Tetrahedron Vol. 49, No. 17, pp. 3677-3689, 1993
Printed in Great Britain

0040-4020/93 $6.00+.00

Between the Secondary Structure and the Tertiary Structure Falls the Globule: A Problem in De Novo Protein Design

Tomikazu Sasaki* and Marya Lieberman

Department of Chemistry, University of Washington
Seattle, WA 98195

(*Received in USA* 9 *March* 1993)

Abstract. Three-helix bundle proteins can be assembled by metal driven association of bipyridine-modified peptides. The metal tris-bipyridine complex offers four different templates on which three helices may be arranged to form a bundle structure. Analysis of one of such proteins by HPLC, NMR, and CD shows that all four templates are used randomly. Therefore, this artificial protein does not have a single well defined tertiary structure, but, instead, exists as an ensemble of at least four closely related structures. Mutant peptides with diversified interior residues did not form a stable three-α-helix bundle protein. Packing interactions of interior hydrophobic residues must be optimized to construct artificial proteins with a well defined tertiary structure. This system not only permits us to study packing interactions of secondary structures in artificial proteins but also offers a unique model system for a molten globule state of native proteins.

INTRODUCTION

Protein tertiary structures can be viewed as an assembly of secondary structures connected by flexible loops. This model has been the basis for the design of several small artificial proteins[1]. The design process is simple; one first creates secondary structure units and then connects them together with appropriate peptide segments. Idealized amphiphilic α-helices are most often used as elements of secondary structure. Many natural proteins include supersecondary structures made up of amphiphilic α-helices. The leucine zipper, which is a coiled-coil of two helices, and the four-helix bundle are two such motifs. The interiors of these native helical proteins are often occupied with multiple

hydrophobic residues, and the hydrophobic interactions[2] may be the primary driving force for the folding. Amphiphilic α-helices are also known to form aggregates in aqueous solution. Many artificial helical proteins have, therefore, been designed to stabilize their tertiary structures by utilizing the hydrophobic interactions to the maximum extent.

A recent strategy for protein design is to replace flexible loop structures with a rigid template molecule. The template may be non-biological and may incorporate useful functionality, such as a spectroscopic probe or reactive group. Many artificial proteins have been synthesized by covalently attaching amphiphilic α-helices to templates[3]. For example, coproporphyrin was used to assemble a four α-helix bundle protein[3a] as a model for cytochrome P-450. A 15-residue amphiphilic helix was covalently attached to the four carboxylate groups of coproporphyrin I. Fe(II) ion was then incorporated into the porphyrin moiety to bind and activate oxygen for the hydroxylation reaction. Porphyrin serves not only as a template but also as a functional site in the designed protein. This molecule was capable of hydroxylating aniline and showed some substrate specificity. Its thermodynamic stability has been reported[3b] to be similar to that of native proteins of similar size. The template approach has been used to create other functional proteins, including ion-channel proteins[3c] and an artificial esterase[3d].

Simple principles, such as the Chou-Fasman parameters[4] for helix-forming tendency, and indexes of hydrophobicity to estimate amphiphilicity, have allowed many workers to design desired secondary structures which show evidence of cooperative interaction to form larger structures. Circular dichroism spectroscopy indicates that predicted secondary structures do indeed form in designed proteins. The thermodynamic stabilities of some synthetic bundle proteins have been reported[5] to be equal to or greater than native proteins of the corresponding size, and just as in native proteins, folding is cooperative. Designed proteins can also show chemical activity as described above. To a first approximation it would, therefore, appear that this approach has been successful. However, there are lingering uncertainties[6] about the packing interactions of core residues in the folded state. Structural information on artificial proteins is quite limited; there are only two X-ray studies[7] and no solution NMR structures in the literature[8].

In this paper, we will describe our recent efforts to construct three-α-helix bundle proteins with a packed interior core using a tris-bipyridine metal complex as a template. Our results suggest that in at least one designed protein, the folded state contains at least four sub-states of almost equal energy, rather than a single unique conformation.

