

Metal Ion–Dependent Modulation of the Dynamics of a Designed Protein

Tracy M. Handel, Scott A. Williams, William F. DeGrado

The peptide α_4 is a designed four-helix bundle that contains a highly simplified hydrophobic core composed exclusively of leucine residues; its tertiary structure is therefore largely dictated by hydrophobic forces. This small protein adopts a structure with properties intermediate between those of the native and molten globule states of proteins: it is compact, globular, and has very stable helices, but its apolar side chains are mobile and not as well packed as in many natural proteins. To induce a more native-like state, two Zn^{2+} -binding sites were introduced into the protein, thereby replacing some of the non-specific hydrophobic interactions with more geometrically restrictive metal-ligand interactions. In the metal-bound state, this protein has properties that approach those of native proteins. Thus, hydrophobic interactions alone are sufficient to drive polypeptide chain folding nearly to completion, but specific interactions are required for a unique structure.

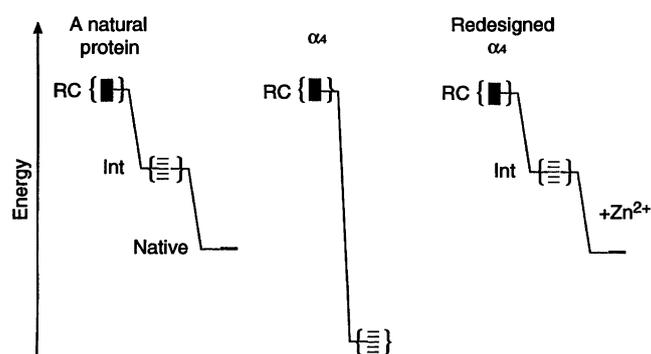
Studies of natural proteins and their variants have suggested that proteins fold via intermediates (1) stabilized by the collapse of their hydrophobic groups (2). The native structure would then form at a later state through the accrual of specific interactions including hydrogen bonding, aromatic interactions, and electrostatic interactions. Although attractive, this hypothesis is difficult to test because of the complexity of natural proteins and the transitory nature of their folding intermediates. Minimalist protein design (3) offers an alternative method to test this and related hypotheses through the design of proteins that are far simpler than natural proteins but that embody the features thought to be responsible for folding.

Among the few proteins designed in recent years (4) is a class of four-helix bundle proteins that includes $\alpha_1\text{B}$, a 16-residue peptide that forms an α -helical tetramer; α_2 , a helix-loop-helix peptide that dimerizes; and α_4 , which consists of four identical $\alpha_1\text{B}$ helices connected by three Pro-Arg-Arg loops (5). The interiors of these proteins consist only of a single hydrophobic amino acid, Leu, and the solvent-accessible faces of the helices consist of charged Lys and Glu residues. Specific side chain–side chain and H-bonded interactions were not included in the design. Thus, the simplified nature of the hydrophobic core made it possible to characterize these proteins with respect to the role of

hydrophobic collapse in protein folding.

These helical bundles had the desired aggregation states, were highly helical, and showed an extraordinarily high thermodynamic stability. Even at low pH, where the glutamates and lysines are protonated and electrostatics should be unfavorable, the assembly of $\alpha_1\text{B}$ into α -helical tetramers was highly favorable (6), indicating the importance of hydrophobic and van der Waals interactions in the folding process. However, not all the properties of these proteins were consistent with the native state: the side chain resonances were not well resolved in the nuclear magnetic resonance (NMR) spectra, and the rates of proton-deuteron exchange were between those of helices in molten globules and native proteins. Thus, these proteins appeared to have some properties associated with molten globules or folding intermediates (7).

Fig. 1. Simplified models for the folding of α_4 and natural proteins. **(Left)** The folding of a natural protein. A protein begins in a random coil configuration containing a large number of conformational states with nearly degenerate energy levels, proceeds to an intermediate state with fewer possible conformational states, and ultimately forms a native-like state. In that proteins probably fold by more complex processes that involve far more than one intermediate, the diagram could be expanded without affecting the concepts of the model. **(Center)** α_4 adopts an extremely stable intermediate-like state consisting of an ensemble of low-energy configurations. **(Right)** To make α_4 more like a natural protein it should be necessary to destabilize its intermediate-like state and to specify a more conformationally well-defined state that will obtain its stability by specific interactions, in this case the binding of Zn^{2+}



Proteins in the molten globule state are compact and have considerable secondary structure, but the interior side chains are not well packed. They do not undergo cooperative thermal transitions and, like folding intermediates, are able to bind hydrophobic dyes (7). The possibility of such nonnative or molten globule-like features in the α_4 proteins could be anticipated by considering a highly simplified folding scheme for a natural protein (Fig. 1). From a random coil configuration with many energetically accessible states, the protein folds into an intermediate state with fewer accessible conformations. Finally, upon accrual of numerous specific tertiary interactions, the protein folds into a well-defined native state. In the case of α_4 , the protein folds into an intermediate-like state stabilized by hydrophobic forces, but because it cannot form specific interactions, it never adopts a truly native state.

To test this hypothesis, we first measured the structural and dynamic properties of the α_4 proteins with NMR, the binding of 1-anilino-8-naphthalene sulfonate (ANS), and measurement of the thermodynamics of unfolding. We then redesigned α_4 so that it could adopt a more native-like state (Fig. 1) by (i) destabilizing the intermediate-like state of α_4 and (ii) introducing more specific interactions. Several apolar Leu residues were changed to metal chelating His residues, thereby replacing some of the relatively nonspecific hydrophobic interactions with more geometry-dependent metal-ligand interactions. We now compare the structural and physical properties of these metalloproteins with those of the parent α_4 protein.

Design, stability, and zinc-binding characteristics of metal-binding proteins. The introduction of three- or four-coordinate Zn^{2+} -binding sites into α_2 and α_4 has been described (8, 9). The three-coordinate Zn^{2+} -binding peptides (9) acquire an

T. M. Handel and W. F. DeGrado are with the Du Pont Merck Pharmaceutical Company, Post Office Box 80328, Wilmington, DE 19880-0328. S. A. Williams is with the Regional Laser and Biotechnology Laboratory, University of Pennsylvania, Philadelphia, PA 19104. W. F. DeGrado is also affiliated with the Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104.

interfacial site when three His residues are introduced into α_2 , two replacing a Leu and a Lys separated by one α -helical turn and the third His replacing a Leu located on a neighboring helix (Fig. 2) (10). This peptide ($H3\alpha_4$) dimerized into a four-helical bundle containing two equivalent metal binding sites. As indicated by NMR data, the Zn^{2+} -binding site in $H3\alpha_4$ is geometrically well defined (9), and all three His residues come within van der Waals contact in the presence of Zn^{2+} . We introduced one or two copies of this site into the full-length α_4 protein. We chemically synthesized $H3\alpha_4$, containing a single Zn^{2+} site, and $H6\alpha_4$, which contains one Zn^{2+} -binding site between helices 1 and 2 and a second site between helices 3 and 4 (Fig. 2) (10).

