state of a protein (Murphy & Freire, 1992; Privalov, 1992). Both ΔC_p and m increase for α_3C versus α_3B, suggesting that α_3C is slightly better packed than its predecessor. This is most likely due to the greater specificity in the packing interactions of the redesigned hydrophobic core, as a control protein that contained the redesigned capping and electrostatic interactions of α_3C but retained the all-leucine

Table 2. *Thermodynamic parameters*[a] *of* α_3B *and* α_3C

Protein	[GdnHCl] (M)	ΔH_{T_m} (kcal mol^{-1})	ΔC_p (cal mol^{-1} K^{-1})	T_m (°C)	$\Delta H_{(\text{van't Hoff})}$ (kcal mol^{-1})
α_3B	2			(T_{ref} = 89)	37.7
	3	24.7	550	70	24.9
	3.5	17.7	526	58	17.4
	4	15.5	616	49	15.7
	4.5	3.7	602	31	2.8
α_3C	0			(T_{ref} = 93)	58.8
	1	36.6	650	74	34.0
	1.5	32.2	633	63	28.1
	2	28.1	593	51	24.1
	2.5	13.1	663	32	12.5
	3	2.2	595	17	1.7

[a] ΔH_{T_m}, ΔC_p, and T_m were determined by nonlinear least-squares analysis, using the Gibbs-Helmholtz equation. $\Delta H_{(\text{van't Hoff})}$ was determined from van't Hoff plots evaluated at T_m (or at T_{ref} for conditions where T_m was not reached in the thermal denaturation curves).

core exhibited thermodynamic behavior essentially identical to that of α_3B (data not shown).

Binding of 8-anilino-naphthalene sulfonic acid (ANS)

Proteins in the molten globule state tend to bind the environment-sensitive fluorescent dye ANS more tightly than do their native counterparts (Semisotnov et al., 1991; Ptitsyn, 1992). Figure 7 illustrates titrations of α_3B, α_3C, lysozyme (as a example of a well-folded native protein), and a designed dimeric four-helix bundle (α_2B) that has previously been determined to form a molten globular state (Handel et al., 1993) into a 1 μM solution of ANS. The titrations of α_2B and α_3B clearly show that the affinity of α_3B for ANS is much lower than that expected for a typical molten globule. The titration of α_3C shows that it has an even lower affinity for ANS than does α_3B and indeed is within the range observed for natural native proteins, providing further evidence that α_3C closely mimics native-like behavior.

NMR spectra

The ^{1}H NMR spectra of native proteins show a large spread of chemical shifts induced by the unique packing of their interior side chains. By contrast, the spectra of proteins in more flexible states such as the molten globule or random coil show less dispersion due to dynamic averaging of the environments of the individual side chains. The NMR spectrum of α_3B shows a limited degree of dispersion, due in part to the redundancy of the Leu residues in its core (see Fig. 8). There is little dispersion of the chemical shifts in the aliphatic region. Nevertheless, several methyl resonances are

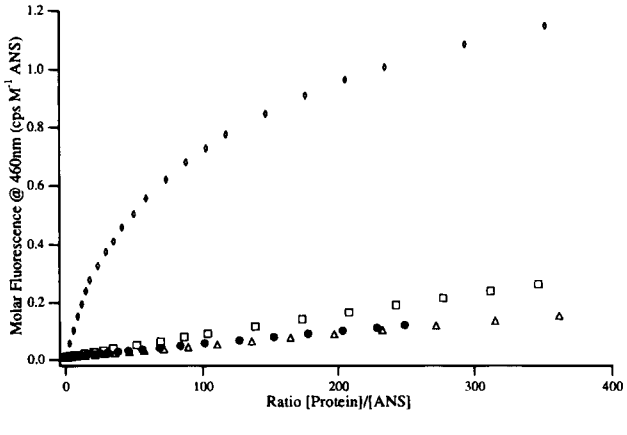

Fig. 7. ANS binding. Data for titrations of α_3B, α_3C, α_3B, and lysozyme into a 1 μM solution of ANS in 20 mM MOPS, pH 7.0. Molar fluorescence ($\times$ 10^{-2} based on actual ANS concentrations) is plotted versus the molar ratio of the proteins to ANS. Strong binding of ANS is observed for α_2B (◇) (a molten globular "minimalist" designed protein (Handel et al., 1993)), while very weak binding is observed for the native protein lysozyme (△). α_3B (□) shows a much lower affinity for ANS than the molten globular α_2B, but somewhat greater affinity than lysozyme, while α_3C (●) shows very weak binding similar to that of the native protein.

