The Hydrogen Bond Mimic Approach: Solid-Phase Synthesis of a Peptide Stabilized as an α-Helix with a Hydrazone Link

Edelmira Cabezas[†] and Arnold C. Satterthwait*[,†]

Contribution from the Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037

Received September 8, 1998. Revised Manuscript Received February 15, 1999

Abstract: Proteins are characterized by extensive hydrogen bonding that defines regular and irregular substructures. However, hydrogen bonds are weak and insufficient for stabilizing peptide conformation in water. Consequently, the biological activity of peptides is reduced. This led us to test whether a covalent mimic of the hydrogen bond could be used to stabilize peptide conformation in water. A solid-phase synthesis is described for replacing a main-chain hydrogen bond (NH → O═CRNH) with a hydrazone link (N−N═CH−CH_2CH_2) in peptides. The synthesis is easy to implement, rapid, and capable of high yields. The replacement of a putative (i + 4 → i) hydrogen bond with the hydrazone at the N terminus of acetyl-GLAGAEAAKA-NH_2 (**1**) to give [JLAZ]AEAAKA-NH_2 (**2**) converts it to a full-length α-helix in water at ambient temperature as indicated by NMR spectroscopy. The observation of weak $d_{\alpha N}(i, i + 3)$, medium $d_{NN}(i, i + 1)$, and strong $d_{\alpha\beta}(i, i + 3)$ NOEs that span **2** establish the formation of a full-length α-helix in water. $J_{\alpha N}$ coupling constants and amide proton chemical shifts and temperature coefficients are consistent with a model involving rapidly equilibrating extended and α-helical conformers. Substituting L-alanine with L-proline to give [JLPZ]AEAAKA-NH_2 (**3**) enhances α-helix nucleation and shifts the equilibrium further toward full-length α-helix. The hydrazone link displays many of the properties required of a hydrogen bond mimic and could find use as a general means for constraining peptides to a range of biologically relevant conformations.

Introduction

Synthetic peptides are a rich source of drug[1,2] and vaccine leads.[3] Biologically active peptides are often searched for by screening peptides derived from proteins[4] or by screening peptide libraries[5] against receptors. However, most peptide fragments from proteins are conformationally heterogeneous in aqueous solution.[6] Since receptors bind peptides in specific conformations, it is important to consider the effect that peptide conformation has on binding energy. This question, first treated formally by Anfinsen,[7] led to the conclusion that the binding energy of a peptide is directly proportional to the free energy required for folding the peptide into the receptor bound conformation. The free-energy barrier to folding a peptide can be considerable. When synthetic peptides are constrained to shapes that mirror the binding pockets of receptors, the energy barrier is reduced, and 10^2-10^5-fold improvements in affinities are achieved.[7−9] Differences of this magnitude are sufficient for identifying new peptide leads and receptors[10] and for

* Address correspondence to this author at the following address: The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, California 92037. Telephone: 619-646-3100, ext. 3658. Fax: 619-646-3195. E-mail: asat@burnham-inst.org.

[†] Present address: The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037.

(1) Hruby, V. J. *Life Sci.* **1982**, *31*, 189−199. Hruby, V. J.; Al-Obeidi, F.; Kazmierski, W. *Biochem. J.* **1990**, *268*, 249−262. Li, B.; Tom, J. Y. K.; Oare, D.; Yen, R.; Fairbrother, W. J.; Wells, J. A.; Cunningham, B. C. *Science* **1995**, *270*, 1657−1660. Wrighton, N. C.; Farrell, F. X.; Chang, R.; Kashyap, A. K.; Barbone, F. P.; Mulcahy, L. S.; Johnson, D. L.; Barrett, R. W.; Jolliffe, L. K.; Dower, W. J. *Science* **1996**, *273*, 458−464.

(2) Müller, K.; Obrecht, D.; Knierzinger, A.; Stankovic, C.; Spiegler, C.; Bannwarth, W.; Trzeciak, A.; Englert, G.; Labhardt, A. M.; Schönholzer, P. *Perspect. Med. Chem.* **1993**, 513−531.

(3) (a) Sela, M.; Schechter, B.; Schechter, I.; Borek, F. *Cold Spring Harbor Symp. Quant. Biol.* **1967**, *39*, 537−545. (b) Lerner, R. A. *Adv. Immunol.* **1984**, *36*, 1−44. (c) Brown, F. *Philos. Trans. R. Soc. London Ser. B* **1994**, *344*, 213−219.

(4) Geysen, H. M.; Meloen R. H.; Barteling, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3998−4002. Houghten, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5131−5135.

(5) (a) Scott, J. K.; Smith, G. P. *Science* **1990**, *249*, 386−390. (b) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−84. (c) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84−86.

(6) (a) Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 795−817. (b) Dyson, H. J.; Sayre, J. R.; Merutka, G.; Shin, H. C.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 819−835.

(7) Sachs, D. H.; Schechter, A. N.; Eastlake, A.; Anfinsen C. B. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 3790−3794. Anfinsen, C. B. *Science* **1973**, *181*, 223−229.

(8) (a) Teicher, E.; Maron, E.; Arnon, R. *Immunochemistry* **1973**, *10*, 265−271. (b) Furie, B.; Schechter, A. N.; Sachs, D. H.; Anfinsen, C. B. *J. Mol. Biol.* **1975**, *92*, 497−506. (c) Komoriya, A.; Hortsch, M.; Meyers, C.; Smith, M.; Kanety, H.; Schlessinger, J. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1351−1355. (d) Judice, J. K.; Tom, J. Y. K.; Huang, W.; Wrin, T.; Vennari, J.; Petropoulos, C. X. J.; McDowell, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13426−13430.

(9) (a) Satterthwait, A. C.; Cabezas, E.; Calvo, J. C.; Wu, J. X.; Wang, P. L.; Chen, S. Q.; Kaslow, D. C.; Livnah, O.; Stura, E. A. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fourteenth American Peptide Symposium;* Mayflower Scientific Ltd.: England, 1996; pp 772−773. (b) Cabezas, E.; Stanfield, R.; Wilson, I. A.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fourteenth American Peptide Symposium;* Mayflower Scientific Ltd.: England, 1996; pp 800−801.

(10) (a) Calvo, J. C.; Perkins, M.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Thirteenth American Peptide Symposium;* ESCOM: Leiden, 1994; pp 725−726. (b) Satterthwait, A. C.; Cabezas, E.; Calvo, J. C.; Chen, S. Q.; Wu, J. X.; Wang, P. L.; Xie, Y. L.; Stura, E. A.; Kaslow, D. C. *Peptides: Biology and Chemistry; Proceedings of the 1994 Chinese Peptide Symposium;* ESCOM: Leiden, 1995; pp 229−233.

10.1021/ja983212t CCC: $18.00
Published on Web 04/07/1999

converting immunologically inactive peptides into active peptides.[10,11] Conversely, this implies that many peptide leads and receptors are lost because peptides unfold. Synthetic vaccine development may be even more dependent on the mimicry of protein structure.[3a,9−12] Consequently, there are compelling reasons for developing methods for constraining peptides to the substructures occupied by their cognate sequences in proteins.

To address this need, we proposed the development of covalent hydrogen bond mimics.[13] On average, more than 60% of the amino acids in globular proteins form intramolecular hydrogen bonds between main-chain peptide bonds.[14] In addition, secondary and irregular structures can be defined by different hydrogen-bonding patterns.[15] We therefore reasoned that the replacement of putative structure-defining hydrogen bonds with a covalent mimic could yield a general approach for constraining peptides to different protein substructures based on the position of the link and the amino acid sequence.

Initially, Arrhenius and Satterthwait[16] developed a solution-phase synthesis for replacing a main-chain hydrogen bond with a hydrazone link. However, the synthesis was lengthy, yields were low, and solubility problems arose with fully protected peptides. In this paper, we report a synthetic protocol for the insertion of the hydrazone link into a peptide by solid-phase synthesis. It allowed us to explore the α-helix-stabilizing potential of the hydrazone link in water. Work on medium-sized loops is being reported separately.[9,17]

The α-helix is an important structure in proteins.[15a,d,18] It presents a definitive test for any main-chain hydrogen bond mimic since structural incompatibilities are more likely to become evident in the α-helix than in less crowded structures such as loops. Several synthetic approaches to helix stabilization have been taken.[19] One approach builds on templates to propagate helices.[2,20,21] A second approach links amino acid side chains in a position-dependent manner to stabilize helices.[22] A third approach uses linkers, bridges or synthetic scaffolds to position peptide chains in proximity for helix-stabilizing interactions.[23]

Table 1. Representative Regular and Irregular Protein Structures Defined by Main-Chain to Main-Chain Amide Hydrogen Bonds[a,b]

protein substructure	hydrogen bond	ring size
3_{10}-helix; β-turns	$i + 3 \rightarrow i$	10
β-hairpin loops, $n > 0$	$i + 3 + n \rightarrow i$	$10 + 3n$
α-helix	$i + 4 \rightarrow i$	13
3:3 loop (rare:irregular)		
2:2 loop (defined)	$i \rightarrow i + 3$	14
2:4 loop (rare, irregular)		
β-hairpin loops, $n > 0$	$i \rightarrow i + 3 + n$	$14 + 3n$
4:4 loop (defined and irregular)	$i + 5 \rightarrow i$	16
3:5 loop (defined and irregular)	$i \rightarrow i + 4$	17
4:6 loop (irregular)	$i \rightarrow i + 5$	20
antiparallel β-sheet, parallel β-sheet	$i \rightarrow i + n$ or $i + n \rightarrow i$	variable or discontinuous

[a] Reference 15. [b] The hydrogen bond is indicated by an arrow. The arrow points from the residue donating the amide NH proton to the residue bearing the accepting carbonyl oxygen atom (NH → O═CRNH). The reference residue, *i*, is N-terminal to the second residue which is separated by the indicated number of amino acids, *n*. The ring size is defined by the number of atoms connected by the hydrogen bond.