The substitution of His residues for hydrophobic Leu residues at partially buried sites in $H3\alpha_4$ and $H6\alpha_4$ destabilized the intermediate-like state of these proteins in the absence of metal ions, as indicated by denaturation with guanidium chloride (GdmCl). The thermodynamic stabilities of these proteins were less favorable than that of α_4 ; extrapolated free energies of folding (11) were -6.2 and -2.5 kcal/mol for $H3\alpha_4$ and $H6\alpha_4$, respectively (Table 1), compared to -15.4 kcal/mol for acetylated α_4 ($Ac-\alpha_4-CONH_2$, a chemically synthesized derivative of α_4 lacking the NH_2 -terminal Met of the cloned α_4 protein but containing the NH_2 -terminal acetyl and the COOH-terminal carboxamide blocking groups as in $H3\alpha_4$ and $H6\alpha_4$). However, in the presence of 1.0 mM Zn^{2+} where the Zn^{2+} -binding sites were fully saturated [K_{diss} for binding $Zn^{2+} \ll 1 \mu M$ for both proteins (12)], the stabilities of $H3\alpha_4$ and $H6\alpha_4$ were enhanced (Table 1); the extrapolated free energies of stabilization ($\Delta\Delta G^0$) were -3.5 kcal/mol for $H3\alpha_4$, and -7.8 kcal/mol for $H6\alpha_4$. Thus, by replacing some of the indiscriminate hydrophobic interactions of Leu residues with more precisely oriented Zn^{2+} -binding sites, we have succeeded in designing a protein whose stability depended on interactions with metal ions.

¹H NMR spectra of α_4 proteins compared to metal-binding proteins. The NMR spectra of native proteins show a large spread of chemical shifts induced by their well-defined tertiary structures, but molten globules or folding intermediates show much less dispersion resulting from structural mobility, which in turn leads to averaged resonance frequencies. The spectra of the downfield (amide proton) region of α_1B , α_2 , and α_4 (13) are well dispersed across approximately 1.5 ppm (Fig. 3), in agreement with earlier findings that most of the backbone forms relatively rigid α -helices, which place each of the individual amides

in a distinct environment. The broadening of the resonances as the number of loops increase is due to the loops, which make each of the helices become slightly nonequivalent in α_2 and α_4 (resulting in a slight nonequivalence in the chemical shifts of the amides at a given position in the helices). In addition, motions that are intermediate on the NMR time scale and aggregation at the high concentrations used could also contribute to resonance broadening. In contrast, the aliphatic region (Fig. 3) was less well dispersed, particularly in the methyl region where all leucine methyl groups (48 in α_4) resonate within an 0.2 to 0.3 ppm envelope centered near the random coil value. This finding is partly due to the absence of aromatic residues, but rapid averaging of the side chain conformations

could also contribute to resonance overlap. These and previous data (6), suggest that the major portion of the backbone of these four-helix bundles adopts a stable α -helical conformation, whereas the side chains undergo conformational averaging.

A similar result was observed for $H3\alpha_4$; both in the presence and absence of Zn^{2+} the spectra showed little dispersion in the aliphatic region of the spectrum (Fig. 4) (13). In contrast, $H6\alpha_4$, which binds two Zn^{2+} ions per bundle, showed a dramatic Zn^{2+} -dependent increase in the chemical shift dispersion of both the leucine methyl groups at ~ 1 ppm and the H_α resonances near 4 to 5 ppm. The general shift of resonances away from random coil values indicates that the side chains adopt more well-defined positions and that the organiz-

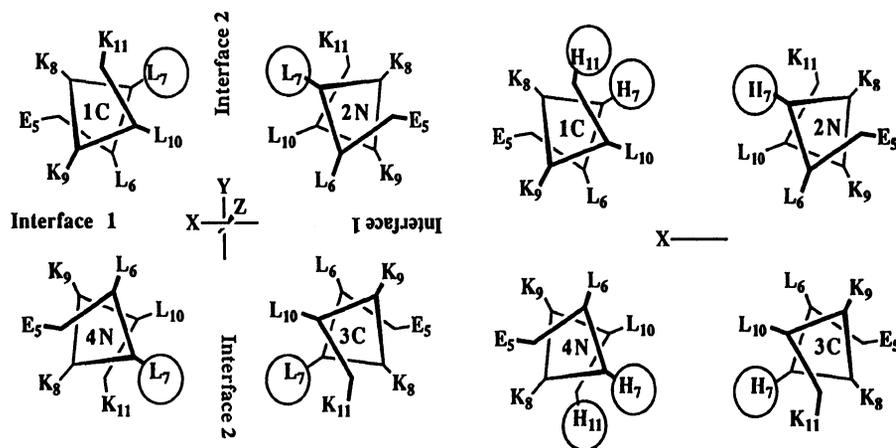


Fig. 2. An axial projection through a seven-residue section of a model of α_4 and $H6\alpha_4$. In the model of α_4 (left) the helices are arranged with 2,2,2 symmetry (having a twofold rotational symmetry axes coincident with the X, Y, and Z coordinates), and hence there are two different helix-helix packings—labeled interface 1 and interface 2. Hydrophobic Leu residues that are changed to His residues in $H6\alpha_4$ are circled. The arrangement of the His residues breaks the 2,2,2 symmetry in $H6\alpha_4$ (right) and there is now only a single twofold symmetry axis. $H6\alpha_4$ contains a Zn^{2+} site between helices 1 and 2 and a second site between helices 3 and 4 and has the sequence: $Ac-G(ELEELHKKLHELLK)GPRRG(ELEELHKKLKELLK)GPRRG(ELEELHKKLKELLK)GPRRG(ELEELHKKLHELLK)G-CONH_2$, where the sequences expected to be helical are in parentheses and the residues expected to ligate Zn^{2+} are italicized. $H3\alpha_4$ contains a Zn^{2+} -binding site between helices 1 and 4, and has the sequence: $Ac-G(ELEELHKKLKELLK)GPRRG(ELEELKKLKELLK)GPRRG(ELEELKKLKELLK)GPRRG(ELEELHKKLHELLK)G-CONH_2$.

Table 1. Summary of GdmCl denaturation curves for $H6\alpha_4$, $H3\alpha_4$, and $Ac-\alpha_4-CONH_2$. The ellipticity at 222 nm was recorded (AVIV 62DS CD) as a function of GdmCl (11). Samples were approximately $10 \mu M$ in 10 mM Hepes pH 7.0, with or without 1.0 mM zinc chloride. The data were fit to the equation $\Delta G_{obs} = \Delta G_{H_2O} - RT\Delta\beta^0[GdmCl]$ where ΔG_{obs} is $-RT \ln(K)$ (K is the equilibrium constant for protein folding at a given GdmCl), ΔG_{H_2O} is the free energy of folding in the absence of denaturant, and $RT\Delta\beta^0$ is the change in the molar cosolvation free energy (which is a measure of the cooperativity of the transition). $\Delta\Delta G$ is the difference in the free energy of folding in the presence or absence of metal for the metal binding proteins. The actual form of the equation used to fit the parameters is given in (11).

| Protein | Midpoint (M) | $RT\Delta\beta^0$ (kcal/mol)* | ΔG_{H_2O} (kcal/mol) | $\Delta\Delta G$ (kcal/mol) |
|------------------------|--------------|-------------------------------|------------------------------|-----------------------------|
| $Ac-\alpha_4-CONH_2$ | 6.3 | 2.6 | -15.4 | |
| $H3\alpha_4 - Zn^{2+}$ | 4.6 | 1.36 | -6.21 | |
| $H3\alpha_4 + Zn^{2+}$ | 5.5 | 1.76 | -9.74 | -3.5 |
| $H6\alpha_4 - Zn^{2+}$ | 3.35 | 0.76 | -2.50 | |
| $H6\alpha_4 + Zn^{2+}$ | 5.5 | 1.89 | -10.3 | -7.8 |

*Per molar GdmCl.

ing effect of metal ligation extends beyond the metal-binding site.

