# A *de Novo* Designed Protein Mimics the Native State of Natural Proteins

Daniel P. Raleigh,† Stephen F. Betz,‡ and William F. DeGrado*,‡,§

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

*Received March 14, 1995*

The *de novo* design of a protein that mimics the properties of natural proteins is an important goal.[1,2] Although some progress has been made in the design of coiled coils,[3] including a *de novo* designed peptide with native-like behavior,[3b] all *globular* proteins designed to date appear to show some of the features of molten globules.[4] Here, we describe a peptide that self-assembles into a protein with native-like physical properties. This was achieved by a strategy which included not only interactions that stabilized the desired three-dimensional structure but also interactions that destabilized potential alternative folds.

Several designed four-helix bundles have been prepared with a strict heptad repeat of $HB_a$-$HP_b$-$HP_c$-$HB_d$-$HB_e$-$HP_f$-$HP_g$ (HB and HP are hydrophobic and hydrophilic amino acids, respectively).[5,6] These include tetrameric peptides (designated $\alpha_1$ peptides), dimeric helix−turn−helix units ($\alpha_2$), and single chain proteins ($\alpha_4$). The $\alpha_2$ dimers behave very similarly to the full length $\alpha_4$ proteins, hence we have concentrated initial studies on the $\alpha_2$ peptides. The first design, $\alpha_2$B, contained Leu as its sole hydrophobic residue,[5a,b] but its core was unusually flexible compared to native proteins. We therefore replaced many of the Leu residues with structurally diverse apolar amino acids that pack together in a geometrically complementary fashion,[6] resulting in $\alpha_2$C, which undergoes a transition from a native-like to a molten globule-like state near room temperature. However, even at low temperature, $\alpha_2$C retained some properties associated with molten globules: it showed broadened NMR resonances and bound the hydrophobic dye 8-anilino-1-naphthalenesulfonate (ANS). An alternate strategy, which has been partially successful, has been to introduce metal-binding sites into $\alpha_2$B and $\alpha_4$.[5c,e] We therefore attempted to combine these strategies by building a metal-binding site into the repacked $\alpha_2$C peptide. The resulting sequence $(\alpha_2D)^7$ contains two His residues at positions 11 and 26 and a Glu residue at position 7. This peptide assembles into a dimeric protein with all the hallmarks of a native protein, even in the absence of bound metal ions.

Analytical ultracentrifugation demonstrates that the peptide exists in a monomer−dimer equilibrium,[8] and the concentration dependence of the far-UV CD signal at 222 nm[8] indicates that the dimers are highly helical whereas the monomers are random coils. The free energy of dimerization of $\alpha_2$D is $-7.4$ kcal mol$^{-1}$ (1 M standard state) at 298 K.

The interior-facing amino acids of $\alpha_2$D are well-ordered as assessed by NMR spectroscopy. Several ring current shifted methyl resonances are observed across a 0.75 ppm window in the folded state versus a 0.2 ppm window for the denatured protein (Figure 1). The upfield region is also relatively sharp and well-resolved. Unfortunately, the fingerprint region of the two-dimensional correlation spectrum is overlapped due to limited dispersion of the $C_\alpha$ protons, precluding sequential assignments. We therefore are cloning and expressing a dimeric form of $\alpha_2$D to permit isotopic labeling for multidimensional NMR analysis.

Titrations with ANS also provide evidence for a well-packed core. $\alpha_2$B binds ANS with a dissociation constant on the order of 50 $\mu$M[5e] while no binding of ANS to $\alpha_2$D could be detected using fluorescence under typical conditions[8] or NMR at millimolar concentrations.[9] Each monomer contains a single Trp residue that exhibits a strong near-UV CD signal, suggesting that the indole side chain is well-ordered. Its CD signal titrates with the p$K_a$ of a His side chain,[8] highlighting the crucial role of the residues at positions 11 and 26. Also, the fluorescence emission spectrum of this residue is consistent with its designed partially-exposed location.[8]

While most small, monomeric natural proteins display cooperative thermal denaturation transitions[10] few designed proteins have done so. $\alpha_2$D exhibits a cooperative thermal denaturation, with coincident loss of both secondary and tertiary structure monitored by following near-UV and far-UV CD signals (Figure 2). The same transition was also observed by monitoring several NMR resonances. These results suggest that

* Author to whom correspondence should be addressed.
† Present address: Department of Chemistry, State University of New York, Stony Brook, NY 11794-3400.
‡ The DuPont Merck Pharmaceutical Company.
§ Johnson Research Foundation.