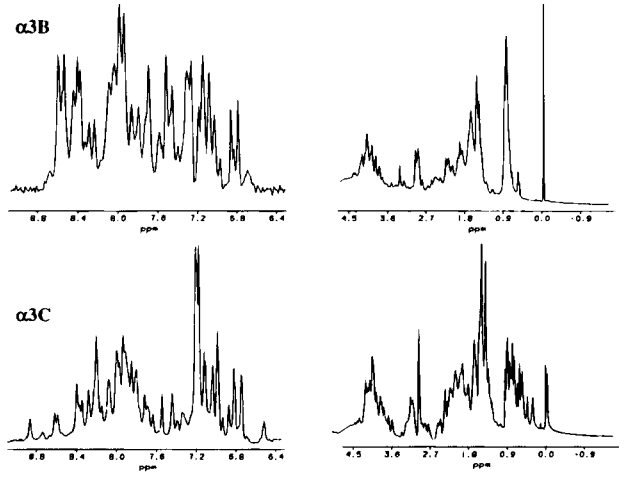

Fig. 8. 1H NMR spectra. The amide/aromatic and aliphatic regions of the one-dimensional 1H NMR spectra of α_3B and α_3C are shown. The spectrum of α_3C exhibits narrower line widths and greater dispersion in both the amide and aliphatic regions.

observed to be shifted upfield by 0.5 ppm relative to the large peak associated with the other Leu methyl resonances, consistent with the model which predicts that the indole side chains of Trp26 and Trp51 partially penetrate the core. Further, the indole resonances are nonequivalent, indicating they each experience unique magnetic environments.

The one-dimensional ^{1}H NMR spectra of α_3C shows much greater dispersion in the amide and aliphatic regions than that of α_3B (see Fig. 8). Further, the aliphatic region shows a number of well-resolved upfield-shifted methyl resonances. The 2D NOESY spectra show good dispersion in the amide region and a large number of well-resolved amide proton to aliphatic proton (NOESY) cross peaks (data not shown). These data suggest that α_3C adopts a unique structure and that it should be possible to determine its structure at high resolution using isotope-edited NMR.

Hydrogen/deuterium exchange (HDX)

We determined the rate of amide hydrogen/deuterium exchange of α_3B at $pD_{read} = 5.0$ by integrating the total intensity of the amide envelope as a function of time after dissolving the protein in D_2O. Several phases were observed, and after several hours peaks associated with approximately ten amide resonances remained. These amides exchanged together as a group with a pseudo-first order rate constant of 0.0014 min^{-1}. This value is 4,500-fold slower than the value expected for a random coil, corresponding to an energetic difference of 5.0 kcal mol^{-1}, which is somewhat lower than the overall stability of the protein (7.2 kcal mol^{-1}) determined by guanidine denaturation. This discrepancy indicates that HDX of even the slowest amide protons occurs through some conformational process that does not require total denaturation of the protein. This behavior contrasts with that observed for native proteins, whose exchange requires global denaturation of the entire protein (Bai et al., 1995; Chamberlain et al., 1996). Instead, by this criterion, α_3B behaves somewhat like molten globules, which undergo particularly facile local unfolding transitions that allow rapid HDX. The HDX of α_3C has been conducted on an ^{15}N-labeled, bacterially expressed derivative, and will be reported elsewhere together with the solution structure of the protein.

Conclusions

These results illustrate a hierarchic approach to the design of native proteins beginning with the coordinates of an idealized coiled coil. The design began with idealized amphiphilic helices that hydrophobically associated into a coiled coil, and proceeded in several steps to include: (1) capping interactions at the ends of the helices; (2) electrostatic interactions between partially exposed side chains at the helix/helix interfaces; (3) a hydrophobic core that incorporated a variety of hydrophobic residues. The proteins were examined by a battery of biophysical techniques, illustrating the hierarchic requirements for adopting a native fold. Each of these steps increased the conformational specificity of the protein; the initially designed α_3A had a strong tendency to aggregate. This problem was addressed in the design of α_3B, but its behavior fell somewhat short of the expectation for a native protein, particularly with respect to HDX of its amide protons. The third design, α_3C, was entirely successful, and appeared to be native-like in all respects examined. This protein has recently been cloned and expressed in bacteria, and its structure is being determined by NMR spectroscopy. It will be interesting to determine how closely the three-dimensional structure of α_3C matches the intended design.