In the absence of Zn^{2+} , the nonexchangeable imidazole protons from the His residues of $H3\alpha_4$ and $H6\alpha_4$ are overlapped (Fig. 4) and appear at chemical shifts typical of the δ and ϵ protons of histidine in a random coil conformation (14). This suggests that the His residues undergo rapid interconversion between multiple conformers. Addition of a single equivalent of Zn^{2+} to $H6\alpha_4$ results in a shift of the resonances away from the random coil positions (Fig. 4); the chemical shifts of the six doublets are nearly identical to those reported for the six resonances in the $H3\alpha_2$ dimer (9). The doubling of resonances in $H6\alpha_4$ reflects the slight nonequivalence of the two sites. Together with earlier data, these findings indicate that $H6\alpha_4$ forms two distinct three-His Zn^{2+} -binding sites. However, addition of Zn^{2+} to $H3\alpha_4$, which has a single Zn^{2+} -binding site, causes the formation of multiple, broad peaks; six appear at the characteristic chemical shifts of $H6\alpha_4$ suggesting there is a population of three-His Zn^{2+} -

binding sites, but the additional resonances must arise from alternative conformations, different modes of binding metal, or aggregation at the high concentrations used for NMR.

Fluorescence. ANS binds to apolar regions of molten globule proteins with dissociation constants in the range of 0.01 to 1.0 mM, presumably because their hydrophobic interiors are not as well packed as in native proteins (15). It does not bind to unfolded proteins or to native proteins unless the protein has a pocket for binding hydrophobic substrates or cofactors (15, 16). Like molten globule proteins, α_1B , α_2 , and α_4 bind ANS (17) with apparent $K_{diss} = 600 \mu M$, $80 \mu M$, and $50 \mu M$ (± 15 percent), respectively. In the case of α_4 , a part of the binding energy appears to arise from favorable electrostatic interactions between positively charged side chains (net charge = +7 at pH 7) and the negatively charged sulfonate group of ANS since 0.20 M NaCl causes an increase in the dissociation constant from 50 to 110 μM . However, two-dimensional NOESY (nuclear

Overhauser effect spectroscopy) spectra of complexes of ANS with α_1B or α_4 indicate that the most apolar half of ANS contacts the methyl groups of several Leu residues confirming that ANS resides mainly in the apolar core of the bundle. Also $H3\alpha_4$ and $H6\alpha_4$ bind ANS in the absence and the presence of Zn^{2+} ; $H3\alpha_4$ binds in a Zn^{2+} -independent manner ($K_{diss} = 70 \mu M$ in the presence and absence of Zn^{2+}), whereas $H6\alpha_4$ binds more tightly in the presence of Zn^{2+} ($K_{diss} = 80 \mu M$) versus in the absence of Zn^{2+} ($K_{diss} = 230 \mu M$) (17).

The emission spectrum of ANS is sensitive to solvent polarity, shifting from 510 nm in water, to 480 nm in methanol, to 460 nm in pentanol (18); binding to α_1B , α_2 , and α_4 , shifts the emission maximum to 461 nm. By comparison, λ_{max} of ANS is 454 nm when bound to native apomyoglobin (16) and between 470 and 492 nm when bound to various molten globule proteins (15). Thus, the polarity of the hydrophobic core of the α_4 proteins appears to be intermediate between a molten globule protein and apomyoglobin. However, ANS binds to these bundle proteins in several distinct orientations, as shown below, and the measurement represents an average. The emission maxima for $H3\alpha_4$ in the absence and presence of Zn^{2+} are 458 nm and 456 nm, respectively; for $H6\alpha_4$ the corresponding values are 459 nm and 455 nm. Thus, it appears that ANS resides in a more apolar environment when bound to apo- $H3\alpha_4$ or apo- $H6\alpha_4$, despite the fact that these proteins are more polar than the α_4 . Addition of Zn^{2+} shifts the emission maxima to values characteristic of apomyoglobin, an indication of a more native-like environment for the metal-binding proteins.

Although these proteins, like molten globules, bind ANS, binding alone is not a criterion for defining the molten globule state since native proteins such as apomyoglobin bind this probe (16). Additional insight into the structural and dynamic properties of the protein to which it is bound can be obtained from quantitative data concerning the isotropic and anisotropic fluorescence decay of the complexed dye. The isotropic fluorescence lifetime of ANS depends on environment and provides additional information about environment polarity. In water, methanol, and pentanol, the fluorescence decays as a single exponential with lifetimes of 0.25, 5.9, and 11.6 ns, respectively (18). Three exponentials are required to describe the decay of ANS when bound to the α_1B , α_2 , or α_4 , suggesting that the ANS occupies multiple binding sites in these proteins. From 15 to 20 percent of the signal has a 0.4-ns lifetime indicating that this fraction of ANS is well hydrated and loosely associated with

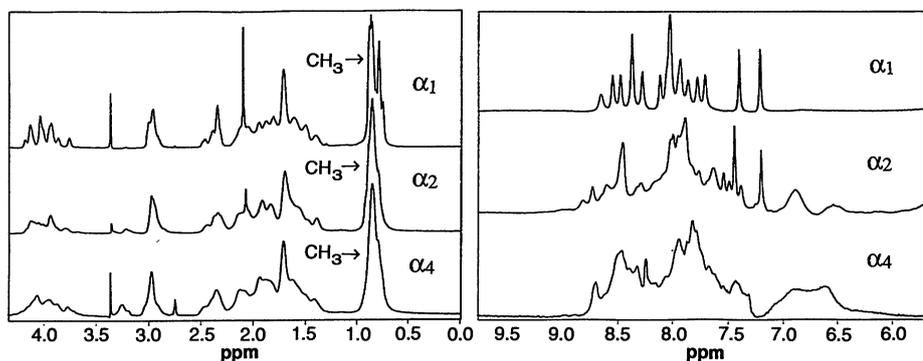
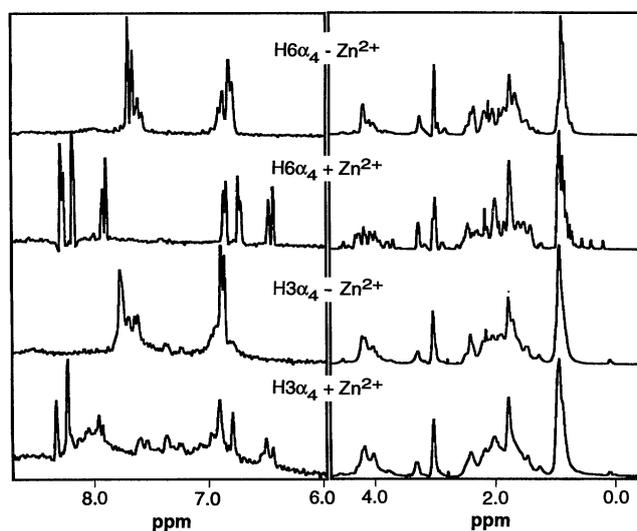


Fig. 3. Proton NMR spectra of α_1B , α_2 , and α_4 : (left) upfield (aliphatic) region; (right) downfield (amide) region (13). The sharp resonances near 7.2 and 7.4 in α_1B and α_2 are due to the carboxamide ($CONH_2$ protons) and the broad features near 6.5 to 7.0 are due to resonances from the Arg residues in the loops. Samples were prepared in 10 percent D_2O , pH 5.0.