RESULTS

Synthesis of a Three Helix Bundle Protein

Recently, we reported[3c] the assembly of a three-helix bundle protein on a tris-bipyridine Fe(II) complex (Figure 1). Bipyridine is known to form stable tris complexes with various metal ions[9]. Fe(II) was chosen as a metal ion template in this study because of its extremely large third binding constant relative to the first and second binding constants. This is important for the assembly of three-α-helix bundle proteins because other metal ions may form a mixture of mono-, bis, and tris-complexes at low peptide concentrations. Spectroscopic titration of bipyridine-modified peptides with Fe(II) showed clear isosbestic points, indicating the absence of mono- and bis-complexes. HPLC analysis of the resulting complexes also detected no evidence for these intermediate species under the conditions we employed in this study.

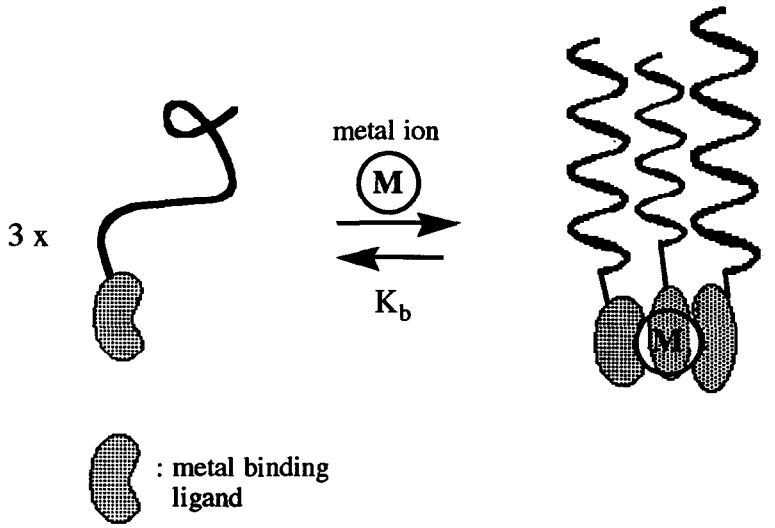


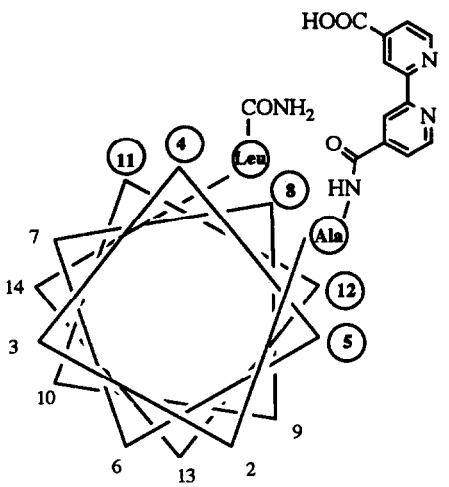
FIGURE 1. Schematic representation of the synthesis of a three-α-helix bundle protein by a metal-driven association of peptides modified with a metal binding ligand.

Bipyridine was covalently attached to the N-terminus of a 15-residue peptide which was designed to form an amphiphilic α–helix. In the presence of Fe(II) ion, the bipyridine-modified peptide (pepy) trimerizes with concomitant increase (35% to 85%) in α–helicity. We attributed the increase in α–helicity to favorable hydrophobic interactions between the peptides when they pack together in a bundle of α–helices. Such favorable interactions should increase the formation constant for the trisbipyridine Fe(II) complex.

a)

peptide	sequence	α-helical hydrophobic moment
	1 5 10 15	
1 (pepy)	**Ala**-Glu-Gln-**Leu**-**Leu**-Gln-Glu-**Ala**-Glu-Gln-**Leu**-**Leu**-Gln-Glu-**Leu**-$CONH_2$	5.52
2	- Gln-Glu-**Ala**-**Leu**-Glu-Gln-**Phe**-Gln-Lys-**Ala**-**Gln**-Lys-Gln - -	6.07
3	- - - **Phe**-**Trp** - - **Phe** - - **Ser**-**Gln** - - - -	5.27
4	- - - **Phe**-**Gly** - - **Gly** - - **Ala**-**Leu** - - - -	5.85

b)



c)