Materials and methods

Design

Molecular design was performed using the InsightII molecular modeling package (Biosym) on a Silicon Graphics Indigo^2 workstation. Energy minimization was carried out using Discover (Biosym). The core repacking programs NBSEARCH and ROC (Desjarlais & Handel, 1995) were compiled and run on a Silicon Graphics Indigo^2 workstation. With the exception of three residues flanking each of the designed loops, the backbone coordinates of the body of all three helices were not changed from those observed in the CoilSer crystal structure (1COS.pdb) until the final minimization of the model of α_3C.

Materials

Fmoc-protected amino acids (Fmoc = 9-fluorenylmethoxy carbonyl), PAL resin (Albericio & Barany, 1987) (PAL = 5-[4-(aminomethyl)-3,5-*bis*(methoxy)phenoxy] valeric acid), TBTU (2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate) and HOBt (N-hydroxybenzotriazole) were purchased from Millipore. All solvents and chemicals used in the peptide synthesis and purification were of the highest available grade and were used without further purification.

Peptide synthesis and purification

Peptides were synthesized on a Milligen 9050 automated peptide synthesizer using PAL resin (Albericio & Barany, 1987) (substitution 0.38 mmol/g) on a 0.2 mmol scale using standard Fmoc-protection solid phase methodology. Following cleavage from the resin with TFA (trifluoroacetic acid), the peptides were left with a C-terminal amide and N-terminal amino groups. The sequences synthesized are listed in Figure 1. The side-chain protections were: Glu(O*t*Bu) (O*t*Bu = *tert*-butyl ester), Trp(Boc) (Boc = *tert*-butoxycarbonyl), Lys(Boc), Ser(*t*-Bu) (*t*-Bu = *tert*-butyl ether), His(Boc), Tyr(*t*-Bu), Arg(PMC) (PMC = pentamethylchromansulfonyl). Fmoc-amino acid derivatives were either used as pentafluorophenyl active esters (OPfp) or as free acids activated by the TBTU/HOBt procedure (Knorr et al., 1989). Amino acid coupling times were 45 min. Double coupling was used for all Asn, Gln, Ile, Val, Trp, and Pro residues, for all residues coupled to a Pro, Ile, or Val residue, and for the C-terminal amino acid. With the exception of Trp residues, all double couplings used -OPfp activated esters for the first coupling and the free acid with TBTU/HOBt activation for the second coupling. Trp and Ser were used only as free acids with TBTU/HOBt activation. Unreacted chains were capped by acetylation at each step of the synthesis using acetic anhydride/pyridine in DMF (4:4.7:91.3 v/v, 5 min). The Fmoc protecting group of the final amino acid was removed with piperidine in DMF (20% v/v, 10 min). Cleavage from the resin and side-chain deprotection were accomplished by reaction with a mixture of TFA/thioanisole/1,2-ethanedithiol/anisole (9:0.5:0.3:0.2, v/v) at room temperature for 2 h. The resins were filtered off and the crude peptides precipitated by partial evaporation and addition of cold diethyl ether. The solid peptides were dried and purified by

reverse-phase HPLC on a preparative Vydac C-18 column using linear gradients of water and acetonitrile containing 0.1% TFA. (α_3C was first partially purified by reverse-phase HPLC using linear gradients of water and acetonitrile containing 0.2% triethylammonium-phosphate, pH 8.0 using a POROS reverse-phase column (PerSeptive); after lyophilization, α_3C was further purified as above.) The peptides were determined to be homogeneous by analytical reverse-phase HPLC. Confirmation of the expected molecular mass was obtained by electrospray ionization mass spectrometry.

Analytical ultracentrifugation

Sedimentation equilibrium analysis was performed with a Beckman XLA analytical ultracentrifuge (Harding et al., 1992). Purified peptides were dissolved in buffer and their concentrations determined spectrophotometrically in 6 M GdnHCl (guanidinium hydrochloride) by absorbance at 280 nm assuming extinction coefficients of 5,700 M^{-1} cm^{-1} for each Trp and 1,300 M^{-1} cm^{-1} for each Tyr residue (Creighton, 1993). Samples were centrifuged at three speeds, 35,000, 42,000, and 48,000 rpm, each equilibrated at least 20 h. Equilibrium was assumed when successive radial scans at the same speed were identical. Data were analyzed with the program IgorPro (Wavemetrics) using a modified version (J.D. Lear, pers. comm.) of the algorithm developed by Brooks et al. (1993). Partial specific volume of the proteins was calculated using the method of Cohn and Edsall (Cohn & Edsall, 1943).