Design

A review of hydrogen bonding in proteins[14,15] shows that different substructures can be defined by their hydrogen-bonding patterns between main-chain peptide bonds on the basis of ring size and orientation (Table 1). The hydrazone link (N−N═CH−CH_2CH_2) was designed to replace the hydrogen bond, (NH → O═C(R)NH) that forms between the main-chain amide proton of an upstream amino acid and the main-chain carbonyl oxygen of a downstream amino acid ($i + n \rightarrow i$). The hydrogen bond is replaced by the N═CH double bond, and the associated

(11) Lairmore, M. D.; Lal, R. B.; Kaumaya, P. T. *Biomed. Pept. Proteins Nucleic Acids* **1995**, *1*, 117−122. Valero, M. L.; Camarero, J. A.; Adeva, A.; Verdaguer, N.; Fita, I.; Mateu, M. G.; Domingo, E.; Giralt, E.; Andreu, D. *Biomed. Pept. Proteins Nucleic Acids* **1995**, *1*, 133−140. Sheth, H. D.; Glasier, L. M.; Ellert, N. W.; Cachia, P.; Kohn, W.; Lee, K. K.; Paranchych, W.; Hodges, R. S.; Irvin, R. T. *Biomed. Pept. Proteins Nucleic Acids* **1995**, *1*, 141−148. Urban, J.; Qabar, M.; Sia, C.; Klein, M.; Kahn, M. *Bioorg. Med. Chem.* **1996**, *4*, 673−676

(12) (a) Jemmerson, R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 9180−9184. (b) Satterthwait, A. C.; Arrhenius, T.; Hagopian, R. A.; Zavala, F.; Nussenzweig, V.; Lerner, R. A. *Vaccine* **1988**, *6*, 99−103. (c) Laver, W. G.; Air, G. M.; Webster, R. G.; Smith-Gill, S. J. *Cell* **1990**, *61*, 553−556. (d) Stanfield, R. L.; Feiser, T. M.; Lerner, R. A.; Wilson, I. A. *Science* **1990**, *248*, 712−719. (e) Müller, S.; Plaue, S.; Samama, J. P.; Valette, M.; Briand, J. P.; Van Regenmortel, M. H. V. *Vaccine* **1990**, *8*, 308−314. (f) Spangler, B. D. *J. Immunol.* **1991**, *146*, 1591−1595. (g) Cachia, P. J.; Glasier, L. M.; Hodgins, R. R.; Wong, W. Y.; Irvin, R. T.; Hodges, R. S. *J. Pept. Res.* **1998**, *52*, 289−299. (h) Verdaguer, N.; Sevilla, N.; Valero, M. L.; Stuart, D.; Brocchi, E.; Andreu, D.; Giralt, E.; Domingo, E.; Mateu, M. G.; Fita, I. *J. Virol.* **1998**, *72*, 739−748

(13) Arrhenius, T.; Lerner, R. A.; Satterthwait, A. C. *Protein Structure, Folding and Design 2;* Alan R. Liss, Inc.: New York, 1987; pp 453−465.

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

(15) (a) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167−339. (b) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221−232. Wilmot, C. M.; Thornton, J. M. *Protein Eng.* **1990**, *3*, 479−493. (c) Sibanda, B. L.; Blundell, T. L.; Thornton, J. M. *J. Mol. Biol.* **1989**, *206*, 759−777. (d) Richardson, J. S.; Richardson, D. C. *Prediction of Protein Structure and the Principles of Protein Conformation;* Plenum Press: New York, 1989; pp 1−98. (e) Donate, L. E.; Rufino, S. D.; Canard, L. H. J.; Blundell, T. L. *Protein Sci.* **1996**, *5*, 2600−2616.

(16) Arrhenius, T.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Eleventh American Peptide Symposium;* ESCOM: Leiden, 1990; pp 870−872.

(17) (a) Cabezas, E.; Wang, P. L.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fourteenth American Peptide Symposium;* Mayflower Scientific Ltd.: England, 1996; pp 734−735. (b) Cabezas, E.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fifteenth American Peptide Symposium;* ESCOM: Leiden, in press. (c) Stanfield, R. L.; Cabezas, E.; Satterthwait, A. C.; Stura, E. A.; Profy, A. T.; Wilson, I. A. *Structure*, **1999**, *7*, 131−142.

(18) Kaiser, E. T.; Kezdy, F. J. *Science* **1984**, *223*, 249−255. Degrado, W. F. *Adv. Protein Chem.* **1988**, *39*, 51−124.

(19) *Houben-Weyl Methods in Organic Chemistry. Synthesis of Peptides and Peptidomimetics;* Georg Thieme Verlag: Stuttgart, 1999; Vol. E22b, manuscript in preparation.

(20) (a) Kemp, D. S.; Curran, T. P. *Tetrahedron Lett.* **1988**, *29*, 4935−4938. (b) Kemp, D. S.; Curran, T. P.; Davis, W. M.; Boyd, J. G.; Muendel, C. *J. Org. Chem.* **1991**, *56*, 6672−6682. (c) Kemp, D. S.; Curran, T. P.; Boyd, J. G.; Allen, T. J. *J. Org. Chem.* **1991**, *56*, 6683−6697. (d) Kemp, D. S.; Allen, T. J.; Oslick, S. L. *J. Am. Chem. Soc.* **1995**, *117*, 6641−6657. (e) Kemp, D. S.; Allen, T. J.; Oslick, S. L.; Boyd, J. G. *J. Am. Chem. Soc.* **1996**, *118*, 4240−4248. (f) Kemp, D. S.; Rothman, J. H. *Tetrahedron Lett.* **1995**, *36*, 4023−4026.

(21) Austin, R. E.; Maplestone, R. A.; Sefler, A. M.; Liu, K.; Hruzewicz, W. N.; Liu, C. W.; Cho, H. S.; Wemmer, D. E.; Bartlett, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 6461−6472.

(22) (a) Felix, A. M.; Heimer, E. P.; Wang, C. T.; Lambros, T. J.; Fournier, A.; Mowles, S. M.; Maines, S.; Campbell, R. M.; Wegrzynski, B. B.; Toome, V.; Fry, D.; Madison, V. S. *Int. J. Pept. Protein Res.* **1988**, *32*, 441−454. (b) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc*. **1990**, *112*, 6046−6051. (c) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1990**, *114*, 6966−6973. (d) Bracken, C.; Gulyás, J.; Taylor, J. W.; Baum, J. *J. Am. Chem. Soc.* **1994**, *116*, 6431−6432. (e) Houston, M. E.; Gannon, C. L.; Kay, C. M.; Hodges, R. S. *J. Pept. Sci.* **1995**, *1*, 274−282. (f) Jackson, D. Y.; King, D. S.; Chmielewski, J.; Singh, S.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 9391−9392. (g) Yu, C.; Taylor, J. W. *Tetrahedron Lett.* **1996**, *37*, 1731−1734. (h) Phelan, J. C.; Skelton, N. J.; Braisted, A. C.; McDowell, R. S. *J. Am. Chem. Soc.* **1997**, *119*, 455−460.

(23) Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535−554. Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825−831. Dawson, P. E.; Kent, S. B. H. *J. Am. Chem. Soc.* **1993**, *115*, 7263−7266. Monera, O. D.; Sonnichsen, F. D.; Hicks, L.; Kay, C. M.; Hodges, R. S. *Protein Eng.* **1996**, *9*, 353−363. Mant, C. T.; Chao, H.; Hodges, R. S. *J. Chromatog. A* **1997**, *791*, 85−98. Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. *Protein Sci.* **1998**, *7*, 1404−1414. Ghirlanda, G.; Lear, J. D.; Lombardi, A.; DeGrado, W. F. *J. Mol. Biol.* **1998**, *281*, 379−391.

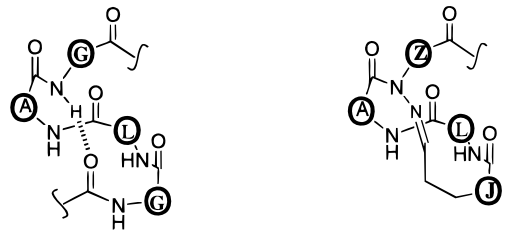


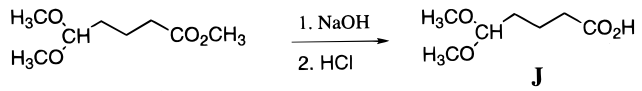
Figure 1. The substitution of an ($i + 4 \rightarrow i$) hydrogen bond in one turn of an α-helix (a) with a hydrazone covalent hydrogen bond mimic (b). The depicted structure (b) is [**J**LA**Z**]-. Circled letters at the α-carbons are one letter codes for amino acids and modified amino acids used at these positions. L replaces $HC(CH_2)CH(CH_3)_2$. A replaces $HC(CH_3)$. G, **J**, and **Z** replace CH_2.

peptide link is replaced with an ethylene group (Figure 1). Although neither hydrogen bond length nor angles are precisely mimicked by the double bond,[24] the attached hydrocarbon chain provides enough flexibility for the hydrazone link to substitute for structure defining hydrogen bonds in molecular models of peptides in many different conformations.

The α-helix is defined by a pattern of sequential ($i + 4 \rightarrow i$) hydrogen bonds.[14,15a,d] When an N-terminal ($i + 4 \rightarrow i$) hydrogen bond is replaced with the hydrazone link, a 13-membered ring is formed that corresponds to the 13-membered hydrogen-bonded turn of an α-helix (Figure 1). The hydrazone is formed by the reaction of an activated acetal (**J**) with a hydrazino derivative (**Z**). The hydrazone link fits compactly into a space-filling Corey−Pauling−Koltum molecular model of an α-helical peptide. **Z** is restricted to an α-methylene carbon (glycine equivalent) to alleviate a steric clash of side chains with the −N═CH− bond that would occur in an α-helix.