Fig. 4. The 1H -NMR spectrum of $H6\alpha_4$ and $H3\alpha_4$ in the presence and absence of Zn^{2+} . (Left) The δ - and ϵ -CH protons of the His residues in the downfield (aromatic) region of the spectra. (Right) The upfield (aliphatic) region of the spectra. The samples were prepared at a concentration of approximately 0.5 mM in 50 mM sodium acetate- d_3 , D_2O , pH 7 (uncorrected meter reading), with and without 1.1 equivalents of zinc nitrate (13).



the protein surface. The remaining two components represent 30 and 50 percent of the signal and have lifetimes of approximately 3- and 12-ns, respectively. Thus, a large portion of the ANS lies within a highly hydrophobic environment like that of apomyoglobin, while other components have surface and interfacial localizations. By comparison, apomyoglobin shows no short-lived component, 11 percent with a 2-ns lifetime and 89 percent with a 14-ns lifetime. Without Zn^{2+} , the isotropic fluorescence decay of $H3\alpha_4$ and $H6\alpha_4$ is similar to that of α_4 . Addition of Zn^{2+} to $H3\alpha_4$ induces only a slight change in isotropic decay. For $H6\alpha_4$, however, the presence of Zn^{2+} causes ANS to populate a more hydrophobic or structured environment, as the components with the two short lifetimes become smaller and the component with the longer lifetime becomes larger (reaching 73 percent of the total compared to 89 percent for apomyoglobin) (19).

The anisotropic fluorescence decay of bound ANS provides information concerning the hydrodynamic properties of the protein, and the rigidity with which the dye is held in its binding site or sites. Two components were observed for the anisotropic decay of ANS bound to α_1B , α_2 , and α_4 . One component ($\Phi_s \sim 0.6$ to 0.9 ns) is very short, indicating ANS undergoes limited but rapid positional fluctuations within its binding sites (Table 2) while the remainder of the polarization (Φ_l) correlates with the overall rotation of the protein (20). The associated order parameter S , which has limiting values of 0.0 and 1.0 for flexible and rigid environments, respectively, ranged between 0.63 and 0.69. By comparison, the probe was highly immobilized within the heme-binding pocket of apomyoglobin as evidenced by an S value approaching 1.0 (Table 2).

The anisotropic fluorescence decay of ANS bound to $H6\alpha_4$ showed the surprising result that, even in the absence of metal ions, ANS is bound more rigidly to $H6\alpha_4$ than to α_4 as assessed by an order parameter of 0.88 (Table 2). Apo- $H3\alpha_4$ with only three His residues has an order parameter of 0.81, intermediate between α_4 and apo- $H6\alpha_4$. This is interesting in light of the large thermodynamic destabilization of the apoproteins relative to α_4 (Table 1). The high order parameter of apo- $H6\alpha_4$ can be attributed to this protein having the greatest number of hydrophilic residues at partially buried positions, which would lead to fewer energetically favorable helix-helix packings, and hence, to less dynamical averaging. In the presence of Zn^{2+} , ANS is held approximately as rigidly in $H6\alpha_4$ as in native apomyoglobin ($S = 0.90$). Furthermore, the addition of metal causes a reduction in the hydrodynamic radius of the

protein as reflected by a 5.8-ns rotational correlation time compared to 6.5 ns for the apoprotein (Table 2). Thus, Zn^{2+} induces a native-like structure in $H6\alpha_4$, which is significantly more compact than the apo form or α_4 (20).

Thermal denaturation. We used thermal denaturation to examine the nature and cooperativity of the interactions stabilizing the proteins. Although the small size and high stabilities makes their unfolding transition relatively noncooperative (21), at high concentrations of GdmCl (needed

to partially destabilize the proteins so that thermal transitions can be observed), both high- and low-temperature unfolding was observed (22). The plot of the equilibrium constant for protein unfolding (K) as a function of temperature for α_4 in 6.0 M GdmCl and T4 lysozyme (lzm) in 3.0 M GdmCl (22, 23) (Fig. 5) show parabolic shapes similar to native proteins, and exhibit both low- and high-temperature unfolding transitions. Expressed on a per residue basis, ΔH_g and ΔS_g for the low- and high-temperature transitions are similar for

Table 2. Time-dependent anisotropic fluorescence decay analysis of ANS/protein complexes (18–20). One short and one long component (Φ_s and Φ_l) were resolved. The order parameter, S , represents the fraction of polarization lost in the slow process relative to fast motions. Standard errors are given in parenthesis. $R(O) = 0.30$, the limiting anisotropy extrapolated to $t = 0$ (time).

| Protein | S | Component 1 | | Component 2 |
|------------------------|-------------|---------------|--|---------------|
| | | Φ_s (ns) | | Φ_l (ns) |
| α_1B | 0.63 (0.08) | 0.68 (0.03) | | 5.7 (0.68) |
| α_2 | 0.69 (0.07) | 0.90 (0.03) | | 6.5 (0.3) |
| α_4 | 0.68 (0.07) | 0.60 (0.06) | | 9.2 (0.8) |
| Apomyoglobin | 0.92 (0.07) | 0.42 (0.30) | | 8.5 (0.2) |
| $H3\alpha_4 - Zn^{2+}$ | 0.81 (0.01) | 1.7 (0.3) | | 9.4 (0.2) |
| $H3\alpha_4 + Zn^{2+}$ | * | * | | * |
| $H6\alpha_4 - Zn^{2+}$ | 0.88 (0.02) | 1.4 (0.3) | | 6.5 (0.2) |
| $H6\alpha_4 + Zn^{2+}$ | 0.90 (0.03) | 1.9 (0.8) | | 5.8 (0.2) |

*The protein showed a distribution of at least four correlation times (20).

Fig. 5. Temperature-dependent unfolding curves for α_4 in 6.0 M GdmCl, $H6\alpha_4$ in 5.25 M GdmCl, and T4 lysozyme in 3.0 M GdmCl; K is the equilibrium constant for unfolding (22). Curve fits to the data were based on the method of Shellman and are represented by (●); actual data are shown as (x). The curve for lysozyme was reproduced from the thermodynamic parameters in (23). For $H6\alpha_4$, it was not possible to accurately fit the data, presumably because of the presence of multiple equilibria due to high-temperature dissociation of the metal.