Chemical denaturation

The loss of the helical CD signal at 222 nm was monitored with increasing concentration of GdnHCl in solutions containing identical concentrations of protein. All solutions contained 25 mM MOPS, 100 mM NaCl, pH 7.3. Samples of α_3B were 19 μM, and those of α_3C were 20 μM. Ellipticity at 222 nm was monitored using an Aviv 62DS Spectropolarimeter with a temperature control unit maintaining a constant temperature of 25 °C. All data were collected using a 0.1 cm path length cell, ellipticity at 222 nm was averaged over 60 s. The data were analyzed using the equation and method described in Santoro and Bolen (1988) with the software package IgorPro (Wavemetrics, Inc., Lake Oswego, OR), allowing for linear variation in both the folded and unfolded baselines.

Thermal denaturation

The loss of the helical CD signal at 222 nm was monitored with increasing temperature for solutions containing identical concentrations of protein and different concentrations of GdnHCl. All solutions contained 25 mM MOPS, 100 mM NaCl, pH 7.3. Samples of α_3B were 167 μM and those of α_3C were 41 μM. Ellipticity at 222 nm was monitored using an Aviv 62DS Spectropolarimeter with temperature control unit, in a 0.02 cm path length cell. The temperature was ramped from 0 to 96 °C in 2° steps. At each temperature the cell chamber was equilibrated for 5 min then data were collected and averaged over 99 s. The data were analyzed using the software package IgorPro (Wavemetrics, Inc.). The baselines for the ellipticity of the folded and unfolded proteins are dependent on both the concentration of GdnHCl and the temperature. Baseline corrected data were converted to fraction folded (FF) using Equation 3:

$$FF(T) = (\theta(T) - \theta_{UN}(T))/(\theta_{FOLD}(T) - \theta_{UN}(T)) \quad (3)$$

where FF(T) is the resulting fraction folded for each temperature T, and $\theta(T)$ is the observed ellipticity at temperature, T. $\theta_{UN}(T)$ is a linear function (Equation 3A) describing the "unfolded baseline" (i.e., the ellipticity of the unfolded protein as a function of temperature), and $\theta_{FOLD}(T)$ is a linear function describing the "folded baseline" (Equation 3B).

$$\theta_{UN}(T) = \theta_{UN}(0) + m_{UN}(T) \quad (3A)$$

$$\theta_{FOLD}(T) = \theta_{FOLD}(0) + m_{FOLD}(T) \quad (3B)$$

The values of $\theta_{FOLD}(0)$ and $m_{FOLD}(T)$ represent, respectively, the ellipticity at $T = 0$, and the change in the ellipticity with respect to temperature; these parameters are easily determined at low guanidine concentrations where the protein is fully folded at low temperatures. At higher concentrations of [GdnHCl], the values of $\theta_{UN}(T)$ were assumed to vary linearly with [GdnHCl] as in the analysis of isothermal GdnHCl-induced denaturation curves (Santoro & Bolen, 1988). The variation in the unfolded baseline with respect to [GdnHCl] at 25 °C is known from the above-mentioned analysis of the isothermal data. Thus, values of $\theta_{UN}(T)$ at a given [GdnHCl] were multiplied by a single scaling factor, such that $\theta_{\mathrm{UN}}(T)$, when evaluated at 25 °C, would match the value extrapolated from an analysis of the GdnHCl denaturation curves. The same procedure was applied to the analysis of the unfolded baseline.

The fraction folded curves for selected concentrations of GdnHCl were plotted vs. temperature (K), and fit to the Gibbs-Helmholtz equation (Equation 2). The function fitted is given by Equation 4:

$$FF = 1/(1 + \exp(-\Delta G_d/RT)) \quad (4)$$

where ΔG_d is given by the Gibbs-Helmholtz expression in Equation 2. The values of ΔH_{T_m}, ΔC_p, and T_m could be obtained with reasonable confidence when the data sets included both heat and cold denaturation (Chen & Schellman, 1989).

A second method was also applied to determine ΔC_p. The van't Hoff enthalpies, $\Delta H_{T_{ref}}$ were determined from the slope of a plot of $R \ln K$ vs. $1/T$ evaluated near the midpoint of the transition and ΔC_p was determined from the slope of the linear plot of $\Delta H_{T_{ref}}$ vs. T_{ref} (Santoro & Bolen, 1992; Agashe & Udgaonkar, 1995).