The hydrazone link was substituted for the putative ($i + 4 \rightarrow i$) hydrogen bond at the N terminus of a linear peptide, acetyl-GLAGAEAAKA-NH_2 (**1**), to give [**J**LA**Z**]AEAAKA-NH_2 (**2**). [**J**LA**Z**]- is depicted in Figure 1. The constrained peptide should form about three turns of an α-helix which is about the average size for an α-helix in a globular protein.[25a] The alanine-rich extension, -AEAAKA-NH_2 was selected on the basis of its potential for α-helix formation and disinclination to aggregate.[26] The conformations of these two peptides were compared using principally NMR spectroscopy. Alanine-based peptides are preferred for NMR studies with helical peptides because alanine is both a helix-stabilizing amino acid[26] and the only amino acid with equivalent β protons. Thus, distances to these protons are independent of side-chain rotamer distributions and consequently predictable for specific conformations. Since α-helices are characterized by a unique pattern of short, sequential distances $d_{\alpha\beta}(i, i + 3)$, alanine-based peptides can be used to test for the α-helical conformation on a per residue basis.[20e] Although NMR signals for alanine often overlap, they can be edited and assigned by means of deuterated alanine analogues.[20e] A second constrained peptide, [**J**LP**Z**]AEAAKA-NH_2 (**3**), was synthesized to test whether L-proline with $\phi = -60°$,[15d] close to the ideal for an α-helix ($\phi = -57°$), could enhance α-helix stabilization. L-proline often initiates α-helix formation in proteins.[25]

Scheme 1



Scheme 2

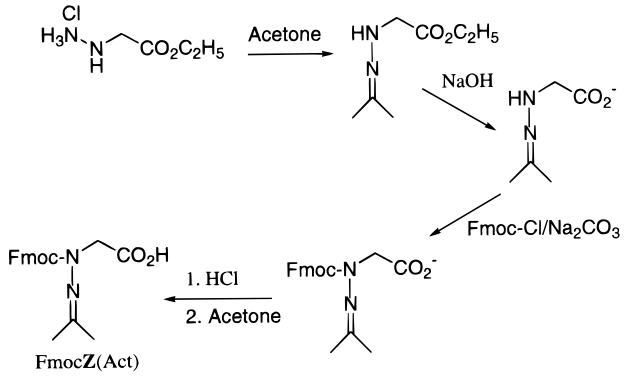


Table 2. Linear and Constrained Peptides Prepared for This Study[a]

1:	acetyl-G1-L2-A3-G4-A5-E6-A7-A8-K9-A10-NH_2
1a:	acetyl-GLA^{1}GAEAAKA-NH_2
1b:	acetyl-GLAGAEA^{1}AKA-NH_2
1c:	acetyl-GLAGAEAA^{1}KA-NH_2
1d:	acetyl-GLAGAEAAKA^1-NH_2
2:	[**J**1-cL2-cA3-**Z**4]-A5-E6-A7-A8-K9-A10-NH_2
2a:	[**J**LA**Z**]A^{4}EA^{4}A^{1}KA^4-NH_2
2b:	[**J**LA**Z**]A^{3}EA^{4}A^{1}KA^4-NH_2
2c:	[**J**LA**Z**]A^{4}EA^{3}A^{4}KA^1-NH_2
3:	[**J**1-cL2-cP3-**Z**4]-A5-E6-A7-A8-K9-A10-NH_2

[a] Amino acids (in bold) are in one-letter code and numbered according to their position in the sequence: c indicates the amino acid is from the cyclized segment of the peptide. **J** and **Z** are identified in Schemes 1, 2. **[J**LA**Z]** is the cyclized peptide in Figure 1b. A^1 is L-alanine-2-d_1; A^3 is L-alanine-3,3,3-d_3; A^4 is L-alanine-2,3,3,3-d_4.

Synthesis

The insertion of the hydrazone link into peptides by solid-phase peptide synthesis requires the prior preparation of 5,5-dimethoxy-1-oxopentanoic acid (**J**), (1-methyethylidene-2-Fmoc)hydrazinoacetic acid (Fmoc-**Z**(Act) and an *N*-α-Fmoc-L-amino acid chloride. **J** was prepared as the methyl ester in three steps according to Stevens and Lee[27] and converted to the acid (Scheme 1). The acid is stable as a neat liquid at −20 °C. Fmoc-**Z**(Act) was prepared in a one-pot synthesis according to Scheme 2. *N*-α-Fmoc-L-amino acid chlorides were prepared from the *N*-α-Fmoc-amino acid and thionyl chloride according to Carpino et al.[28]

A list of peptides synthesized for this study is in Table 2. Linear peptides including **1** were synthesized using standard Fmoc synthesis. [**J**LA**Z**]AEAAKA-NH_2 (**2**) and its deuterated analogues were synthesized on Rink's amide resin according to Scheme 3. Standard Fmoc chemistry was used for coupling *N*-α-Fmoc-amino acids, Fmoc-**Z**(Act) and **J**. The most critical reaction in peptide assembly was the coupling of *N*-α-Fmoc-L-alanine to the secondary amine, **Z**(Act). Several coupling procedures were tested including prior reaction of *N*-α-Fmoc-L-alanine with EDC to form the anhydride and activation of

(24) The mean hydrogen bond length and angles for parallel and antiparallel β-sheets and α-helix are NH···O, 2.0 Å; NH···*O*═C, 155°; N−*H*···O, 160°; Chothia, C. *Annu. Rev. Biochem.* **1984**, *53*, 537−572. The C═N bond length for pyridine, etc. is 1.34 Å; Gordon, A. J.; Ford, R. A. *The Chemists Companion; A Handbook of Practical Data, Techniques, and References;* John Wiley: New York, 1972; p 108.

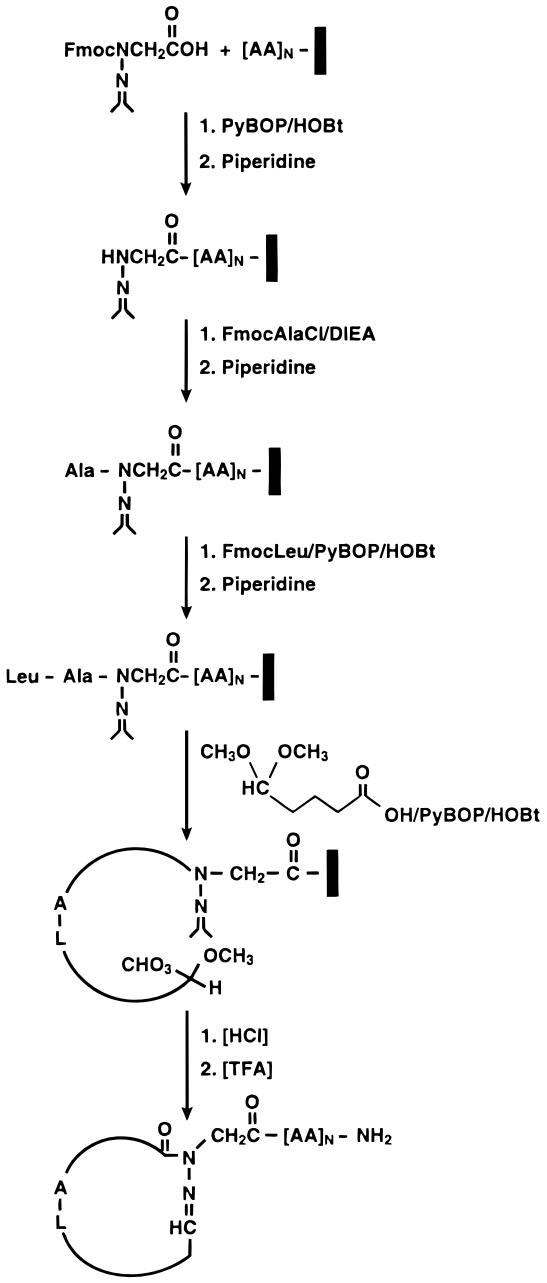
(25) (a) Presta, L. G.; Rose, G. D.*Science* **1988**, *240*, 1632−1641. (b) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648−1652.

(26) (a) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8898−8902. (b) Scholtz, J. M.; Marqusee, S.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Santoro, M.; Bolen, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2854−2858.

(27) Stevens, R. V.; Lee, A. W. M. *J. Am. Chem. Soc.* **1979**, *101*, 7032−7035.

(28) Carpino, L. A.; Cohen, B. J.; Stephens, K. E., Jr.; Sadat-Aalaee, S. Y.; Tien, J. H.; Langridge, D. C. *J. Org. Chem.* **1986**, *51*, 3732−3734.

Scheme 3



N-α-Fmoc-L-alanine with PyBOP/HOBt or DIC/HOBt. However, none of these methods proved as effective as coupling with *N*-α-Fmoc-alanine chloride. For the most efficient coupling, the acyl chloride in *N*-methylpyrrolidone was preequilibrated with **Z**(Act)-Rink-resin before adding diisopropylethylamine. The coupling reaction is complete within 15 min. Either premixing the acyl chloride with diisopropylethylamine or adding the base to **Z**(Act)-Rink-resin prior to the acyl chloride reduced the yield of the final product. Premixing of acyl chloride with diisopropylethylamine is expected to yield a quaternary amide that while highly activated could be sterically hindered from reaction with **Z**(Act). Steric hindrance could also account for reduced yields from all coupling reagents other than acyl chlorides.

The assembled peptide was cyclized on the solid support with catalytic quantities of hydrochloric acid in 20% 2,2,2-trifluoroethanol/dichloromethane. Cyclization is complete within 15 min. Several events must occur for cyclization to occur including activation of the acetal group, removal of the acetone-protecting group, and hydrazone formation (Scheme 3). The rapid cyclization suggests that hydrochloric acid catalyzes the formation of an oxocarbonium ion that in turn reacts either with the imine nitrogen or the free amine. The acetone-protecting group is removed by hydrolysis. Since similar yields were obtained with either 1 or 5 equiv of hydrochloric acid, the poorly basic *N*-amino group is not significantly protonated under the reaction conditions. The methoxocarbonium ion may undergo an exchange reaction with 2,2,2-trifluoroethanol to give the more electrophilic, trifluroethoxocarbonium ion. However, 2,2,2-trifluoroethanol is not necessary for cyclization since the yields of **2** were only partially reduced when the cyclization reaction was carried out either in *N*-methylpyrrolidone or *N*,*N*-dimethylformamide.

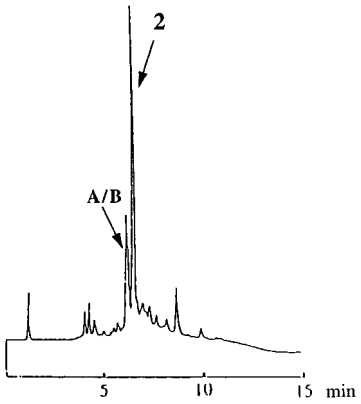


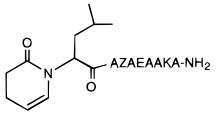
Figure 2. HPLC chromatogram of the crude product from the solid-phase synthesis of **2**. The major product is **2**. Two minor products, **A/B** are indicated. **A** (peak) is a dimer of **2**; **B** (shoulder) is an isomer of **2**.

Following cyclization, the peptide was cleaved from the Rink resin and protecting groups removed with trifluoroacetic acid. Analytical HPLC of the crude product (Figure 2) showed a single major product, **2**. Peptide **2** was purified by HPLC to give a final yield of 47%. Two minor products (**A/B**) were further purified and identified as a dimer (**A**) and an isomer of **2 (B)** by mass spectroscopy. The dimer displays an NMR signal corresponding to that for a hydrazone proton while the isomer does not. Apparently, when **J** is activated, it reacts with **Z** on an adjacent chain. This is followed by cyclization to yield the "bis" compound. The isomer **B** is unidentified. However, a reaction between the oxocarbonium ion and the cL2 amide nitrogen would give a 6-membered heterocycle with a mass identical to **2**.[29]

[**J**LP**Z**]AEAAKA-NH_2 (**3**) was synthesized in the same manner as **2**. However, the final yield of **3** following two purification steps was lower (4.7%) than that for **2**; the major product was a dimer of **3**. This suggests that L-proline slows the cyclization rate of the monomer.