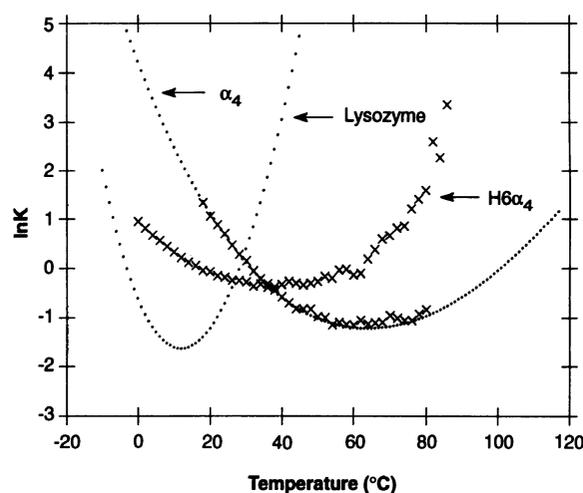


Table 3. Thermodynamic parameters for the thermal unfolding of α_4 and T4 lysozyme (22, 23). The subscript g (ΔH_g) indicates that parameters were evaluated at the midpoints of the thermal denaturation curves where $\Delta G = 0$. Quantitatively similar data was found for α_2 as for α_4 .

| Condition | T (°C) | ΔC_p (cal mol ⁻¹ K ⁻¹ ·res ⁻¹) | ΔH_g (cal mol ⁻¹) | ΔS_g (cal mol ⁻¹ deg ⁻¹) | ΔG cal mol ⁻¹ |
|-----------------------------------|----------|--|--|---|-------------------------------------|
| <i>T4 lysozyme in 3.0 M GdmCl</i> | | | | | |
| Low-temperature transition | -3 | 11.2 | -199 | -0.74 | 0 |
| High-temperature transition | 28 | 11.2 | 212 | 0.70 | 0 |
| Temperature of maximum stability | 12.5 | | | | 926 |
| <i>alpha_4 in 6.0 M GdmCl</i> | | | | | |
| Low-temperature transition | 31 | 5.3 | -192 | -0.63 | 0 |
| High-temperature transition | 98 | 5.3 | 223 | 0.60 | 0 |
| Temperature of maximum stability | 62.5 | | | | 677 |

α_4 and I_{zm} (Table 3). However, α_4 is predicted to have a melting temperature of 98°C in 6.0 M GdmCl, demonstrating that it is a very stable protein.

A major difference in the unfolding transitions of I_{zm} and α_4 is in the cooperativity, which is related in part to ΔC_p . This parameter is believed to be a measure of the difference in the exposure of apolar side chains in the folded and unfolded states to water, and is relatively constant for single domain proteins (10 to 15 mol⁻¹ K⁻¹ per residue) (21). For many molten globules, it is believed to be near zero (7). ΔC_p

for α_4 (Table 3) (and also α_2) is 2.5 times lower than for native proteins, although it is much greater than for many molten globules. A strict interpretation on this parameter is not possible because it was obtained for α_4 at high concentrations of GdmCl, and the relation between ΔC_p and the concentration of GdmCl has not been established. Nevertheless, the value of ΔC_p determined for T4 lysozyme in 3.0 M GdmCl (Table 3) is close to the value expected for native proteins (23, 24) and that for α_2 in 4.4 M GdmCl is nearly identical to ΔC_p for α_4 in 6.0 M GdmCl,

suggesting that the effect of GdmCl will not be large enough to affect the qualitative conclusion that ΔC_p for α_4 is about half that expected for a native protein.

The unfolding of molten globules is a largely noncooperative process (7), regardless of whether the denaturant is GdmCl or temperature. The cooperativity of the thermal unfolding of α_4 is intermediate between that expected for the native and molten globule states, whereas the denaturant-induced unfolding is highly cooperative. One explanation for this difference is that the cooperativity of GdmCl denaturation is a measure of change in the accessibility of the folded compared to the unfolded form of this protein to the guanidinium ion (4 to 5 Å), and ΔC_p is a measure of accessibility to water, which is a much smaller probe. Thus if the unfolded state is a true random coil and fully solvated (a likely possibility with the 4.0 to 8.0 M GdmCl used with thermal denaturation), then the difference probably lies in the greater accessibility of the unfolded state to water.

We assessed the unfolding of metal-bound $H6\alpha_4$ in the presence of Co^{2+} because the Zn^{2+} complex showed unfavorable solubility at high temperatures. The unfolding was somewhat more cooperative than α_4 , especially at high temperature where the cooperativity was that expected for a protein of this size. However, the data were difficult to fit to a single two-state transition presumably because the metal dissociates at high temperature.

Control of global topology. When viewed down their helical axes, four-helix bundle proteins can be considered as clockwise or counterclockwise (25) (Fig. 6). The His ligands composing the metal-binding sites in $H6\alpha_4$ have been so arranged that they serve as probes of the topology of the protein. If the bundle is clockwise, we obtain two, nearly equivalent three-His Zn^{2+} -binding sites as envisioned in the design (Fig. 6, row 2, left); but in a counterclockwise winding, one two-His and one four-His site would be formed (Fig. 6, row 2, right). To show that the counterclockwise form would give rise to sites spectroscopically distinct from the proposed three-His site, we prepared derivatives of α_2 that either contained a two-His (9) or a four-His Zn^{2+} -binding site (Fig. 6, row 1) similar to those that would be formed if $H6\alpha_4$ wound in a counterclockwise manner. Circular dichroism (CD) and size exclusion chromatography indicate that these peptides dimerize to form four-helix bundles in the absence and presence of Zn^{2+} . NMR shows that these proteins bind Zn^{2+} , but the spectra in the aromatic region of the two-His (9) and the four-His peptides (Fig. 6) were complex and consistent with the formation of numerous species that intercon-