The self-consistency of the chemical and thermal denaturation data was checked by comparing the values of $\Delta G_{\mathrm{unfolding,25\,°C}}$ calculated by the Gibbs-Helmholtz expression using the values of T_m, ΔH_{T_m}, and ΔC_p derived from fitting the thermal denaturation data at a given [GdnHCl] to the values calculated from the chemical denaturation curves at the same [GdnHCl]. [GdnHCl] concentrations were chosen near the beginning of the chemical denaturation curves where extrapolations would be minimal: 3.5 M GdnHCl for α_3B and 2 M GdnHCl for α_3C. The values of $\Delta G_{\mathrm{unfolding,25\,°C}}$ agreed within 0.5 kcal $mole^{-1}$ for both proteins.

Fluorescence spectroscopy

Fluorescence spectra and titrations were measured using a Spex Fluorolog fluorimeter at 298 K using a fixed amount of 8-anilino-1-naphthalene sulfonic acid (ANS) of 1 μM starting concentration. The proteins were titrated into that solution and the fluorescence at 460 nm monitored with excitation at 370 nm. Emission spectra

were recorded on the titration endpoint solutions in which protein concentrations were in great excess over that of ANS and are corrected for dilution. Emission spectra were recorded from 400–550 nm with excitation at 370 nm. Buffer conditions were 20 mM MOPS, pH 7.3.

H-D exchange

One-dimensional ^{1}H NMR spectra were recorded as a function of time after dissolution of protein lyophilized from buffered aqueous solutions into ammonium acetate buffered D_2O. Raw data were transformed and phased using Felix 1.0 software (Hare, Woodinville, WA). Integration over the amide proton region was plotted against time and fit to a triple exponential decay using IgorPro software (Wavemetrics, Inc.). Random coil intrinsic exchange rates (k_3) were calculated as previously described (Englander et al., 1979; Bai et al., 1993). The structural unfolding model (Hvidt & Nielsen, 1966) was used to describe the exchange process, assuming the EX2 limit ($k_{folding} \gg k_{unfolding} \gg k_3$), so that the observed exchange rate, k_{obs}, is given by Equation 5:

$$k_{obs} = k_{unfolding} * k_3/k_{folding} = K_{op} * k_3. \quad (5)$$

The protection factor $P = k_3/k_{obs}$, or $P = 1/K_{op}$, and the free energy of opening is then given by Equation 6:

$$\Delta G_{op} = -RT \ln K_{op} = -Rt \ln(k_{obs}/k_3) \quad (6)$$

where K_{op} is the equilibrium constant for the transient opening of the structure and ΔG_{op} is the free energy difference between the folded and open states. If exchange requires global unfolding of the protein, then ΔG_{op} should correspond to the energy of global unfolding, ΔG_D, determined, for example, by chemical denaturation.

NMR spectroscopy

^{1}H NMR spectra of α_3B and α_3C were recorded with a Bruker AMX-600 spectrometer at 298 K on samples that were approximately 1 mM in 90% H_2O/10% D_2O, pH 7.1. One-dimensional spectra were recorded with 8k points, 2D TOCSY (Wuthrich, 1986) and NOESY experiments were recorded with 2k points in the F2 dimension, and 512 increments in F1. TOCSY mixing time was 61 ms, NOESY mixing times were 100 and 180 ms. NMR data were processed using the program AZARA (W. Boucher, unpubl. data) on a Silicon Graphics Indigo computer.

Note added in proof

Very recently, Johansson and coworkers described a closely related approach to the design of a single-chain three-helix bundle protein: Johansson JS, Gibney BR, Skalicky JJ, Wand AJ, Dutton PL. 1998. A native-like three-α-helix bundle protein from structure-based redesign: A novel maquette scaffold. *J Am Chem Soc 120*:3881–3886.

Acknowledgments

We thank Stephen Betz, Karyn O'Neil, Hongxing Zhou, and David Eisenberg for helpful discussions, Peter Domaille, Sharon Archer, and Scott Walsh for NMR assistance, Barbara Larsen for H-D exchange by ESI-MS, and the DuPont Merck Postdoctoral Program for support. TMH is an NSF Young Investigator and JRD is a fellow of the Jane Coffin Childs Memorial Fund for Medical Research. This work was supported by NIH grant GM54616.