Since the biological activity of peptides is normally measured in aqueous media, and synthetic vaccines must withstand weeks of exposure to physiological conditions, the chemical stability of **2** was assessed under physiological conditions of temperature and pH by monitoring it in water at pH = 7.4, 37 °C by HPLC. After one month, no change was observed in the retention time, and there was little detectable degradation. No change was observed in **2** after 1 day in 10% D_2O/90% H_2O, pH 2.9,

(29) Proposed structure for **B**. Proposed by Barry Noar.



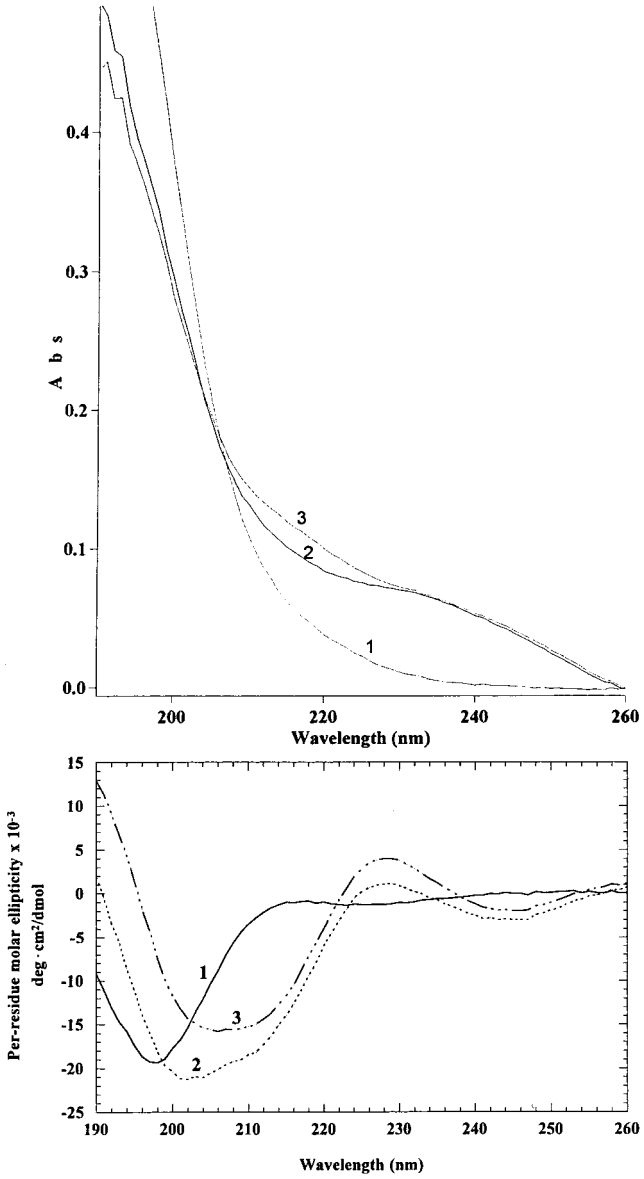


Figure 3. UV spectra for 10 μM **1**−**3** at pH 2.9, ambient temperature; path length is 1 cm (upper panel). CD spectra for 100 μM **1**−**3** at pH 2.9, 22 °C (lower panel).

conditions used for 2D NMR spectroscopy. However, **3** showed slow decomposition at pH 2.9 which became evident in NMR experiments. The lower yield and faster rate of decomposition of **3** may reflect ring strain. Peptides **2** and **3** showed little degradation under the acidic peptide cleavage conditions. However, the addition of ethanedithiol to the cleavage mixture rapidly destroys the hydrazone link and should be avoided in peptide deprotection reactions.

Conformational Analysis

CD Spectroscopy. Peptides **1**−**3** were analyzed by UV, circular dichroism (CD), and NMR spectroscopy. UV and CD spectra for **1**−**3** are compared in Figure 3. The CD spectrum for **1** in water at ambient temperature shows little evidence for helix while the CD spectra for **2** and **3** are complex. CD spectra are often analyzed in terms of UV spectra.[30] The difference in UV spectra for **1**−**3** indicates that the acylhydrazone link absorbs strongly in the UV with a broad band(s) centered at about 230 nm. The slightly stronger absorbance by **3** between 204 and 232 nm compared with **2** may reflect conformational effects on absorption by the conjugated acylhydrazone link.[31] The UV absorption band(s) attributable to the acylhydrazone are reflected in the CD spectra by positive mean residue ellipticity at 229 nm (**2**: $\Theta_{229} = 1064$) and 228 nm (**3**: $\Theta_{228} = 3990$) and negative mean residue ellipticity at 243 nm (**2**: $\Theta_{243} = -3005$) and 246 nm (**3**: $\Theta_{246} = -2030$).

The α-helix is characterized by a positive ellipticity at 190 nm and a double minimum at about 208 and 222 nm.[26,30] The CD spectra for **2** and **3** could include bands from both the hydrazone link and α-helix which would be difficult to separate since they would overlap. Also, the higher mean residue ellipticity for **3** compared with **2** at 228 and 246 nm suggests that the hydrazone link like the peptide bond contributes to CD spectra in a conformationally dependent manner that would be difficult to correct for. However, $[\Theta]_{190}$ which increases from −9161 (**1**) to 1857 (**2**) to 13003 (**3**) and a shifting minimum from 198 (**1**) to 202 (**2**) to 206 nm (**3**) are consistent with increasing α-helicity.

NMR Spectroscopy. Peptides **1**−**3** were principally analyzed by NMR spectroscopy. Several criteria have been developed for identifying α-helical peptides including upfield shifts in signals for main-chain amide NH protons, decreases in amide proton temperature coefficients and H−D exchange rates, decreases in $J_{\alpha N}$ coupling constants, and specific distance patterns in two-dimensional (2D) proton−proton nuclear Overhauser enhancement (NOE) spectra.[32,33] NOEs which appear as cross-peaks in 2D NOESY spectra provide the most important data since α-helices are characterized by sequential distances, $d_{NN}(i, i + 1)$, $d_{\alpha N}(i, i + 3)$, and $d_{\alpha\beta}(i, i + 3)$, which allow conformational analysis on a per residue basis.[33] NOE data are the least subject to mitigating factors that often complicate the interpretation of other spectral data and is considered the most reliable indicator of peptide structure in solution.[6,33,34]

NMR spectra were acquired on 1−20 mM peptides at pH 2.9, 22 °C. Although lower pH is expected to decrease the α-helical propensity of the peptide,[26] it optimally separates signals for the amide protons in NMR spectra which is important for the unambiguous assignments of NOEs. NMR studies on helical peptides are generally carried out at 5 °C [6] since the α-helix is significantly stabilized at lower temperatures.[26] However, low-temperature NMR requires specialized equipment that was not available to us. A comparison of chemical shifts and line shapes of the amide protons for 1 and 20 mM of **2** showed no discernible differences, ≤0.01 ppm, indicating that self-association is not significant at these concentrations. The resonance positions for all protons in **1**−**3** were assigned by using standard approaches to 1D ^{1}H NMR and 2D COSY, TOCSY, and ROESY (rotating-frame NOESY) spectra of undeuterated and deuterated analogues (see the Experimental Section and Supporting Information). Deuterated alanine analogues of the peptides (Table 2) were used to identify alanine amide proton signals for **1** and to isolate $d_{\alpha\beta}(i, i + 3)$ NOEs for **2**. The use of deuterated analogues of **2** had an additional advantage of reducing noise in the crowded regions of the ROESY spectra which enhanced signal/noise ratios for weak

(30) (a) Holzwarth, G.; Doty, P. *J. Am. Chem. Soc.* **1965**, *87*, 218−228. (b) Sears, D. W.; Beyechok, S. *Physical Principles and Techniques of Protein Chemistry;* Academic Press: New York, 1973; pp 460−466.

(31) Buchwald, M., Jencks, W. P. *Biochemistry* **1968**, *7*, 844−859.

(32) (a) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519−538. (b) Waltho, J. P.; Feher, V. A.; Merutka, G.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1993**, *32*, 6337−6347.

(33) Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715−740.

(34) Bradley, E. K.; Thomason, J. F.; Cohen, F. E.; Kosen, P. A.; Kuntz, I. D. *J. Mol. Biol.* **1990**, *215*, 607−622.

Table 3. Summary of NMR Data of **1**−**3**[a]

Acetyl	G	L	A	G	A	E	A	A	K	A	NH_c	NH_t
Temp.Coef. (δppb/K)	#	-6.9	-7.3	#	-4.3	-6.0	-6.3	-6.3	-6.2	-6.4	-5.0	-4.8
$J_{\alpha N}$,Hz	*	*	5.4	*	5.5	*	*	*	*	5.8	-	-

	[J	L	A	Z]	A	E	A	A	K	A	NH_c	NH_t
Temp.Coef (δppb/K)	-	-5.0	-7.0	-	-2.5	-5.9	-4.3	-4.4	-3.7	-4.8	-4.1	-5.6
$J_{\alpha N}$,Hz	-	6.1	6.5	-	4.8	6.1	5.3	*	*	5.9	-	-
$d_{NN}(i, i+1)$			NA					*	*			
$d_{\alpha N}(i, i+1)$			NA					*	*			
$d_{\beta N}(i, i+1)$			NA	NA				*				
$d_{\alpha N}(i, i+3)$												
$d_{\alpha\beta}(i, i+3)$												

	[J	L	P	Z]	A	E	A	A	K	A	NH_c	NH_t
Temp.Coef. (δppb/K)	-	-5.6	-	-	-1.0	-7.6	*	-3.1	*	-4.2	-3.7	-5.6
$J_{\alpha N}$,Hz	-	6.7	-	-	5.2	5.9	*	*	*	6.2	-	-
$d_{NN}(i, i+1)$			NA					*	*	*		
$d_{\alpha N}(i, i+1)$			NA					*	*	*		
$d_{\beta N}(i, i+1)$			NA					*	*			
$d_{\alpha N}(i, i+3)$												
$d_{\alpha\beta}(i, i+3)$												
$d_{\alpha\delta}(i, i+3)$												
$d_{\alpha N}(i, i+4)$												

[a] Temperature coefficients were determined from linear regression analysis of five measurements between 21−45 °C with $R > 0.996$; the data is plotted in the Supporting Information. # indicate that the G1/G4 αN cross-peaks for **1** overlap in the TOCSY spectra; temperature coefficients for G1 and G4 are similar to those for E6. NH_c and NH_t are the carboxyl terminal cis and trans carboxamide protons. $J_{\alpha N}$ coupling constants are from 1D NMR spectra with a resolution of 0.33 Hz/datapoint. * and dashed lines indicate overlapping signals. Hatched bars are for NOEs from cL2 NH to cP3 δ CH_2, from cL2 β to cP3 δ CH_2 and from Z4 N═CH to A5 NH. NA, not applicable. Strong, medium, and weak NOEs are identified by the height of the filled bars; thick lines correspond to strong NOEs thin lines to weak NOEs.