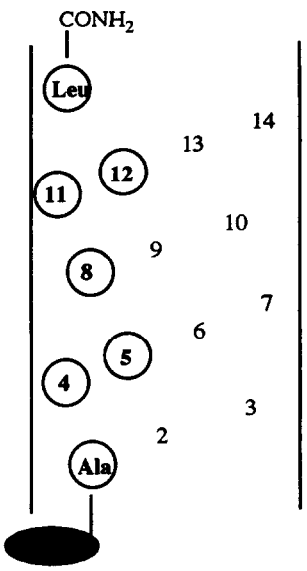


FIGURE 2. a). Amino acid sequences. Ala = Alanine, Glu = Glutamic acid, Phe = Phenylalanine, Gly = Glycine, Lys = Lysine, Ser = Serine, Gln = Glutamine, and Trp = Tryptophan. Identical residues are represented by dash. Anticipated interior core residues are highlighted by bold letters. b). Helical wheel diagrams. Hydrophobic residues are circled. c). Helical net diagrams. Hydrophobic residues are circled.

Indeed, the formation constant for the $Fe(pepy)_3$ complex was found to be significantly higher than that for a model compound, bipyridine-modified alanine.

Design of Mutated Peptide Sequences

In order to examine hydrophobic packing interactions, we synthesized several mutant peptides with different core residues. Figure 2 shows the sequence of the mutants used in this work. The hydrophobic surfaces were designed to present different shapes and hydrophobicities. The hydrophobic moments[10] calculated for the sequences of the mutants are comparable to pepy and to those of membrane-binding helical segments, which are among the highest amphiphilicities found in native proteins. The three mutants would have the same hydrophilic surface were they to adopt an α–helical conformation. An ion pair is expected to form between Lys_{10} and Glu_6 to stabilize[11] an α-helical conformation. Other Lys and Glu residues are placed at the C- and N-termini to neutralize the helix dipole, which would otherwise be expected[12] to destabilize the head-to-head packing interactions of the α-helices. The 15-residue peptides are all predicted[4] to have a large helix-forming propensity.

Fe(II) Complexes of Pepy and Its Mutants

All bipyridine-modified peptides formed stable Fe(II) complexes on mixing with an aqueous solution of ammonium ferrous sulfate at pH 4.8. Table 1 shows absorption maxima and extinction coefficients of pepy mutants and their Fe(II) complexes. The absorption spectra are similar to that of the Fe(II) complex of bipyridine-modified alanine, a model compound that lacks peptide-peptide interactions, indicating that the electronic structure of the tris-bipyridine Fe(II) core is not significantly perturbed by an attached protein domain.

TABLE 1. Electronic absorption maxima of synthetic peptides and their Fe(II) complexes, and formation constant for Fe(II) complexes of ligands

peptide	λ_{max} (nm, ε x 10^4 M^{-1} cm^{-1})	Fe(II) complex (λ_{max}) (nm, ε x 10^4 M^{-1} cm^{-1})	binding constant (x 10^{16} M^{-3})
1, pepy	294, 1.12	542, 1.47	9(4)
2	296, 1.14	540, 1.39	2.4(8)
3	286, 1.67	540, 1.80	3(1)
4	294, 1.21	540, 1.53	1.0(5)
bipy-Ala	293, 1.32	540, 1.68	2(1)