(1) (a) DeGrado, W. F.; Wassermann, Z. R.; Lear, J. D. *Science* **1989**, *243*, 622. (b) Mutter, M.; Vulleumies, S. *Angew. Chem., Int. Ed. Eng.* **1989**, *28*, 535. (c) DeGrado, W. F.; Raleigh, D. P.; Handel, T. M. *Curr. Opin. Struct. Biol.* **1991**, *1*, 984. (d) Richardson, J.; Richardson, D. C.; Tweedy, N. B.; Gernert, K. M.; Quinn, T. P.; Hecht, M. H.; Erickson, B. W.; Yan, Y; McClain, D. M. E.; Donlan, M. E.; Surles, M. C. *Biophys. J.* **1992**, *63*, 1186. (e) Betz, S. F.; Raleigh, D. P.; DeGrado; W. F. *Curr. Opin. Struct. Biol.* **1993**, *3*, 601. (f) Sasaki, T.; Lieberman, M. *Tetrahedron* **1993**, *49*, 3677. (g) Tuchscherer, G.; Doemer, B.; Sila, U.; Kamber, B.; Mutter, M. *Tetrahedron* **1993**, *49*, 3559.

(2) (a) Goraj, K.; Renard, A.; Martial, J. A. *Protein Eng.* **1990**, *3*, 259. (b) Kaumaya, P. T.; Berndt, K. D.; Heidorn, D. B.; Trewhella, J.; Kèzdy, J. F.; Goldberg, E. *Biochemistry* **1990**, *29*, 13. (c) Hecht, M. H.; Richardson, J. S.; Richardson, D. C.; Ogden, R. C. *Science* **1990**, *249*, 884. (d) Ghadiri, R. M.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825. (e) Fedorov, A. N.; Dolgikh, D. A.; Chemeris, V. V.; Chernov, B. K.; Finkelstein, A. V.; Schulga, A. A.; Alakhov, Y. B.; Kirpichnikov, M. P.; Ptitsyn, O. B *J. Mol. Biol.* **1992**, *225*, 927. (f) Mutter, M.; Tuchscherer, G. C.; Miller, C.; Altman, K.-H.; Carey, R. I.; Wyss, D. F.; Labhardt, A. M.; Rivier, J. E. *J. Am. Chem. Soc.* **1992**, *114*, 1463. (g) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht M. F. *Science* **1993**, *262*, 1680. (h) Fezoui, Y.; Weaver, D. L.; Osterhout, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3675. (i) Roberston, D.; Farid, R.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J.; Wand, A. J.; DeGrado, W. F.; Dutton P. L. *Nature* **1994**, *368*, 425. (j) Yan, Y.; Erickson, B. W. *Protein Sci.* **1994**, *3*, 1069. (k) Quinn, T. P.; Tweedy, N. B.; Williams, R. W.; Richardson, J. S.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8747. (l) Kitakuni, E.; Kuroda, Y.; Oobatake, M.; Tanaka, T.; Nakamura, H. *Protein Sci.* **1994**, *3*, 831.

(3) (a) Hodges, R. S. *Curr. Biol.* **1992**, *2*, 122. (b) O'Shea, E. K.; Lumb, K. J. ; Kim, P. S. *Current Biology* **1993**, *3*, 658. (c) Myszka, D. G.; Chaiken, I. M. *Biochemistry* **1994**, *33*, 2363. (d) Monera, O. D.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1994**, *33*, 3862.

(4) (a) Kuwajima, K. *Proteins* **1989**, *6*, 87. (b) Semisotnov, G. V.; Rodionova, N. A.; Razgulyaev, O. I.; Uversky, V. N.; Gripas, A. F.; Gilmanshin, R. I. *Biopolymers* **1991**, *31*, 119.