NOEs. $J_{\alpha N}$ coupling constants and amide proton temperature coefficients for **1**−**3** and NOEs for **2** and **3** are summarized in Table 3.

Amide NH Region. Portions of 1D NMR spectra for **1**−**3** are compared in Figure 4. Signals for the main-chain amide protons of **2** and **3** show major differences in chemical shifts from corresponding protons for **1** with signals for the AAKA-NH_{cis} amide protons shifting upfield. The trend in signals for **2** are accentuated in **3** (Figure 5). The upfield shifts for the AAKA-NH_{cis} amide protons are predicted for hydrogen bonding that accompanies α-helix formation[35] with the larger differences for **3** indicative of greater helix stabilization.

The chemical shifts for the amide protons for **2** and **3** that are located near the acylhydrazone link shift in a discordant manner (Figure 5) which as discussed later may reflect shielding effects from the cyclic portion of the molecule. Differences in the signal for the hydrazone N═CH proton are also evident (Figure 4). Coupling constants for the hydrazone proton change from **2** (3.6, 8.2 Hz) to **3** (2.5, 9.5 Hz), indicating that the linker adjusts its conformation in response to the substitution of alanine (**2**) with proline (**3**).[36] These changes in conformation at the amino terminal end of **2** and **3** are reflected by changes at the carboxyl terminal end. NMR signals for the two C-terminal carboxamide protons (Figure 4) converge from a difference of 0.50 (**1**) to 0.42 (**2**) to 0.38 ppm (**3**). Both coupling constants and signal convergence provide sensitive and easily measurable parameters for conformational changes that must be occurring through the full length of **2** and **3**.

Hydrazone Link. Secondary amides often form mixtures of cis and trans isomers that interconvert slowly on the NMR time scale and are distinguishable in NMR spectra.[37] The data, however, provides strong evidence that the acylhydrazone (main-chain C−N link) adopts the trans configuration which relegates the hydrogen bond mimic (N−N link) to the required "*cis*" position (Figure 1). First, only one set of NMR signals is observed for **2** and **3**, indicating that only one geometric isomer is present (Figure 4). Second, no NOE was observed between the cA3 α or β protons and the **Z** methylene protons for **2**, whereas strong $d_{\alpha\alpha}(i, i+1)$ NOEs are observed for cis peptide bonds.[38] Likewise, no NOE was found between the cP3 β protons and the **Z** α protons. It is possible that the hydrazone link could enforce a cis bond at another position, the most vulnerable being the cL2-cP3 peptide bond in **3**. The strong NOE between cL2 NH and cP3 δCH_2 for **3** (Figure 7), however, indicates that the cL2-cP3 peptide bond is trans.[39] There is also no evidence for cis isomerization along the main-chain peptide bonds in either the cyclic or linear portion of **2** or **3**.

Additional NOEs agree with the targeted α-helical conformation (Figure 1). Strong NOEs between the hydrazone N═CH proton and the **Z** methylene proton(s) (not shown) indicate that the vinyl N−N═CH proton is trans to the main-chain amide nitrogen. Also, an NOE between the N═CH proton and the A5 NH proton (Figure 6, **2** and **3**) and a medium strength cL2-cA3 $d_{\alpha N}(i, i+1)$ NOE (Figure 7, **2b**) with similar intensity to the cL2-cA3 $d_{NN}(i, i+1)$ NOE (Figure 6, **2**) support the structure.

Distances. Helices are characterized by sequential $d_{NN}(i, i+1)$, $d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$ NOEs.[33] ROESY spectra for **2** and **3** show $d_{NN}(i, i+1)$ NOEs, in every case where signal separation permits, for cL2-cA3, A5-E6-A7 and K9-A10-NH_{cis}, for the whole length of the peptide in accordance with helix stabilization (Figure 6). The strongest evidence for a helix is provided by the midrange $d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$ NOEs that span the length of **2** (Table 3). These NOEs span cL2-cA3-Z4 in the cyclic peptide and A5-E6-A7 in the appended peptide (Figures 7, 8a). Although $d_{\alpha\beta}(i, i+3)$ NOEs for the cL2 and cA3 α protons are unambiguous, those for the α protons of the -ZAEAAKA- extension are potentially overlapped with other NOEs. To isolate these signals, alanine α and β protons were deuterated (**2a)** and a spectrum acquired to identify the remaining K9 NOEs (Figure 8b). Proton pairs from A5 (α

(35) Reymond, M. T.; Huo, S.; Duggan, B.; Wright, P. E.; Dyson, H. J. *Biochemistry* **1997**, *36*, 5234−5244.

(36) In a further test, glycine was substituted for L-leucine in **2** to give [**J**GA**Z**]AEAAKA-NH_2. A comparison of NMR data for [**J**GA**Z**]AEAAKA-NH_2 ($J_{N=CH-CH2}$ = 6.0 Hz, 6.0 Hz; C-terminal carboxamide protons, Δppm = 0.48) with that for **1**−**3** confirms the flexibilty of the linker. It indicates that glycine relaxes the conformation of the NucSite and reduces its ability to propagate a helix. Cabezas, E.; Satterthwait, A. C., unpublished data.

(37) Grathwohl, C.; Wüthrich, K. *Biopolymers* **1976**, *15*, 2025−2041.

(38) Chazin, W. J.; Wright, P. E. *J. Mol. Biol.* **1988**, *202*, 623−636.

(39) Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley: New York, 1986; pp 122−125.

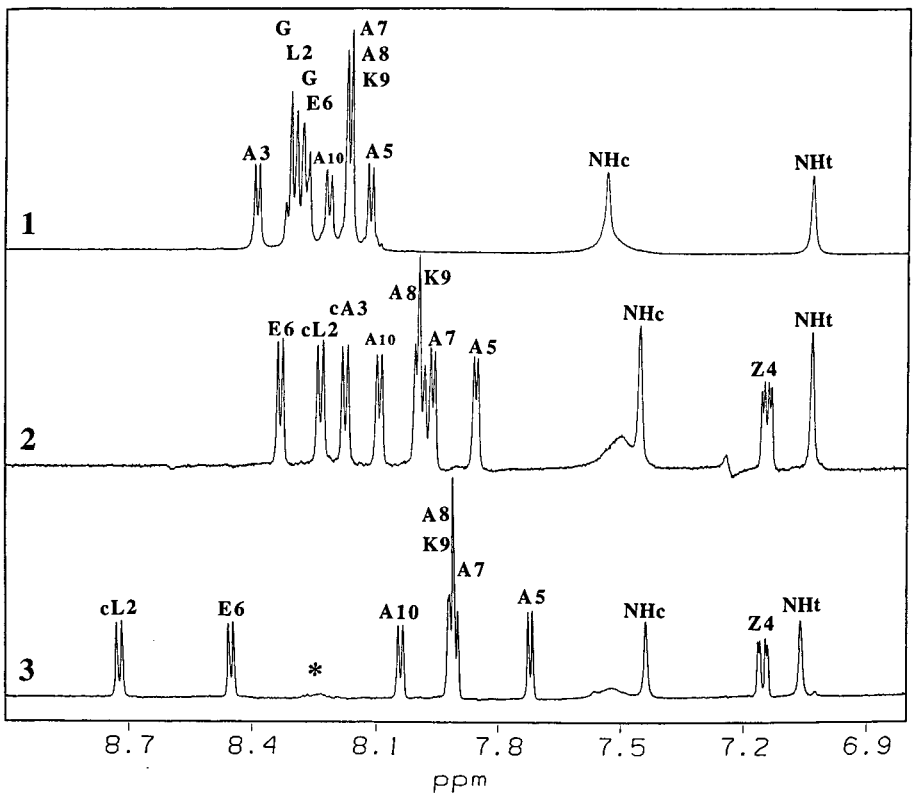


Figure 4. Amide proton region of 1D NMR spectra of **1**−**3** in 10% D_2O/H_2O at pH 2.9, 22 °C. Main-chain amide proton signals are identified for each amino acid according to the numbering scheme in Table 2. The signal for **Z**4 is from the N=CH proton. NH_c and NH_t are for the C-terminal cis and trans carboxamide protons.* Hydrolysis product that forms slowly at pH 2.9.

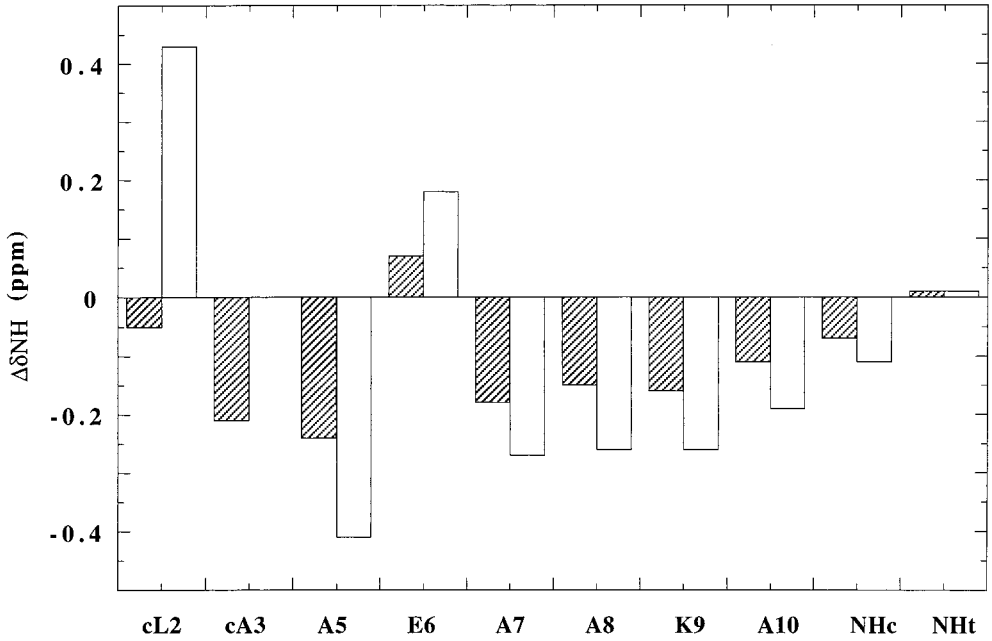


Figure 5. Differences in the chemical shifts for the main-chain amide protons of **2** (shaded bars) and **3** (open bars) compared with those for **1**. Data are from the spectra in Figure 4

proton) and A8 (β protons) and from A7 (α proton) and A10 (β protons) were then added back to give respectively **2b** and **2c** (Table 2) to unambiguously identify $d_{\alpha\beta}(i, i + 3)$ NOEs for these proton pairs (Figure 8c,d).