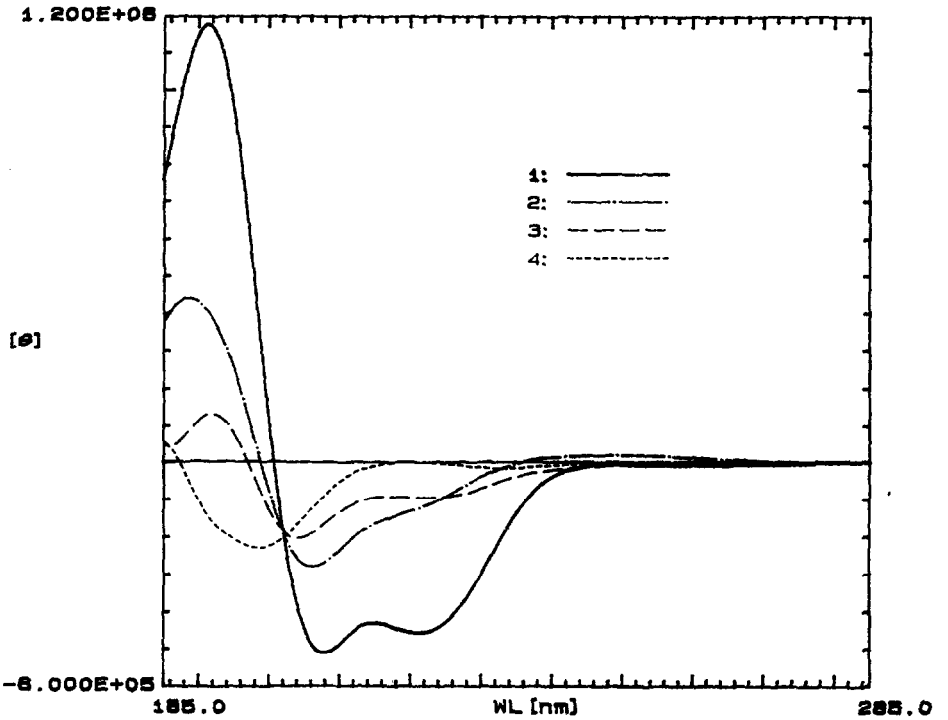


FIGURE 3. Circular dichroism spectra of Fe(II) complexes of bipyridine-modified peptides in 10 mM acetate buffer, pH 4.8. All spectra were recorded at 25 °C and smoothed by software provided by JASCO. **1** (solid), **2** (center), **3** (dashed), and **4** (dotted).

The overall stability (ΔG_T) of the artificial protein can be expressed by the following general equation where ΔG_C , ΔG_P , and ΔG_L are Gibbs free energies for metal coordination, peptide-peptide interactions, and connection of bipyridine and peptide moieties, respectively[13].

$$\Delta G_T = \Delta G_C + \Delta G_P - \Delta G_L$$

ΔG_C and ΔG_L for all pepy mutants are considered to be the same since all mutants contain the same metal-binding moiety. The formation constants of Fe(II) complexes of pepy mutants, therefore, reflect the extent of inter-peptide interactions.

A large increase in α-helicity (35% to 85%) was observed with pepy in the presence of Fe(II) as reported previously. Mutants **2** and **3** , however, showed only a small increase in α-helicity (19% to 23%, and 14% to 17% for **2** and **3**, respectively), and mutant **4** showed no change in its α-helicity (<2%). Fig. 3 shows CD spectra of Fe(II) complexes of pepy and its mutants. No clear correlation was observed between hydrophobic moments and α-helicities of bundle proteins of pepy derivatives, suggesting that hydrophobic interaction alone is not sufficient to stabilize an α-helix bundle structure. There is a rough correlation between degree of α-helicity of three α-helix bundle proteins and the formation constants of the Fe(II) complex of the corresponding peptides.

Isomers of Fe(II)(pepy)$_3$

Since the bipyridine moiety of pepy and its mutants is unsymmetrical, four stereoisomers, *fac*-Λ, *fac*-Δ, *mer*-Λ and *mer*-Δ, are possible at the Fe(II) center (Figure 4). The structural differences between the *fac* and *mer* templates are significant; for example, in the *fac* isomers, the bases of the helices are positioned 9 Å apart, while in the *mer* isomers, two helices are 14 Å distant from one another and each is 9 Å away from the third helix. These structural differences affect the geometry of the bundle in which three peptides are packed with each other. The chiral peptides will interact diastereomerically with the Δ and Λ enantiomers at the metal center of each *fac* and *mer* isomer. The protein domain, therefore, has a choice of four distinct template structures on which to arrange its three α-helices. Increased preference for one of the isomers is expected for a three-helix bundle protein with a better packed interior core.