References

Agashe VR, Udgaonkar JB. 1995. Thermodynamics of denaturation of barstar: Evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry 34*:3286–3299.

Albericio F, Barany G. 1987. An acid-labile anchoring linkage for solid-phase synthesis of C-terminal peptide amides under mild conditions. *Int J Peptide Protein Res 30*:206–216.

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

Axe DD, Foster NW, Fersht A. 1996. Active barnase variants with completely random hydrophobic cores. *Proc Natl Acad Sci USA 93*:5590–5594.

Bai Y, Milne JS, Mayne L, Englander SW. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct Funct Genet 17*:75–86.

Bai Y, Sosnick T, Mayne L, Englander S. 1995. Protein folding intermediates: Native-state hydrogen exchange. *Science 269*:192–197.

Becktel WJ, Schellman JA. 1987. Protein stability curves. *Biopolymers 26*:1859–1877.

Betz S, Fairman R, O'Neil K, Lear J, DeGrado W. 1995. Design of two-stranded and three-stranded coiled coil peptides. *Philos Trans R Soc London B 348*: 81–88.

Betz SF, Raleigh DP, DeGrado WF. 1993. De novo protein design: From molten globules to native-like states. *Curr Opin Struct Bio 3*:601–610.

Boice JA, Dieckmann GR, DeGrado WF, Fairman R. 1996. Thermodynamic analysis of a designed three-stranded coiled coil. *Biochemistry 35*:14480–14485.

Booth D, Sunde M, Bellotti V, Robinson C, Hutchinson W, Fraser P, Hawkins P, Dobson C, Radford S, Blake C, Pepys M. 1997. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature 385*:787–793.

Brooks IS, Soneson KK, Hensley P. 1993. Development and use of a Mac based data-analysis package for equilibrium sedimentation data from the analytical ultracentrifuge. *Biophys J 64*:A244.

Bryson JW, Betz SF, Lu HS, Suich DJ, Zhou HH, O'Neil KT, DeGrado WF. 1995. Protein design: A hierarchic approach. *Science 270*:935–941.

Cedergren L, Andersson R, Jansson B, Uhlén M, Nilsson B. 1993. Mutational analysis of the interaction between staphylococcal protein A and human IgG1. *Protein Eng 6*:441–448.

Chamberlain AK, Handel TM, Marqusee S. 1996. Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nature Struct Biol 3*:782–787.

Chen B-L, Schellman JA. 1989. Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry 28*:685–691.

Cohn EJ, Edsall JT. 1943. *Proteins, amino acids, and peptides as ions and dipolar ions*. New York: Reinhold Publishing Corp. pp 370–377.

Cordes MHJ, Davidson AR, Sauer RT. 1996. Sequence space, folding and protein design. *Curr Opin Struct Biol 6*:3–10.

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

Creighton TE. 1993. *Proteins: Structures and molecular properties*. New York: W. H. Freeman and Company.

Dahiyal BI, Mayo SL. 1997. De Novo protein design: Fully automated sequence selection. *Science 278*:82–87.

Dasgupta S, Bell JA. 1993. Design of helix ends: Amino acid preferences, hydrogen bonding and electrostatic interactions. *Int J Peptide Protein Res 41*:499–511.

DeGrado WF, Wasserman ZR, Lear JD. 1989. Protein design, a minimalist approach. *Science 243*:622–628.

Deisenhofer J. 1981. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 1.8 Å resolution. *Biochemistry 20*:2361–2370.

Desjarlais JR, Handel TM. 1995. De novo design of the hydrophobic cores of proteins. *Protein Sci 4*:2006–2018.

Efimov AV. 1991. Structure of α-α hairpins with short connections. *Protein Eng 4*:245–250.

Efimov AV. 1993. Patterns of loop regions in proteins. *Curr Opin Struct Bio 3*:379–384.

Englander JJ, Calhoun DB, Englander SW. 1979. Measurement and calibration of peptide group hydrogen-deuterium exchange by ultraviolet spectrophotometry. *Anal Biochem 92*:517–524.

Fezoui Y, Connolly PJ, Osterhout JJ. 1997. Solution structure of alpha-t-alpha, α helical hairpin peptide of de novo design. *Protein Sci 6*:1869–1877.