Midrange $d_{\alpha N}(i, i + 3)$ and $d_{\alpha\beta}(i, i + 3)$ NOEs that span cL2-cP3-Z4 and A5-E6-A7 were also observed for **3** (Figure 7,9) indicating that helix nucleation occurs in **3** in the same manner as in **2**. Overlapping signals prevented the assignment of midrange NOEs within the appended peptide of **3** except for a weak NOE that was observed between E6 αCH and K9 δCH_2. However, a similar pattern of upfield shifts in signals for the amide protons for AAKA-NH_{cis} indicates that **3** propagates a helix in the same way as **2** (Figure 5).

The 3_{10}-helix, α-helix and hybrid 3_{10}-α-helices are closely related structures that are difficult to distinguish by NMR spectroscopy.[20e,40] However, a comparison of distances for the

(40) Osterhout, J. J.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1989**, *28*, 7059−7064.

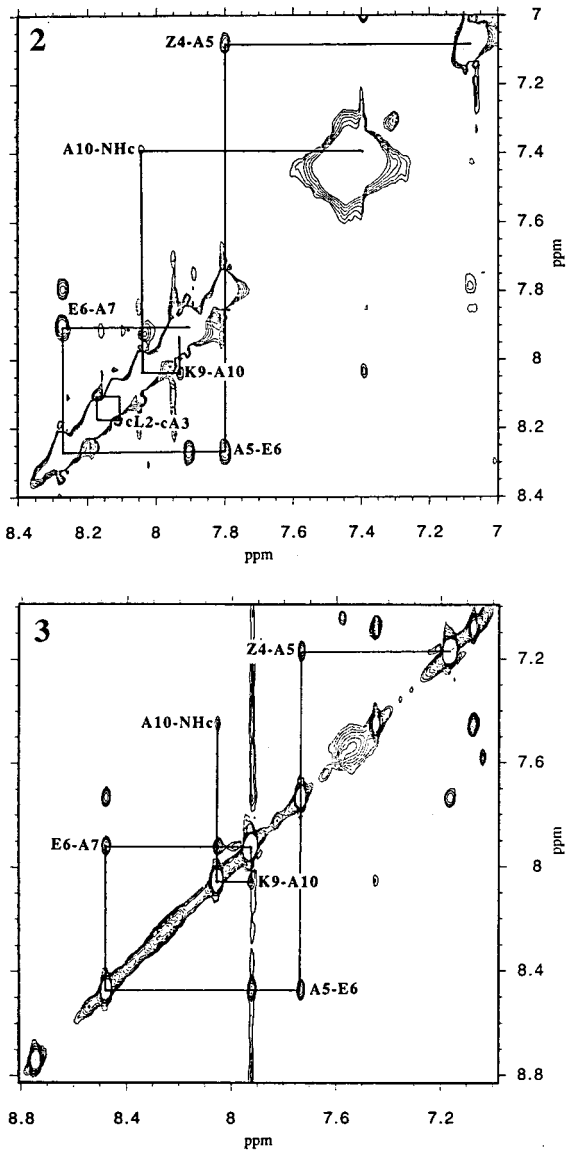


Figure 6. The amide proton region from a ROESY spectra of **2** in 10% D_2O/H_2O at pH 2.9, 19 °C (upper panel) and **3** at 22 °C (lower panel). Cross-peaks are for $d_{NN}(i, i+1)$ NOEs identified according to the numbering scheme in Table 2. The **Z**4-A5 NOE is for $Z4^{N=CH}$-$A5^{N}$.

3_{10}- and α-helix lead to several predictions[20e,41] that can be tested with the NOE data for **2** and **3**. Two sets of distances for the 3_{10}- and α-helix, derived from similar models, were converted into relative NOE intensities (Table 4) for comparison. Since NOE intensities are to a first approximation proportional to $1/d^6$,[34] even small differences in models can affect the predicted NOE intensities. However, these uncertainties do not affect the following predictions.[41] (1) α-helices but not 3_{10}-helices are predicted to display very weak $d_{\alpha N}(i, i+4)$ NOEs. (2) 3_{10}-helices are predicted to display very weak $d_{\alpha N}(i, i+2)$ NOEs and weak $d_{\alpha N}(i, i+3)$ NOEs. $d_{\alpha N}(i, i+2)$ NOEs are predicted to be negligible for the α-helix. (3) For an α-helix, the relative intensities of NOEs are predicted to be $d_{\alpha\beta}(i, i+3) \geq d_{NN}(i, i+1) > d_{\alpha N}(i, i+3)$ NOEs. For a 3_{10}-helix, the predicted relative intensities are $d_{NN}(i, i+1) > d_{\alpha\beta}(i, i+3) \approx 1.5 \times d_{\alpha N}(i, i+3)$ NOEs.

The NOE data for **2** and **3** strongly support an α-helical structure with no indication of a 3_{10} structure. First, a very weak

(41) Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley: New York, 1986; p 162−166.

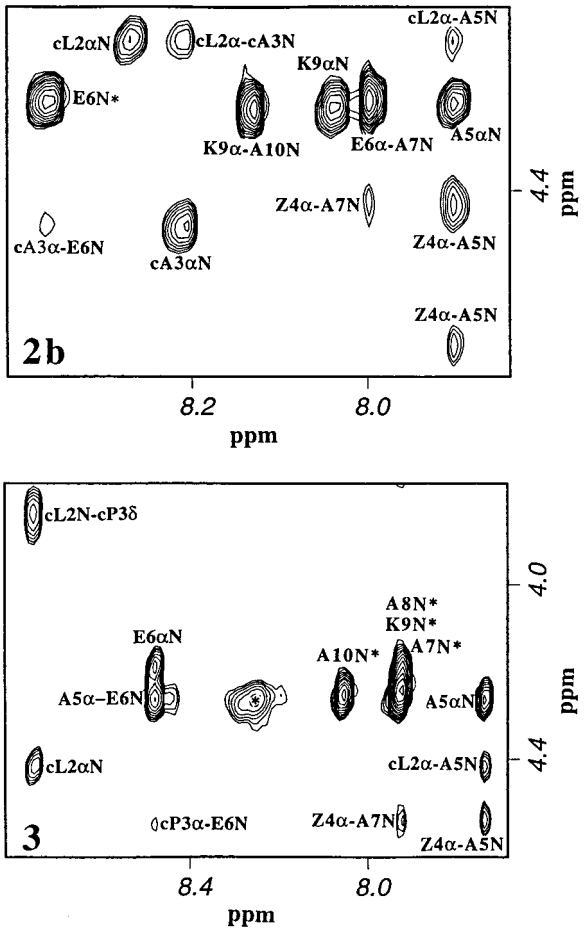


Figure 7. The fingerprint region from a ROESY spectra of **2b**, [**J**LA**Z**]-$A^3EA^4A^1KA^4$-NH_2 (upper panel) and **3** (lower panel) in 10% D_2O/H_2O at pH 2.9 at 22 °C. ***** indicates signals from a small amount of hydrolysis product of **3** that forms at low pH. N* indicates overlapping $d_{\alpha N}$ and $d_{\alpha N}$ $(i, i+1)$ NOEs to the indicated NH.

$d_{\alpha N}(i, i+4)$ NOE for cL2-E6 (**3**) is observed (see Supporting Information). Second, for **2** and **3**, $d_{\alpha\beta}(i, i+3)$ NOEs for cL2-A5, cA3/cP3-E6, and Z4-A7 are stronger than $d_{NN}(i, i+1)$ NOEs for A5-E6-A7 and far stronger than the corresponding $d_{\alpha N}(i, i+3)$ NOEs (Figures 7−9, see Supporting Information for stack plot for **3**). Overlapping signals prevented a determination of whether $d_{\alpha N}(i, i+2)$ NOEs were present or absent. Although $d_{NN}(i, i+1)$ NOEs can potentially arise from unfolded forms and other conformers,[32b] any contributions from these sources would be additive. This would lower the observed ratio of $d_{\alpha\beta}(i, i+3)/d_{NN}(i, i+1)$ NOEs. Instead, $d_{\alpha\beta}(i, i+3)$ NOEs are more intense than $d_{NN}(i, i+1)$ NOEs, indicating that the α-helix is a major conformer. Since strong $d_{\alpha\beta}(i, i+3)$ NOEs span [**J**L(A/P)**Z**] and -A5-E6-A7- of the appended peptide, the observed NOE ratios constitute strong evidence for α-helix nucleation with no apparent distortion from the acylhydrazone link. Similarly, strong $d_{\alpha\beta}(i, i+3)$ NOEs for A5-A8 and A7-A10 in **2** indicate that the α-helix is propagated to the carboxyl terminal end.

In addition to the evidence for an α-helix, the observation of strong $d_{\alpha N}(i, i+1)$ NOEs for -AEAAKA-NH_{cis} for **2** and **3** (Table 3, Figure 7) indicates that a fraction of the appended peptide is unfolded.[34] Strong $d_{\alpha N}(i, i+1)$ NOEs are characteristic of β-extended forms (Table 4) and not predicted for an α-helix.[33] The NMR data for the constrained peptide is most readily interpreted in terms of a model that includes unfolded