Reverse phase HPLC was used to separate four isomers of Fe(II)(pepy)$_3$ as shown in Figure 5. The isomers slowly interconvert, yielding a mixture of isomers at equilibrium that is the same as the mixture that forms on mixing pepy and iron. The isomer ratio was 9.5(1): 17(2): 43(1): 29(1) for *fac*-Λ, *fac*-Δ, *mer*-Λ and *mer*-Δ isomers, respectively. The

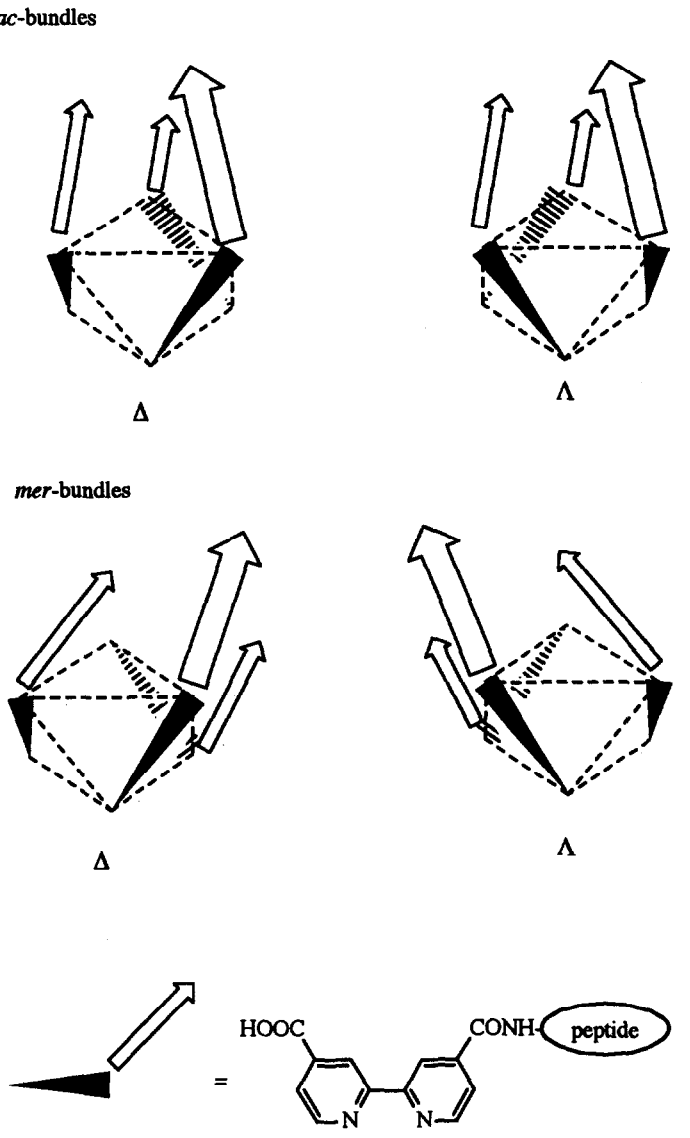


FIGURE 4. Schematic representation of the formation of a three-α-helix bundle protein on four isomeric templates. Three bipyridines bind in octahedral coordination.

observed ratio is close to the predicted ratio(12.5: 12.5: 37.5: 37.5)) for a statistically random mixture, which means that the formation constants for the four isomers are almost identical.

The *fac* and *mer* stereoisomers were identified by ^{1}H-NMR, and the absolute chirality of each separated isomer was determined by CD spectroscopy[14]. The curve shapes for all four of the isomers indicate a high α-helical content, consistent with the helicity observed in the original mixture.