(5) (a) Ho, S. W.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751. (b) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976. (c) Handel, T. M.; DeGrado, W. F. *J. Am. Chem. Soc.* **1990**, *112*, 6710. (d) Osterhout, J. L.; Handel, T. M.; Na, G.; Toumadje, A.; Long, R. C.; Connolly, P. J.; Hoch, J. C.; Johnson, W. C.; Live, D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 331. (e) Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879. (f) Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, P. L.; Robertson, D. E.; DeGrado, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 856.

(6) Raleigh, D. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 1079.

(7) $\alpha_2$D, Ac-GEVEELEKKFHELWKGPRRGEIEELHKKFKELIKG-$NH_2$, was prepared as described[5f] and homogeneous as assessed by FAB mass spectrometry and HPLC.

(8) See supporting information.

(9) A concentrated solution of $\alpha_2$D (pH 7.3 in $D_2O$) was titrated into 500 $\mu$L of a 0.200 mM solution of ANS in $D_2O$, pH 7.3. No shifts in the NMR resonances of either the dye or the peptide were observed, even when the peptide was present in 10-fold excess.

(10) (a) Privalov, P. L.; Gill, S. J. *Adv. Protein Chem.* **1988**, *39*, 4. (b) Alexander, P.; Frahnestock, S.; Lee, T.; Orban, J.; Bryan, P. *Biochemistry* **1992**, *31*, 3597.

0002-7863/95/1517-7558$09.00/0 

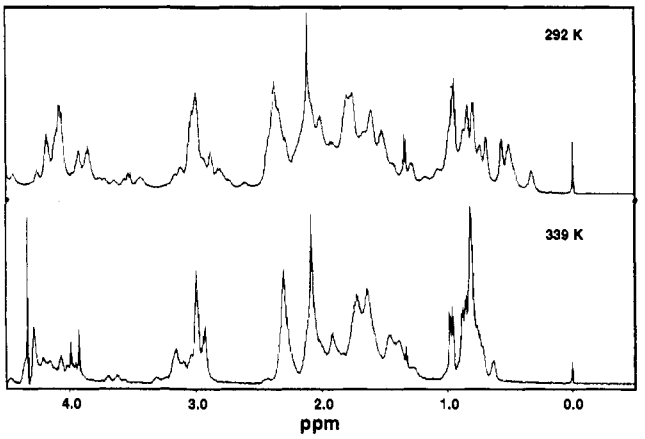


**Figure 1.** The upfield region of the NMR spectrum of $\alpha_2$D at 292 K (top spectrum) and 339 K (bottom spectrum) in $D_2O$ at pH 7.3. Spectra were recorded on a Bruker AMX-600 MHz spectrometer with weak irradiation to saturate residual HOD. Chemical shifts are given in ppm from TMS.

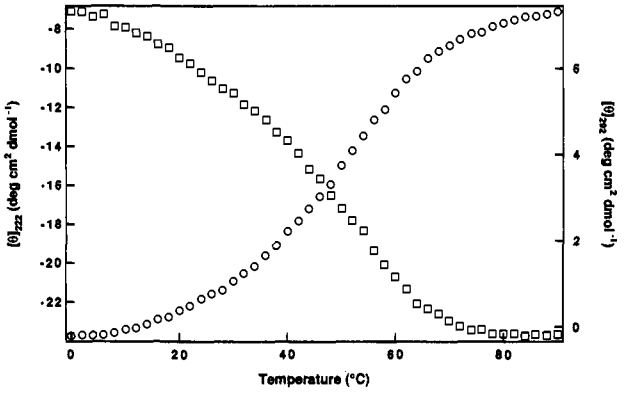


**Figure 2.** Temperature dependence of the near- and far-UV CD signals of $\alpha_2$D at 222 (○) and 292 nm (□) at 124 $\mu$M in 50 mM MOPS, pH 6.9.

denaturation occurs in a cooperative two-state process. The van't Hoff enthalpy[8] at 494 $\mu$M and 326 K is 26 ± 1 kcal (mol of monomer)$^{-1}$, typical for natural proteins this size, and $\Delta C_p$ is 9.4 ± 0.9 cal mol$^{-1}$ K$^{-1}$ per residue, near the values typically observed for small globular proteins (10−15 cal mol$^{-1}$ K$^{-1}$ per residue).[10]