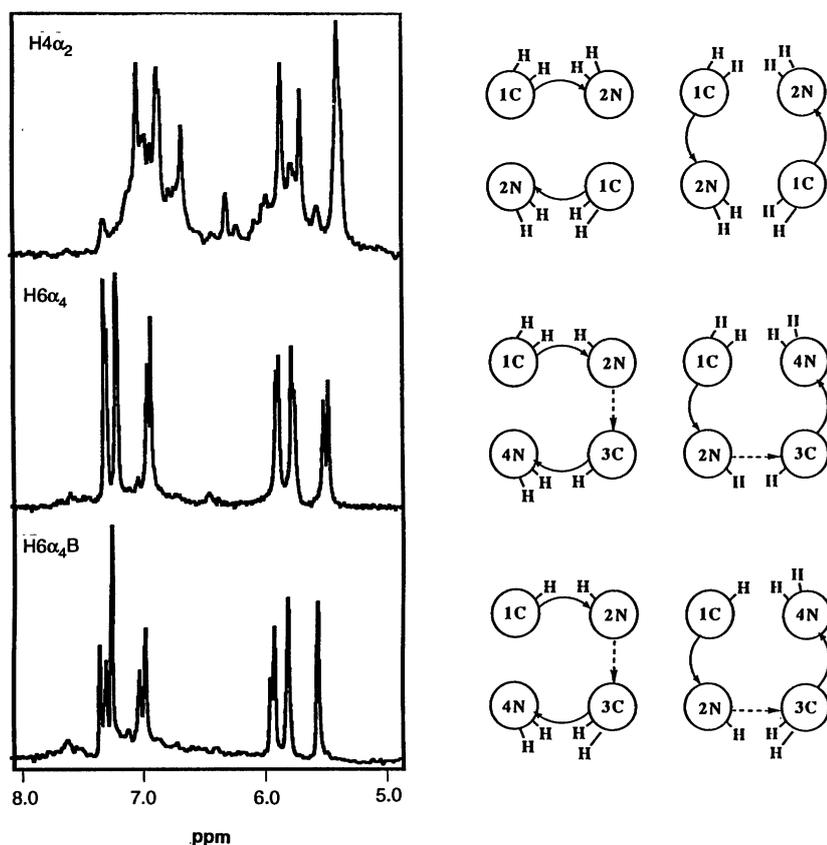


Fig. 6. (Left) The δ - and ϵ -CH protons of the His residues in the downfield region of the spectra of $H4\alpha_2$ (top), $H6\alpha_4$ (center), and $H6\alpha_4B$ (bottom). The samples were prepared in 50 mM sodium acetate- d_3 , D_2O , pH 7 (uncorrected), with 1.1 equivalents of zinc nitrate. (Right) A highly schematic drawing of the two possible topologies of $H4\alpha_2$ (top), $H6\alpha_4$ (center), and $H6\alpha_4B$ (bottom). The helices are numbered and their NH_2 - and $COOH$ -terminal ends are designated as N and C. The schematic on the left (center row) shows the clockwise orientation of $H6\alpha_4$, which orients the His residues appropriately for two three-His metal-binding sites. The right panel shows that a counterclockwise orientation of the helices produces a four-His binding site and a two-His binding site, both of which would produce complicated NMR spectra on the basis of model compounds such as $H4\alpha_2$. The bottom row shows the clockwise (left) and counterclockwise orientation of $H6\alpha_4B$ (which contains three-His Zn^{2+} -binding sites between helices 1 and 4, and between helices 2 and 3). In contrast to $H6\alpha_4$, a counterclockwise orientation of the helices in $H6\alpha_4B$ properly orients the His residues for two three-His binding sites. The top schematics shows two possible orientations of $H4\alpha_2$ [sequence: Ac-G(ELEELHAKKLHELLK)GPRRG(ELEELHAKKLHELLK)G-CONH₂], both of which produce four-His binding sites. The NMR spectra to the left illustrates the complex appearance of this type of metal binding site. The simplicity of the NMR spectra of $H6\alpha_4$ and $H6\alpha_4B$, which show one resonance for each of the 12 His protons, is characteristic of the three-His site (9) and suggests that the clockwise orientation (center, left) is the correct topology for $H6\alpha_4$, and the counterclockwise topology (bottom right) is correct for $H6\alpha_4B$. Monte Carlo lattice simulations of the folding of α_4 indicate that it can fold in either a clockwise or counterclockwise topology (27), consistent with the above results.

vert at slow rates on the millisecond time scale. The simplicity of the spectrum of H6 α_4 indicates that this protein has a single clockwise topology (Fig. 6).

The finding that H6 α_4 forms a protein with a clockwise topology does not necessarily imply that the parent α_4 protein has a single, preferred clockwise topology. In fact, we have prepared a derivative of H6 α_4 (H6 α_4 B) with a Zn²⁺-binding site between helices 1 and 4, and a second Zn²⁺-binding site between helices 2 and 3 (26). This protein should form the desired three-His binding site only if the protein adopts a counterclockwise topology (Fig. 6, bottom row). This protein also binds Zn²⁺, and the NMR and thermodynamic properties of the adduct are similar to those of H6 α_4 (26). Thus, the loops in α_4 can adopt left-handed and right-handed crossovers, and may exist as a mixture of two topological species.

Origins of the thermodynamic stability and the dynamic properties of α_4 . Although the tertiary design of α_4 was based exclusively on van der Waals and hydrophobic interactions, many of the properties of α_4 are as envisioned in its design and similar to properties of native proteins: It folds into a compact structure consisting of flexible loops and helices of approximately the desired length, it has a very cooperative GdmCl-induced unfolding transition, and it has low and high temperature unfolding transitions with reasonably normal enthalpies and entropies. However, α_4 also shows many properties generally associated with molten globules or protein folding intermediates, such as the binding of ANS, very poor dispersion of side chain resonances in its NMR spectra, and a value of ΔC_p that is intermediate between those expected for a native and a molten globule protein. Further, to the extent that H6 α_4 mimics α_4 , it seems likely that α_4 exists in at least two folding topologies, including left-handed and right handed bundles; this conclusion is supported by computer simulations which indicate the two topologies are isoenergetic (27). Thus, α_4 has a structure which is more mobile than the native state, but also more stable and structured than many folding intermediates or molten globules.

Our data support the hypothesis that hydrophobic interactions play an important role in collapsing an unfolded protein into a state that is close to the native state. Various experiments have suggested that one of the initial steps in folding is the formation of α helices and β sheets (1, 28) that derive their stabilities through clustering of apolar side chains in the interior of the protein (2). Hydrophobic forces appear ideal for holding the protein together at this stage—this force is long range and relatively insensitive to the geometries of the interacting groups. Thus, a protein might be

flexible at this stage to facilitate the search for and formation of specific H-bonded and electrostatic interactions between groups that are far apart in sequence. Such interactions would lock the protein into its native conformation, and indeed frequently form late in the folding process (28). The design of α_4 did not include any specific interhelical hydrogen bonds or electrostatic interactions, explaining much of its intermediate-like behavior.

Paradoxically the lack of these specific interactions might account for the extreme stability of α_4 . In native proteins a number of polar residues become partially desolvated in the process of forming a hydrophobic core. Such polar residues generally form hydrogen-bonded and electrostatic interactions that maintain the specific conformations of folded proteins. However, these favorable specific interactions are made at the expense of the unfavorable dehydration of the side chains and are often not energetically favorable (29). Because of its idealized design, α_4 can presumably fold without fully burying any polar side chains, leading to a protein with extreme stability but also with a conformation that is more dynamic and less precisely structured than the native states of most proteins.

The flexibility of Leu, whose side chain has a large number of low-energy rotamers (30), may also contribute to the stability of α_4 . Monte Carlo simulations indicate that the persistence of contacts between pairs of Leu side chains last for only very short times during simulations (27). In the original design of α_4 , we assumed that a single low-energy conformer would predominate for each side chain. Instead we suggest that the entropic advantage associated with equilibrating conformers provides additional thermodynamic stability to the protein. Thus, the sequence of α_4 does not contain sufficient specific interactions to specify the tight packing of side chains that is the hallmark of native structure. Other designed proteins (31–33) probably also adopt molten globule-like states, although none of these have been characterized in great detail. Our work and related studies (34) now show that it is possible to define more native-like structures through the introduction of specific interactions.

Unlike α_4 , H6 α_4 in the metal-bound form appears to adopt a single topology, shows considerable dispersion in the NMR, and binds ANS in a relatively rigid manner, more similar to a native protein than a molten globule. Furthermore, apo-H6 α_4 with four His residues substituted for Leu, behaved more like a native protein than α_4 ; it bound ANS more rigidly and had a smaller hydrodynamic radius. A likely explanation is that H6 α_4 has fewer low energy packings than α_4 ; some helix-helix pack-

ings that are favorable in α_4 might lead to the burial of the His imidazole group in apo-H6 α_4 . A similar rationale has been forwarded to explain the occurrence of a conserved Asn in GCN4 (35).