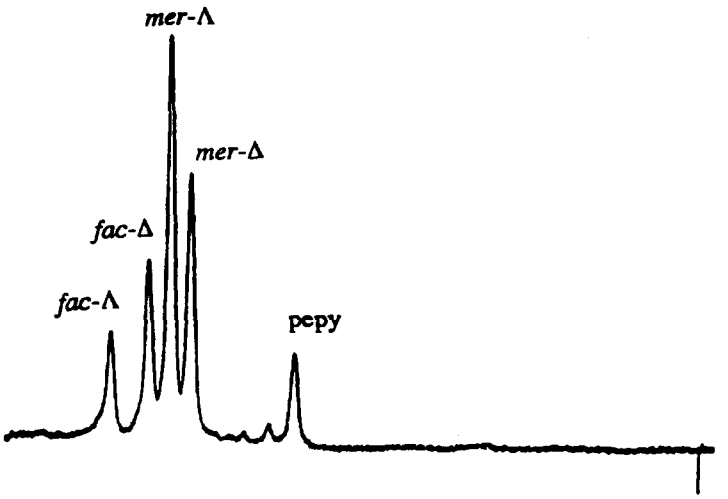


FIGURE 5. HPLC of four isomers of Fe(pepy)$_3$. A C_4 reversed-phase semi-preparative column was used, with a flow rate of 4 ml/min and a linear gradient of 40% B to 100%B in 20 minutes. Eluent A was 80% water, 20% CH_3CN, 0.1% TFA; eluent B was 20% water, 80% CH_3CN, 0.1% TFA. The trace shows absorbance at 294.5 nm, which is the isosbestic point for pepy and Fe(pepy)$_3$.

DISCUSSION

The protein portion of Fe(pepy)$_3$ appears to be very adaptable in terms of the template structures that it will accept. Had we used a flexible template, it is likely that we would have seen a time averaged ensemble of such states. By having four isolable template structures that are in thermodynamic equilibrium, we can show that the potential energy

surface for this three-helix bundle is surprisingly broad, with at least four states of different structure but similar energy.

Under certain conditions, some native proteins appear to lose their well-defined tertiary structures while retaining most of their secondary structure content. This has been termed a molten globule state[15]. It is believed that the hydrophobic regions of the protein remain well buried, but that the secondary structure elements slide around among a welter of isoenergetic states. The molten globule has been proposed as an intermediate stage in protein folding. It appears from our work that the artificial protein $Fe(pepy)_3$ resembles a molten globule more than it resembles a native protein, and this may be true for other designed proteins as well. Recently, De Grado and co-workers used NMR spectroscopy of the amide NH region to observe what appears to be a thermally induced interconversion between a molten globule-like state and a better packed structure[16]

What does a small native protein have that $Fe(pepy)_3$ lacks, that allows it to fold into a well-defined tertiary structure? The immediate answer seems to be diversity of residues, but it is clear that not just any residues will do. Mutant peptides, **2**, **3**, and **4** did not form a stable helix-bundle protein in spite of the increased diversity in the interior core residues. Specific packing interactions within the protein core are apparently required for unique tertiary structures. Past studies of molten globules have been hindered by a lack of the most rudimentary knowledge about the potential energy surface that constrains possible structures. If the structures can not be distinguished, only one point corresponding to an averaged structure can be examined. The $Fe(pepy)_3$ system gives us four points on that energy surface, opening a new window on folding states that are intermediate between secondary and tertiary structure.

EXPERIMENTAL

Synthesis

Peptides were synthesized on an ABI model 430A peptide synthesizer using p-MBHA resin and t-BOC amino acids. 4,4'-dicarboxylic acid-2,2'-bipyridine was attached to the N-termini of the peptides on the resin support as previously described[3c] Peptides were deprotected and cleaved using TFMSA/thioanisole (low-high conditions for tryptophan-containing peptides) as described in the ABI manual[17], isolated by ether precipitation, desalted on a G-25 gel filtration column (eluant 5% HOAC), and purified by HPLC. Peptides were characterized by UV-Vis spectroscopy, electrospray ionization mass spectroscopy and amino acid analysis. Peptide concentrations were initially determined by amino acid analysis with an added valine standard, and thereafter by UV-Vis spectroscopy.