Freire E. 1995. Thermodynamics of partly folded intermediates in proteins. *Annu Rev Biomol Struct 24*:141–165.
Gouda H, Torigoe H, Saito A, Sato M, Arata Y, Shimada I. 1992. Three-dimensional solution structure of the B domain of staphylococcal protein A: Comparisons of the solution and crystal structures. *Biochemistry 31*:9665–9672.
Handel TM, Williams SA, DeGrado WF. 1993. Metal ion-dependent modulation of the dynamics of a designed protein. *Science 261*:879–885.
Harbury PB, Kim PS, Alber T. 1994. Crystal structure of an isoleucine-zipper trimer. *Nature 371*:80–83.
Harding SE, Rowe AJ, Horton JC. 1992. *Analytical ultracentrifugation in biochemistry and polymer science*. Cambridge: The Royal Society of Chemistry.
Harper ET, Rose GD. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry 32*:7605–7609.
Hill CP, Anderson DH, Wesson L, DeGrado WF, Eisenberg D. 1990. Crystal structure of α_1: Implications for protein design. *Science 249*:543–546.
Hvidt A, Nielsen SO. 1966. Hydrogen exchange in proteins. *Adv Protein Chem 21*:287–386.
Ilyina E, Roongta V, Mayo KH. 1997. NMR structure of a de novo designed, peptide 33mer with two distinct, compact beta-sheet folds. *Biochemistry 36*:5245–5250.
Knorr R, Trzeciak A, Bannwarth W, Gillessen D. 1989. TBTU/HOBt activation for peptide synthesis. *Tetrahedron Lett 30*:1927–1930.
Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.
Krylov D, Mikhailenko I, Vinson C. 1994. A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *EMBO J 13*:2849–2861.
Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins Struct Funct Genet 6*:87–103.
Lau SYM, Taneja AK, Hodges RS. 1984. Synthesis of a model protein of defined secondary and quaternary structure. *J Biol Chem 259*:13253–13261.
Lazaridis T, Archontis G, Karplus M. 1995. *Adv Protein Chem 47*:231–306.
Lesk AM, Chothia C. 1980. How different amino acids sequences determine similar protein structures: The structure and evolutionary dynamics of the globins. *J Mol Biol 136*:225–270.
Lovejoy B, Choe S, Cascio D, McRorie DK, DeGrado WF, Eisenberg D. 1993. Crystal structure of a synthetic triple-stranded alpha-helical bundle. *Science 259*:1288–1293.
Lumb KJ, Kim PS. 1995. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry 34*:8642–8648.
Lumb KJ, Kim PS. 1996. Interhelical salt bridges, coiled-coil stability and specificity of dimerization. *Science 271*:1137–1138.
Lyu P, Zhou HX, Jelveh N, Wemmer DE, Kallenbach NR. 1992. Position-dependent stabilizing effects in alpha-helices: N-terminal mapping in synthetic model peptides. *J Am Chem Soc 114*:6560-6562.
Lyu PC, Liff MI, Marky LA, Kallenbach NR. 1990. Side chain contributions to the stability of alpha-helical structure in peptides. *Science 250*:669–673.
Makhatadze GI, Privalov PL. 1990. Heat capacity of proteins I: Partial molar heat capacity of individual amino acid residues in aqueous solution: Hydration effect. *J Mol Biol 213*:375–384.
McGregor MJ, Islam SA, Sternberg MJE. 1987. Analysis of the relationship between side chain conformation and secondary structure in globular proteins. *J Mol Biol 198*:295–310.
Mihara H, Takahashi Y. 1997. Engineering peptides and proteins that undergo α-to-β transitions. *Curr Opin Struct Biol 7*:501–508.
Mirny LA, Shakhnovich EI. 1996. How to derive a protein folding potential? A new approach to an old problem. *J Mol Biol 264*:1164–1179.
Murphy KP, Freire E. 1992. Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv Protein Chem 43*:313–361.
Ogihara NL, Weiss MS, Degrado WF, Eisenberg D. 1997. The crystal structure of the designed trimeric coiled coil Coil-V_aL_d: Implications for engineering crystals and supramolecular assemblies. *Protein Sci 6*:80–88.
O'Neil KT, DeGrado WF. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*:646–651.
Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG. 1996. Protein folding funnels: The nature of the transition state ensemble. *Folding Design 1*:441–450.
O'Shea EK, Klemm JD, Kim PS, Alber TA. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded coiled coil. *Science 254*:539–544.