Structurally, H6 α_4 is about as complex as some simple Ca²⁺-binding proteins such as calbindin D_{9k} (36) or the individual domains of calmodulin (37). Both H6 α_4 and these Ca²⁺-binding proteins consist of two highly homologous bihelical metal-binding domains related by an approximately twofold symmetry axis, both have flexible loops connecting the bihelical motifs, and both bind metal ions with similar dissociation constants. Further, like calbindin, H6 α_4 undergoes a major decrease in flexibility upon binding to Zn²⁺ (38). Although this feature was not intentionally designed, H6 α_4 behaves like calmodulin's individual domains in that it binds ANS more tightly in the metal-bound form than in the apo-form (39). Thus, although several designed proteins have been described (5, 8, 31, 32, 40), H6 α_4 is alone in that extensive physical methods have been applied to show that it has ligand binding and structural properties approaching those of a native protein. Nevertheless, H6 α_4 retains some intermediate-like features including the binding of ANS, and amide proton-deuteron exchange rates that are somewhat fast when compared to many natural proteins. These features may be a consequence of the fact that the metal-ligand interactions constrain only one of the two distinct helix-helix interfaces present in the α_4 dimer (Fig. 2, interface 2). Introduction of specific interactions into the remaining interface (Fig. 2, interface 1), which is still composed of Leu side chains, should result in a protein with fully native properties.

REFERENCES AND NOTES

1. R. L. Baldwin and H. Roder, *Curr. Biol.* **1**, 218 (1991); M. Bycroft, A. Matouschek, J. T. Kellis, L. Serrano, A. Fersht, *Nature* **346**, 488 (1990); J. B. Udgaonkar and R. L. Baldwin, *ibid.* **335**, 694 (1988); H. Roder, G. A. Elöve, W. S. Englander, *ibid.*, p. 700; M. S. Briggs and H. Roder, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2017 (1992); J. Lu and F. W. Dahlquist, *Biochemistry* **31**, 4749 (1992); S. E. Radford, C. M. Dobson, P. A. Evans, *Nature* **358**, 302 (1992); M. F. Jeng, S. W. Englander, G. A. Elöve, A. J. Wand, H. Roder, *Biochemistry* **29**, 10433 (1990); F. M. Hughson, P. E. Wright, R. L. Baldwin, *Science* **249**, 1544 (1990).
2. K. Dill and D. Shortle, *Annu. Rev. Biochem.* **60**, 795 (1991); K. Dill, *Biochemistry* **29**, 7133 (1990); O. B. Ptitsyn, *J. Prot. Chem.* **6**, 273 (1987).
3. W. F. DeGrado, Z. R. Wasserman, J. D. Lear, *Science* **243**, 622 (1989); E. T. Kaiser and F. J. Kezdy, *Proc. Natl. Acad. Sci. U.S.A.* **223**, 249 (1984); S. F. Michael, V. J. Kilfoil, M. H. Schmidt, B. T. Amann, J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4796 (1992).
4. W. F. DeGrado, D. P. Raleigh, T. M. Handel, *Current Opinions Struct. Biol.* **1**, 984 (1991).
5. W. F. DeGrado, L. Regan, S. P. Ho, *Cold Spring Harbor Symp. Quant. Biol.* **52**, 521 (1987); S. P. Ho and W. F. DeGrado, *J. Am. Chem. Soc.* **109**, 66751 (1987); L. Regan and W. F. DeGrado,