One criterion that we have suggested for determining the "native-ness" of a designed protein is that its amide proton exchange rates should be on the order of those in small, globular proteins,[1e] and this has been observed for a designed heterodimeric coiled coil.[3b] The dissociation constant for $\alpha_2$D is on the order of micromolar. At the millimolar concentrations necessary for NMR, ~5% of $\alpha_2$D is a nonhelical monomer. Unfortunately, rapid exchange from the monomer precluded determination of slowed exchange from the dimer.

The nature of the polar residues at positions 7 and 26 is not critical; a variant of $\alpha_2$D containing His at each of these positions behaves similarly (D.P.R. and W.F.D., unpublished observations). However, both polar groups are required, as

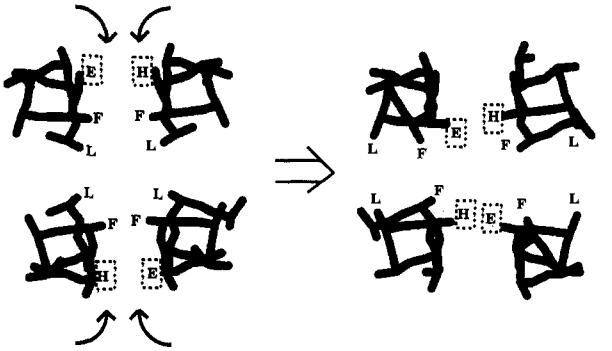


**Figure 3.** Axial view of the central portion (seven-residue slice) of the working model of the $\alpha_2$D dimer (left panel). Residues in "e" positions (boxed) are hydrophilic in $\alpha_2$D, but hydrophobic in earlier versions. A counterrotation of the helices (indicated by the curved arrows) results in burial of the "e" positions. This alternate packing should be highly energetically unfavorable for $\alpha_2$D, but almost isoenergetic for $\alpha_2$B.

peptides in which one of these residues is replaced by Leu exhibit diffuse thermal denaturation transitions. Furthermore, versions of $\alpha_2$ that contain only Leu residues in the core but contain the interfacial polar groups display molten globule-like properties in the apo state.[5c,e] Thus, *both* the presence of the repacked helix−helix interface as well as breaks in the repeating heptad pattern of hydrophilic and hydrophobic residues are required for conformational specificity.[11] Interestingly, a conserved hydrophilic Asn on the otherwise uninterrupted hydrophobic side of an amphiphilic α-helix is similarly important for the conformational specificity of several two-stranded coiled coils related to GCN4.[12] The breaks in the hydrophobic periodicity of $\alpha_2$D may play a similar role as described in Figure 3.

Recently, a family of random proteins based on the same heptad pattern as $\alpha_4$ (with a variety of apolar residues at "a", "d", and "e" positions) has been described.[2g] It will be interesting to see if any of these randomly generated molecules exhibit native-like behavior without the addition of polar residues at "a", "d", or "e" positions.

**Acknowledgment.** We thank Jim Lear for assistance with the ultracentrifuge experiments and Karyn O'Neil, Rob Fairman, and Jim Bryson for helpful discussions.

**Supporting Information Available:** Data and fits to analytical ultracentrifuge experiments, concentration dependent CD experiments, tryptophan and ANS fluorescence emission measurements, and analysis of pH and thermal denaturation data (9 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA950834V

(11) Dill, K. A. *Curr. Opin. Struct. Biol.* **1993**, *3*, 99.
(12) (a) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber T. A. *Science* **1991**, *254*, 539.(b) Betz, S.; Fairman, R.; O'Neil, K.; Lear, J.; DeGrado, W. *Philos. Trans. R. Soc. London B* **1995**, *348*, 81. (c) O'Shea, E. K.; Lumb, K. J.; Kim, P. S. *Curr. Biol.* **1993**, *3*, 658. (d) Wendt, H.; Berger, C.; Baici, A.; Thomas, R. M.; Bosshard, H. R. *Biochemistry* **1995**, *34*, 4097.