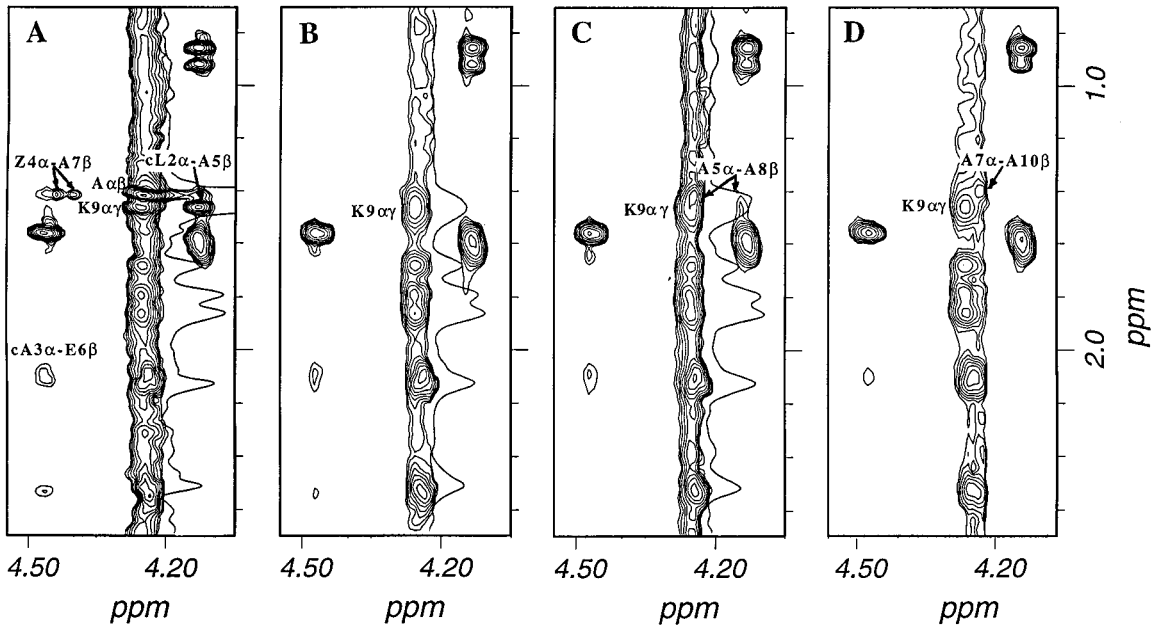


Figure 8. The $\alpha\beta$ region from ROESY spectra for **2** and deuterated analogues of **2** in 10% D_2O/H_2O at pH 2.9 at 22 °C. A 1D slice centered on the column for the K^{α} proton is superimposed on NOEs from the column to indicate the relative strengths of the NOE cross-peaks. A: **2**, identifies $cL2^{\alpha}$-$A5^{\beta}$, $cA3^{\alpha}$-$E6^{\beta}$, and $Z4^{\alpha}$-$A7^{\beta}$ NOEs. B: **2a**, [**J**LA**Z**]$A^4EA^4A^1KA^4$-NH_2, identifies the $K^{\alpha\gamma}$ NOEs by eliminating overlapping NOEs. C: **2b**, [**J**LA**Z**]$A^3EA^4A^1KA^4$-NH_2, identifies the $A5^{\alpha}$-$A8^{\beta}$ NOE. D: for **2c**, [**J**LA**Z**]$A^4EA^3A^4KA^1$-NH_2, identifies the $A7^{\alpha}$-$A10^{\beta}$ NOE.

Table 4. Distances in Å Between Designated Pair of Protons for Polyalanine in an Antiparallel β-Strand, 3_{10}-Helix and α-Helix[a]

distances, Å	β-strand[a]	3_{10}-helix[a]	3_{10}-helix[b]	α-helix[a]	α-helix[b]
$d_{NN}(i, i+1)$	4.3 (1.00)	2.6 (1.00)	2.59 (1.00)	2.8 (1.00)	2.57 (1.00)
$d_{\alpha N}(i, i+1)$	2.2 (55.00)	3.4 (0.20)	3.47 (0.17)	3.5 (0.26)	3.51 (0.15)
$d_{\alpha N}(i, i+2)$		3.8 (0.10)	3.79 (0.10)	4.4 (0.07)	4.32 (0.04)
$d_{\alpha N}(i, i+3)$		3.3 (0.24)	3.42 (0.24)	3.4 (0.31)	3.43 (0.17)
$d_{\alpha N}(i, i+4)$			5.54 (0.01)	4.2 (0.09)	4.18 (0.05)
$d_{\alpha\beta}(i, i+3)$		3.1 (0.35)	3.25 (0.26)	2.5 (1.97)	2.65 (0.83)

[a] Data is from Wüthrich et al.[33]; the $d_{\alpha\beta}(i, i+3)$ distance is for the closest approach of a β-methyl proton. [b] Based on helices generated by MSI Insight soft-ware. The alanine β-methyl protons are staggered relative to the alanine α-proton. The distance $d_{\alpha\beta}(i, i+3)$ is a weight-averaged distance to the three alanine β-methyl protons. Predicted intensities, in brackets, are relative to $d_{NN}(i, i+1)$ and based on $1/d^6$.

states in dynamic equilibria with folded states, including a full-length α-helix.

Dihedral Angles. Since $J_{\alpha N}$ coupling constants follow the Karplus equation, they are predicted to fall from >6 Hz for randomly coiled peptide (9 Hz for the extended chain) to about 4 Hz for an ideal α-helix.[42] The $J_{\alpha N}$ coupling constants for cL2 (6.1 Hz) and cA3 (6.5 Hz) in **2** are higher than predicted and indicate that the carbonyl oxygen atoms for **J**1 and cL2 tilt out about 20° on a weight-averaged basis from the axis of an ideal α-helix.[42] This tilt could reflect contributions from nonhelical forms of **2** which are not stabilized by hydrogen bonding. In **3**, proline ($\phi \approx -60°$)[15d] replaces alanine and shifts the equilibrium further toward the α-helix. The accommodation of a bulky L-proline with a constrained ϕ angle by **3** provides an important test for the targeted structure (Figure 1) that is well met. Proline's ability to enhance α-helix nucleation is probably due in part to better alignment of cL2 carbonyl oxygen along the helix axis.

The $J_{\alpha N}$ coupling constants for the appended peptide on **2** and **3** are not significantly reduced from those observed in **1** (Table 3). However, coupling constants are weight-averaged values for mixtures of different conformers that could contribute in different ways to the observed values. It has been noted before that $J_{\alpha N}$ values are the least sensitive NMR parameter for detecting structure.[32]

(42) Calculated on the basis of the Karplus relationship that applies to ϕ angles in proteins; Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley: New York, 1986; pp 166−168.

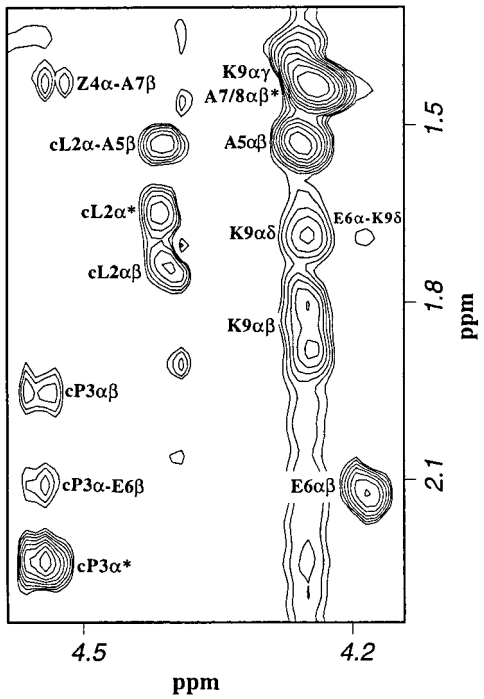


Figure 9. The $\alpha\beta$ region from a ROESY spectrum of **3** in 10% D_2O/H_2O at pH 2.9 at 22 °C. Identifies $cL2^{\alpha}$-$A5^{\beta}$, $cP3^{\alpha}$-$E6^{\beta}$, $Z4^{\alpha}$-$A7^{\beta}$ and a weak $E6^{\alpha}$-$K9^{\gamma}$ NOEs. α* indicates overlapping $\alpha\beta$ and $\alpha\gamma$ NOEs. $\alpha\beta$* indicates overlapping A7 and $A8^{\alpha\beta}$ protons.

Temperature Coefficients. α-Helices are characterized by sequential (i, $i+4$) hydrogen bonds that are associated with lowered H−D exchange rates, upfield shifts in signals for main-chain amide protons and decreases in the magnitude of amide proton temperature coefficients (<6 ppb/K).[32,35] Temperature

coefficients for **1**−**3** are summarized in Table 3; the data is plotted in the Supplementary Information.

The temperature coefficients for **1** are ≥ 6 ppb/K except for the A5 amide proton (4.3 ppb/K). This indicates that a fraction of the A5 amide proton is sequestered, probably by hydrogen bonding. The structure formed by **1** may be more local than propagated, although a small amount of helix cannot be ruled out.[43] The upfield shifts in signals for the amide protons of -AAKA-NH$_{cis}$ for **2** and **3** (Figure 5) are accompanied by lowered temperature coefficients from approximately 6.3 ppb/K (**1**) to about 4 ppb/K (**2**, **3**) as expected of a peptide that has been stabilized as an α-helix. However, changes in chemical shifts and temperature coefficients for the preceding A5-E6 amide protons diverge from the norm (Figure 5). Specifically, the signal for the E6 amide proton moves downfield, and its temperature coefficient decreases slightly from 6.0 (**1**) to 5.9 (**2**) and then **increases** to 7.6 ppb/K (**3**). This would normally indicate that the E6 amide proton of **3** increases its exposure to water compared to **1**, which contradicts the conclusion from the NOE data that **3** nucleates α-helix formation. The corresponding data for A5 shows large upfield shifts in the amide proton signals (Figure 5) and a reduction in temperature coefficients to low values of 2.5 (**2**) and 1.0 ppb/K (**3**). This trend to divergence for the A5-E6 protons is accentuated further for a related peptide in 2,2,2-trifluoroethanol.[44] However, H−D exchange rates for A5-E6 amide protons were **both** depressed to identical low values in 2,2,2-trifluoroethanol.[44] Thus the large differences in temperature coefficients for the A5-E6 amide protons do not necessarily reflect differences in their exposure to solvent. Instead, it is probable that the chemical shifts for the amide protons of A5-E6 (**2**, **3**) are influenced by shielding effects from the hydrazone link. In an α-helical conformation the A5-E6 amide protons abut the cyclic peptide. This might account for downfield shifts in the signal for the E2 amide proton and enhanced upfield shifts in signals for the A5 amide proton. The higher temperature coefficients for E6 amide protons (**2**, **3**) could then arise from the added effect of a decrease in α-helix and shielding as temperature is raised (enhanced upfield shifts with increasing temperature) while an opposite effect is predicted for A5 amide protons (dampened upfield shifts).

The amide protons for -AAKA-NH$_2$ (**2** and **3**) which lie beyond the shielding effects of the cyclic peptide shift upfield as predicted for an α-helix. This agrees with the NOE data that indicates [**J**L(A/P)**Z**]AE nucleates α-helix formation. If the N termini of **2** or **3** were nonhelical, the observed upfield shifts for the -AAKA-NH$_2$ amide protons would not be expected. Data for the amide proton chemical shifts and temperature coefficients are thus consistent with increasing α-helix stabilization for **1**−**3**. Indeed, the divergence in chemical shifts and temperature coefficients for the A5-E6 amide protons which increases from **2** to **3**, more likely indicates the extent of α-helix nucleation.