- Science* **241**, 76 (1988). The sequences of these designed proteins are as follows: the helices are composed of the sequence GELEELKLLKELK-LKG, the loops consist of three residues, PRR. α_1 B = Ac-helix-CONH₂; α_2 = Ac-helix-loop-helix-CONH₂; α_4 = Met-helix-loop-helix-loop-helix (recombinant protein) or Ac-helix-loop-helix-loop-helix-loop-helix-CONH₂ (the chemically synthesized protein).
- Osterhout *et al.*, *J. Am. Chem. Soc.* **114**, 331 (1992).
 - Kuwajima, *Proteins* **6**, 87 (1989); M. Ikeguchi, K. Kuwajima, M. Mitani, S. Sugai, *Biochemistry* **25**, 6965 (1986); O. B. Ptitsyn, R. H. Pain, G. V. Semisotnov, E. Zerovnik, O. L. Razgulyaev, *FEBS Lett.* **262**, 20 (1990).
 - Regan and N. D. Clarke, *Biochemistry* **29**, 10878 (1990).
 - Handel and W. F. DeGrado, *J. Am. Chem. Soc.* **112**, 6710 (1990).
 - The peptides were synthesized (Milligen 9050) by solid phase methods with Fmoc (9-fluorenylmethoxycarbonyl)-protected amino acids [Fmoc-L-Arg(PMC)-OH], and Fmoc L-amino acid pentafluorophenyl (OPfp) esters (Fmoc-L-Glu(Ot Bu)-OPfp, Fmoc-L-Leu-OPfp, Fmoc-L-His(Boc)-OPfp, Fmoc-L-Lys(Boc)-OPfp, Fmoc-L-Gly-OPfp, and Fmoc-L-Pro-OPfp) (Milligen). Synthetic cycles included 45-minute coupling reactions with 10-minute capping steps in 0.5 M pyridine and 0.5 M acetic anhydride in dimethylformamide (Gly, Glu, Leu, Lys). The Arg, His, and Pro residues were not capped but were double-coupled for an additional 45 minutes. The peptides were cleaved for 2 hours in a mixture of trifluoroacetic acid (TFA), anisole, thioanisole, and ethanedithiol (9:0.2:0.5:0.3), and subsequently precipitated with cold ether. The peptides were purified by reversed-phase HPLC on a preparative Vydac C4 column (gradient, acetonitrile containing 0.1 percent TFA). The products were at least 95 percent pure by analytical HPLC and electrospray mass spectrometry.
 - M. M. Santoro and D. W. Bolen, *Biochemistry* **27**, 8063 (1988).
 - The Zn²⁺-binding proteins were stabilized toward guanidine denaturation in the presence of Zn²⁺ because this ion binds more tightly to the native state than the denatured state. The degree of stabilization increases with increasing Zn²⁺ concentration up to 0.1 to 0.3 mM after which it levels off. At this high concentration, Zn²⁺ is bound both to the native as well as the unfolded state and the degree of stabilization is determined by the differences in affinities of the two states for metal ion according to the equation $\Delta\Delta G^\circ = -RT \ln(K_f/K_u)$, where K_f and K_u represent the dissociation constants for binding Zn²⁺ to the folded and the unfolded states, respectively. This equation was derived from an analysis of a thermodynamic cycle comparable to the derivation of equation 1 of J. Sancho and A. R. Fersht [*Biochemistry* **31**, 2253 (1992)]. Because the stabilization levels off near 0.1 mM, it is clear that K_u is approximately 0.1 mM. Thus, K_f is calculated to be 0.3 μ M for H3 α_4 , on the basis of the experimentally determined value of the free energy of stabilization, $\Delta\Delta G^\circ = -3.5$ kcal/mol at 1.0 mM Zn²⁺ (Table 1). Similarly, $\Delta\Delta G^\circ$ for H6 α_4 is -7.8 kcal/mol and the mean dissociation constant for the two sites is about 0.2 μ M.
 - Spectra were recorded (Bruker AM-600) at 22°C with presaturation of the water resonance; 2048 complex points were collected and processed with a phase shifted sine bell.
 - Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
 - G. V. Semisotnov *et al.*, *Biopolymers* **31**, 119 (1991); P. M. Mulqueen and M. J. Kronman, *Arch. Biochem. Biophys.* **215**, 28 (1982); Y. Goto and A. L. Fink, *Biochemistry* **28**, 945 (1989).
 - L. Stryer, *J. Mol. Biol.* **13**, 482 (1965).
 - Fluorescence spectra were measured with a Spex Fluorolog fluorometer. For calculation of K_{diss} , peptides were titrated into 10 μ M 1-anilino-8-naphthalene sulfonate (ANS) (10 mM Hepes, pH 7.5) in a 1-cm quartz cuvette with an excitation wavelength of 370 nm and an emission wavelength of 460 nm. The data were fit (16) with $n = 2$ (n = the number of dye molecules per four-helix bundle), based on the stoichiometry determined from NMR via titration of α_4 into ANS (0.21 mM in D₂O, pH 7.0).
 - For experiments with pure solvents, ANS was dissolved in water, or the appropriate alcohol. For ANS-protein complexes, the concentration of protein was 1 mM α_1 B, 1 mM α_2 , 120 μ M α_4 , and 150 μ M apomyoglobin (Sigma); 5 μ M ANS in 10 mM Hepes, pH 7.5. The methanol and pentanol were not rigorously dried. For isotropic and anisotropic experiments, the χ^2 for fitted parameters did not exceed 1.1.
 - Measurement of the isotropic fluorescence decay of H6 α_4 gave lifetimes of 0.38 ns (36 percent), 3.0 ns (28 percent), and 13.3 ns (36 percent) (the numbers in parentheses refer to the relative amplitudes of each component); in the presence of Zn²⁺ these values shifted to 0.38 ns (12 percent), 2.8 ns (15 percent), 15.5 ns (73 percent). The corresponding values for H3 α_4 in the absence of Zn²⁺ were 0.27 ns (19 percent), 3.3 ns (18 percent); 15.0 ns (62 percent), and in the presence of Zn²⁺ they were 0.25 ns (21 percent), 3.3 ns (23 percent), 14.7 ns (56 percent). The concentrations for isotropic and anisotropic fluorescence experiments were: H3 α_4 and H6 α_4 , 50 μ M; ANS, 2 μ M; Zn²⁺, 2 mM; EDTA (for experiments without Zn²⁺), 3 mM in 10 mM Hepes, pH 7.5.
 - The fact that a larger rotational correlation time (~ 9 ns) was observed for α_4 was due to some aggregation at the high concentrations required for complete complexation of ANS. Similarly, for apo-H3 α_4 the correlation time was ~ 9 ns, because of a less compact state or some degree of aggregation. Addition of Zn²⁺ to H3 α_4 resulted in several distinct long correlation times that could only arise from aggregates of varying association states.
 - P. Alexander, S. Fahnestock, T. Lee, J. Orban, P. Bryan, *Biochemistry* **31**, 3597 (1992). K. P. Murphy, P. L. Privalov, S. J. Gill, *Science* **247**, 559 (1990); R. L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8069 (1986).
 - The ellipticity at 222 nm (AVIV 62DS CD with an HP temperature controller) was recorded on 8.5 μ M samples in 50 mM tris, pH 7.5. For H6 α_4 , the sample included 1.0 mM Co²⁺. The monomolec-
 - ular equilibrium constant for the unfolding transition is given by $K = (1 - f)/f$. To calculate f we assume that the ellipticity of the folded protein (at a given temperature) is the ellipticity in the absence of GdmCl, and the ellipticity of the unfolded form (at a given temperature) is that in the presence of 8.0 M GdmCl. The theoretical temperature-dependent unfolding curve for α_4 in 6.0 M GdmCl was based on fits to the CD data according to the method of Schellman (23) and that for lysozyme was reproduced from thermodynamic parameters in (23). For H6 α_4 , accurate fitting of the data was not possible, presumably because of the presence of multiple equilibria due to high temperature dissociation of the metal.
 - B. Chen and J. A. Schellman, *Biochemistry* **28**, 685 (1989).
 - P. L. Privalov, *Crit. Rev. Biochem. Mol. Biol.* **25**, 281 (1990).
 - J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1981).
 - The derivative of α_2 containing four-His residues has the sequence Ac-G(ELEELHKLLHLLK)G-PRRG(ELEELHKLLHLLK)GPRRG(ELEELHKLLHLLK)G-CONH₂. For this protein, ΔG_{120} was -10.3 and -1.94 kcal mol⁻¹, the midpoint was 5.6 and 3.75 M GdmCl and $RT\Delta\beta$ was 1.86 and 0.55 kcal mol⁻¹ M⁻¹, in the presence and absence of 1.0 mM Zn²⁺, respectively.
 - A. Godzik, A. Kolinski, J. Skolnick, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
 - A. Matouschek, L. Serrano, E. M. Meiering, M. Bycroft, A. R. Fersht, *J. Mol. Biol.* **224**, 837 (1992).
 - K. Langsetmo, J. Fuchs, C. Woodward, *Biochemistry* **30**, 7609 (1991).
 - M. McGregor, S. Islam, M. Sternberg, *J. Mol. Biol.* **198**, 295 (1987).
 - M. H. Hecht, J. S. Richardson, D. C. Richardson, R. C. Ogden, *Science* **249**, 884 (1990).
 - K. W. Hahn, W. Klis, J. M. Stewart, *ibid.* **248**, 1544 (1990).
 - H. Morii, K. Ichimura, H. Uedaira, *Proteins* **11**, 133 (1991).
 - D. Raleigh and W. F. DeGrado, *J. Am. Chem. Soc.* **114**, 10079 (1992).
 - E. K. O'Shea *et al.*, *Science* **254**, 539 (1991).
 - D. M. E. Szebenyi and K. Moffat, *J. Biol. Chem.* **261**, 8761 (1986); M. Akke, T. Drakenberg, W. J. Chazin, *Biochemistry* **31**, 1011 (1992).
 - Y. S. Babu, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **204**, 191 (1988).
 - M. Akke, T. Drakenberg, W. J. Chazin, *Biochemistry* **220**, 173 (1991).
 - R. F. Steiner, *Biopolymers* **22**, 1121 (1984).
 - M. Mütter and S. Vuilleumies, *Angew. Chem. (Int. Ed.)* **28**, 535 (1989).
 - We thank A. Rockwell for the purification of α_4 , S. Anthony-Cahill for providing the results on Ac- α_4 -CONH₂, and S. Ho and P. Kienker for critical reading of the manuscript. Supported by a grant from the Office of Naval Research.

9 November 1992; accepted 12 May 1992