Binding Constant

Binding constants were determined spectroscopically by titration of ligands with iron. At low ligand concentrations, the binding curve deviates significantly from ideality, especially around the equivalence point. The concentration of FeL_3 was determined using

the extinction coefficient at 540-545nm. The concentrations of free ligand and iron were calculated from the concentration of FeL_3 and the known total ligand and iron concentrations. The binding constant is $[FeL_3]/[Fe][L]^3$.

In a typical experiment, 1.078 ml of 25 mM pH 4.8 acetate buffer was used to zero the spectrophotometer, and 23 µl of a concentrated bipyridine-modified peptide was added to give a 1.47 x 10^{-5} M solution. The exact concentration was determined from the known ε_{294} value. A small aliquot (1-6 µl) of a 2.043 x 10^{-4} M solution of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ was added, the cuvette was mixed, and A_{540} was recorded. This was repeated to give 12 points, in which [Fe]:[L] ranged from 0 to 1. 0.3333 is the equivalence point. Formation of the complex ion after addition of an aliquot of iron may take up to two hours, so it is important to wait until the reading at 540 nm has stabilized. Iron(II) solutions could be stored at 4 °C for up to 12 hours, but as they slowly air-oxidize (after 12 hours, titration with excess o-phenanthroline showed that a 2 x 10^{-4} M solution of Fe(II) had declined about 3% in concentration), fresh solutions were made frequently.

CD experiments

A JASCO-700 circular dichroism spectrometer was used for all CD measurements. In order to make measurements down to 184 nm, short path lengths (0.1-0.5mm), low buffer concentrations (<12mM), and nitrogen flow rates of 5 liter / min were used. Peptide concentrations of 0.5 - 2 mM were used. Cell blanks were subtracted and spectra were smoothed; as this procedure often alters the appearance of spectra (blunts sharp peaks), the smoothed spectrum was always compared to the unsmoothed spectrum. The molar ellipticity per backbone amide at 222 nm was used to estimate the percent helicity, with 100% helicity estimated[18] as 33,500 deg mol^{-1} cm^{-2}. This procedure assumes a two-state helix-coil model.

Acknowledgment. This work was supported in part by the donors of the Petroleum Research Fund, administered by the American Chemical Society. M.L. was supported by a fellowship from the Fannie and John Hertz Foundation.

REFERENCES

(1) (a) DeGrado, W. F. *Adv. Protein Chemistry* **1988**, *39*, 51.

(b) Hecht, M. H.; Richardson, J. S.; Richardson, D. C.; Ogden, R. C. *Science* **1990**, *249*, 884.

(c) Gutte, B.; Daumingen, M.; Wittschieber, E. *Nature* **1979**, *281*, 650.

(d) Tanaka, T.; Anaguchi, H.; Hayashi, M.; Fukuhara, K.; Hubbard, T. J. P.; Nakamura, H.; Ikehara, M. *Proceedings of the 12th American Peptide Society Symposium*, **1991**, 376.

(e) Unson, C. B.; Erickson, B. W.; Richardson, D. C.; Richardson, J. S. *Fed. Proc.* **1984**, *43*, 1837.

(f) McClain, R. D.; Yan, Y.; Williams, R. W.; Donlan, M., E.; Erickson, B. W. *Proceedings of the 12th American Peptide Society Symposium*, **1991**, 364.

(2) (a) Rose, G. D.; Geselowitz, A. R.; Lesser, G.J.; Lee, R. H. Zehfus, M. H. *Science* **1985**, *229*, 834.

(b) Rose, G. D.; Roy, S. *Proc. Natl. Acad. Sci., USA* **1980**, *77*, 4643.

(c) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105.

(3) (a) Sasaki, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1989**, *111*, 380.