O'Shea EK, Lumb KJ, Kim PS. 1993. Peptide "Velcro": Design of a heterodimeric coiled coil. *Curr Biol 3*:658–667.
Padmanabhan S, Marqusee S, Ridgeway T, Laue TM, Baldwin RL. 1990. Relative helix forming tendencies of nonpolar amino acids. *Nature 344*:268–270.
Parry DA, Dixon RW, Cohen C. 1992. Analysis of the three-α-helix motif in the spectrin superfamily of proteins. *Biophys J 61*:858–867.
Pascual J, Pfuhl M, Rivas G, Pastore A, Saraste M. 1996. The spectrin repeat folds into a three-helix bundle in solution. *FEBS Lett 383*:201–207.
Presta LG, Rose GD. 1988. Helix signals in proteins. *Science 240*:1632–1641.
Privalov PL. 1992. Physical basis of the folded conformation of proteins. In: Creighton TE, ed. *Protein folding*. New York: W.H. Freeman and Co. pp 83–126.
Privalov PL, Makhatadze GI. 1990. Heat capacity of proteins 2. *J Mol Biol 213*:385–391.
Ptitsyn OB. 1992. The molten globule state. In: Creighton TE, ed. *Protein folding*. New York: W.H. Freeman and Co. pp 243–299.
Ptitsyn OB, Pain RH, Semisotnov GV, Zerovnik E, Razgulyaev OL. 1990. Evidence for a molten globule state as a general intermediates in protein folding. *FEBS Lett 262*:20–24.
Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of a helices. *Science 240*:1648–1652.
Richardson JS, Richardson DC. 1989. The de novo design of protein structures. *Trends Biochem Sci 14*:304–309.
Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry 27*:8063–9068.
Santoro MM, Bolen DW. 1992. A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. *Biochemistry 31*:4901–4907.
Schafmeister CE, LaPorte SL, Miercke LJW, Stroud RM. 1997. A designed four helix bundle protein with native-like structure. *Nature Struct Biol 4*:1039–1046.
Schneider JP, Lear JD, DeGrado WF. 1997. A designed buried salt bridge in a heterodimeric coiled coil. *J Am Chem Soc 119*:5742–5743.
Scholtz JM, York EJ, Stewart JM, Baldwin RL. 1991. A neutral, water-soluble, α-helical peptide: The effect of ionic strength on the helix-coil equilibrium. *J Am Chem Soc 113*:5102–5104.
Schrauber H, Eisenhaber F, Argos P. 1993. Rotamers: To be or not to be? *J Mol Biol 230*:592–612.
Semisotnov GV, Rodionova NA, Razgulyaev OI, Uversky VN, Gripas AF, Gilmanshin RI. 1991. Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers 31*:119–128.
Shalongo W, Stellwagen E. 1995. Incorporation of pairwise interactions into the Lifson-Roig model for helix prediction. *Protein Sci 4*:1161–1166.
Shaw A, McRee DE, Vacquier VD, Stout CD. 1993. The crystal structure of lysin, a fertilization protein. *Science 262*:1864–1867.
Sliz P, Engelmann R, Hengstenberg W, Pai EF. 1997. The structure of enzyme $IIA^{lactose}$ from *Lactococcus lactis* reveals a new fold and points to possible interactions of a multicomponent system. *Structure 5*:775–788.
Speicher DW, Marchesi VT. 1984. Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature 311*:177–180.
Spolar RS, Livingston JR, Record MR Jr. 1992. Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry 31*:3947–3955.
Struthers MD, Cheng RP, Imperiali B. 1996. Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science 271*:342–345.
Thompson KS, Vinson CR, Freire E. 1993. Thermodynamic characterization of the structural stability of the bZIP transcription factor GCN4. *Biochemistry 32*:5491–5496.
Uhlén M, Guss B, Nilsson B, Gatenbeck S, Philipson L, Lindberg M. 1984. Complete sequence of the staphylococcal gene encoding protein A. *J Biol Chem 259*:1695–1702.
Vetter I, Baase W, Heinz D, Xiong J, Snow S, Matthews B. 1996. Protein structural plasticity exemplified by insertion and deletion mutants in T4 lysozyme. *Protein Sci 5*:2399–2415.
Woolfson DN, Alber T. 1995. Predicting oligomerization states of coiled coils. *Protein Eng 4*:1596–1607.
Wuthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.
Zhou HX, Lyu P, Wemmer DE, Kallenbach NR. 1994. Alpha helix capping in synthetic model peptides by reciprocal side chain–main chain interactions: Evidence for an N-terminal "capping box." *Proteins Struct Funct Genet 18*:1–7.