(b) Sasaki, T.; Kaiser, E. T. *Biopolymers* **1990**, *29*, 79.

(c) Lieberman, M.; Sasaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 1470-1471.

(d) Lieberman, M.; Tabet, M.; Tahmassebi, D.; Zhang, J.; Sasaki, T. *Nanotechnology* **1991**, *2*, 203.

(e) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825.

(f) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000.

(g) Mutter, M.; Altmann, K.-H.; Vorherr, T. *Z. Naturforsch.* **1986**, *41b*, 1315.

(h) Mutter, M.; Altmann, E.; Altmann, K.-H. G.; Viulleumier, S.; Gremlich, H.-U.; Muller, K. *Helv. Chim. Acta* **1988**, *71*, 835.

(i) Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535.

(j) Hahn, K. W.; Klis, W. A.; Stewart, J. M. *Science* **1990**, *248*, 1544.

(k) Montal, M.; Montal, M. S.; Tomich, J. M. *Proc. Nat. Acad. Sci. U.S.A.* **1990**, *87*, 6929.

(4) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976.

(5) Chou, P. Y.; Fasman, G. D. *Biochem.* **1974**, *13*, 211.

(6) DeGrado, W. *203rd ACS*, **1992**, Eli Lily Award Address.

(7) (a) Eisenberg, D.; Wilcox, W.; Eshita, S. M.; Pryciak, P. M.; Ho, S. P.; DeGrado, W. F. *Proteins: Struct. Funct. Genet.* **1986**, *1*, 16.

(b) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751.

(8) Recently, Mutter *et al* reported a portion of a NOESY spectrum of a template assembled protein. See Mutter, M.; Carey, R. I.; Dorner, B.; Ernest, I.; Flogel,

R.; Giezendanner, U.; Rivier, J. E.; Servis, C.; Sigel, C.; Steiner, V.; Tuchsherer, G.; Vuilleumier, S.; Wyss, D. *Proceedings of the 12th American Peptide Society Symposium*, **1991**, 326.

(9) Irving, H.; Mellor, D. H. *J. Chem. Soc.* **1962**, 5222.

(10) Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 140.

(11) (a) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. USA* **1987**, 84, 8898. (b) Lyu, P. C.; Marky, L. A.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1989**, *111*,2733.

(12) (a) Gilson, M. K.; Honig, B. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1524. (b) Wierenga, R. K.; De Maeyer, M. C. H.; Hol, W. G. *Biochemistry*, **1985**, *24*, 1346.

(13) Jencks, W. P. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046.

(14) Lieberman, L.; Sasaki, T. unpublished results.

(15) (a) Kim, P. S.; Baldwin, R. L. *Ann. Rev. Biochem.* **1982**, *51*, 459. (b) Ptitsyn, O. B. *J. Prot. Chem.* **1987**, *6*, 273. (c) Roder, H.; Elove, G. A.; Englander, S. W. *Nature* **1988**, *335*, 700. (d) Udgaonkar, J. B.; Baldwin, R. L. *Nature* **1988**, *335*, 694. (e) Semisotnov, G. V.; Rodionova, N. A.; Razgulyaev, O. I.; Uversky, V. N.; Gripas, A. F.; Gilmanshin, R. I. *Biopolymers* **1991**, *31*, 119.

(16) Raleigh, D. P.; DeGrado, W. F. . *Am. Chem. Soc.* **1993**, *114*,10079.

(17) Applied Biosystem Co. Manual for peptide synthesis and purification. Also see Fujii, N.; Otaka, A.; Ikemura, O.; Akagi, K.; Funakoshi, S.; Hayashi, Y.; Kuroda, Y.; Yajima, H. *J. Chem. Soc., Chem. Commun.* **1987**, 274.

(18) (a) Johnson, C. *Proteins: Struct., Funct., Genet.* **1990**, *7*, 205. (b) Chang, C. T.; Wu, C.-S.; Yang, J. T. *Anal. Biochem.* **1978**, *91*, 13.