(43) Percent α-helix is generally calculated from the mean residue ellipticity, Θ_{222}, and an extinction coefficient based on the length of the peptide (21). The reference state for 0% helix is $\Theta_{222} = 0$. Since $\Theta_{222} = -1268$ for **1** at pH 2.9, 22 °C, **1** forms 4% α-helix by the standard calculation: calculated from a ratio of $\Theta_{222}/\Theta_{max}$, where $\Theta_{max} = -39\,500 \times [1 - (2.57/n]$ ($= -29\,349$ for $n = 10$).

(44) A comparison of chemical shifts, temperature coefficients and H−D exchange rates for the amide protons of acetyl-A5-E6(OEt)-E7(OEt)-E8-(OEt)-E9(OEt)$_2$ (**4**) was made with those for [**J**LA**Z**]-A5-E6(OEt)-E7(OEt)-E8(OEt)-E9(OEt)$_2$ (**5**) in 2,2,2-trifluroethanol and 2,2,2-trifluroethanol-d_1. The E6 amide proton of **5** shows a large disproportionate downfield shift relative to other amide protons when compared with **4**. The amide proton temperature coefficients for -AEEEE (**5**) are 0, 7.8, 4.0, 3.2, and 3.9 ppb/K respectively; Arrhenius, T.; Satterthwait, A. C., unpublished observations. The relative deuterium exchange rates for -AEEEE (**4**) are >10, >10, >10, >10, 10, respectively, while the corresponding rates for **5** are 1, 1, 2, 2, and 5, respectively.[16]

Table 5. Data is for **1**−**3** at pH 2.9, 22 °C[a]

parameters	peptide **1**	peptide **2**	peptide **3**
CD Spectral Data			
1. $[\Theta]_{190}$, deg·cm^2/dmol	−9161	1857	13003
2. λ_{min}, nm	198	202	206
NMR Spectral Data			
1. δ NH upfield shift, Δppm			
A7 NH		0.18	0.27
NH$_{cis}$		0.07	0.11
2. $J_{N=CH-CH2}$, Hz		3.6, 8.2	2.5, 9.5
3. C-terminal carboxamide protons, Δppm	0.50	0.42	0.38
4. $R_{NN} = d_{NN}(i, i+1)/d_{\alpha N}(i, i+1)$			
for A5-E6			0.39
for E6-A7		0.23	
for A10-NH$_{cis}$			0.19
5. $R_{\alpha\beta} = d_{\alpha\beta}(i, i+3)/d_{\alpha N}(i, i+1)$			
for cP3$^{\alpha}$-E6$^{\beta}$/A5$^{\alpha}$-E6^{N}			1.15

[a] CD spectral data are from Figure 3. Differences in the chemical shifts for the amide protons for **2** and **3** are relative to **1**; data is from Figure 5. Coupling constants are from spectra in Figure 4. The differences in chemical shifts for the carboxamide protons were derived from temperature dependence data. The distance data are based on NOE volumes from spectra in Figures 6−9.

Correlates of α-Helicity. A comparison of the CD and NMR data for **1**−**3** reveals several measures that indicate or correlate with increasing α-helicity and provide measures for estimating the extent of α-helicity. These data which were noted above are compared in Table 5. Increasing $[\Theta]_{190}$ and λ_{min} in the CD spectra are characteristic of increasing α-helicity.[30] Upfield shifts in signals for the -AAKA-NH$_{cis}$ amide protons are characteristic of α-helical peptides.[35] Coupling constants for the hydrazone N═CH proton diverge, whereas chemical shift differences between the two C-terminal carboxamide protons converge, reflecting conformational change over the full-lengths of **2** and **3**. The $d_{NN}(i, i+1)$ and $d_{\alpha\beta}(i, i+3)$ NOE intensities, representative of an α-helix, increase relative to $d_{\alpha N}(i, i+1)$ NOE intensities which are representative of unfolded forms.[32,34] The observation that $d_{\alpha\beta}(i, i+3)$ NOE intensities are similar to a $d_{\alpha N}(i, i+1)$ intensity (Table 5, stack plots for **3** in Supporting Information) despite the shorter $d_{\alpha N}(i, i+1)$ for extended forms (Table 4) suggests a very considerable fraction of α-helix. The general trend of the data indicates a progressive shift for **1**−**3** toward a larger fraction of α-helix. It indicates that the substitution of L-proline for L-alanine in **2** to give **3** improves α-helix stabilization by up to 2-fold. The decrease in upfield chemical shifts for the A10-NH$_{cis}$ amide protons (Figure 5, Table 5) indicates that the α-helix frays at the end.

Estimates of the percentage of helix are generally based on CD measurements − the mean residue ellipticity at 222 nm, $[\Theta]_{222}$.[21,26] For **1**, $[\Theta]_{222} = -1268$. This could indicate a minor population of α-helix.[43] However, the determination of the percentage of helix from low ellipticity values is very dependent on a reference value for the unfolded or "coiled" state. It is not clear what this reference value should be for **1**. Different alanine-based peptides level off at different negative values of $[\Theta]_{222}$ as temperatures are raised, which suggests that the reference value for 0% helix is peptide-dependent.[26] In addition, the basic relationship between CD and helicity is not well understood.[45] A theoretical treatment by Manning et al.[45] suggests that a minimum length (2−3 turns) is required for an α-helix CD spectrum, that $[\Theta]_{222}$ is length-dependent for short peptides, and that intensities are significantly reduced by outward tilting of the carbonyl oxygen group from the helix axis. These factors could explain several differences between CD spectra and other

(45) . Manning M. C.; Illangasekare, M.; Woody, R. W. *Biophys. Chem.* **1988**, *31*, 77−86.

measurements.[6a,45] For **2** and **3**, additional problems arise, since the acylhydrazone link very likely contributes to $[\Theta]_{222}$ in a conformationally dependent manner that would be difficult to correct for.

The CD data leaves uncertain the status of **1** and the degree to which the hydrazone link stabilizes the peptide as an α-helix. The NMR data for **1**−**3**, however, identifies a new measure that is precise, easy to acquire, and very sensitive to changes in α-helical content. The decrease in Δppm for the two C-terminal carboxamide amide protons for **1**−**3** is striking (Table 5). Since **1**−**3** differ at their N termini, far removed in space from the C termini, the changes in the carboxamide protons are almost certainly related to the degree to which a full-length α-helix is formed. As the percentage of full-length helix increases from **1**−**3**, the difference in the chemical shifts for the carboxamide protons decreases from 0.50 to 0.38 ppm. Consequently, this simple measure can be used to detect small changes in the amount of full-length α-helix.

This prompted us to measure the differences in chemical shifts for the C-terminal carboxamide protons of acetyl-AEAAKA-NH_2. A small increase from Δppm = 0.50 ppm for **1** to Δppm = 0.52 ppm for acetyl-AEAAKA-NH_2 was observed for measurements made under the same conditions. This small difference, ΔΔppm = 0.02, indicates that acetyl-GLAG has very little effect on the C-terminal end of **1**. It provides a reference for measuring the effectiveness of [**J**LA**Z**]- and [**J**LP**Z**]- at propagating a full-length α-helix. The absolute change in ΔΔppm for the C-terminal carboxamide protons increases from **1** (0.02) to **2** (0.10) to **3** (0.14) indicating significant stabilization of the peptide as an α-helix in aqueous solution at 22 °C by the hydrogen bond mimic.

Discussion

The development of a solid-phase synthesis for replacing main-chain hydrogen bonds with the hydrazone link reduces the assembly time of α-helical peptides from 6 weeks to 3 days and improves overall yields at least 10-fold. The assembly process is also amenable to automation and multiple peptide syntheses. It requires little more effort to assemble an α-helical peptide than a peptide. Although yields depend on the individual peptide, the solid-phase synthesis makes accessible constrained peptides that were previously inaccessible. This development provides the wherewithal for assessing whether the hydrazone link is a hydrogen bond mimic and whether it can be used in a general sense to constrain peptides to a range of conformations defined by different hydrogen-bonding patterns (Table 1).

Hydrazone Link. Conformational analysis of **1**−**3** by NMR spectroscopy clearly indicates that the hydrazone link stabilizes a peptide as a full length α-helix in water. These studies further establish that the −N═CH− bond, like the hydrogen bond, positions itself "*cis*" to the main-chain peptide bond (Figure 1). NOE patterns for amino acids that span the hydrazone link (Table 3) indicate that it fits well into an α-helix in agreement with model building. A further indicator of the preference of the mimic for a "*cis*" position is provided by **3**. Whereas X-Pro peptide bonds form cis,trans equilibria,[37] the Leu-Pro peptide bond shifts to an all-trans geometry in **3** indicating that the hydrazone link satisfies the need for a cis bond better than L-proline. The hydrazone link also positions itself "*cis*" in larger loops,[17] demonstrating that this preference in **2** and **3** is not dictated by the small ring size. Presumably, this preference is steric; the −N═CH− link is less bulky than $-CH_2-CO-$.

An important property of the hydrogen bond is its unique flexibility. This provides the hydrogen bond with a capacity to adapt to different angles and lengths required by the different conformations that it defines in proteins.[14] The propylene extension of the hydrazone is also flexible and provides the linker with a capacity to adjust to different amino acids and peptide conformations.[17,36] While the −N═CH− bond length and angles are fixed and differ somewhat from the mean values for hydrogen bonds in proteins,[24] its preference for the "cis" position, small size, and the flexibility of the linker are sufficient for satisfying many of the properties required of a hydrogen bond mimic.

α-Helix Nucleation. The observation that a hydrazone link stabilizes an α-helix implies that enforcing one hydrogen bond is sufficient for stabilizing an α-helical conformation in water. The hydrazone link creates a nucleation site (NucSite) overcoming the most significant energy barrier to α-helix formation in agreement with the Zimm−Bragg model for α-helix formation.[46] Once the initial turn of an α-helix is made, further turns are propagated. The propagation of the hexapeptide extension, -AEAAKA-NH_2, as a full-length α-helix in water sets a lower limit on the number of turns (<2) that can be propagated as a full-length α-helix in water.

NucSites are functionally related to templates.[2,20,21] However, they differ in design. Indeed, template designs differ, reflecting a rich diversity of possibilities. NucSites are formally characterized by a main-chain link that can encompass different amino acids. Templates include side-chain linked di-[20a−e] and triprolines[20f] and nonpeptidomimetic helix inducers.[2,21] The type and extent of helix propagation is a function of the number of hydrogen bond acceptors on the nucleation site, their disposition in space, and the strength of hydrogen bonds between the appended peptide and the nucleation site versus the strength of hydrogen bonding to solvent.

NucSites and templates share a common feature, three carbonyl oxygen groups. Three hydrogen bond accepting groups can potentially propagate in addition to the α-helix, several other helices including the 3_{10}-, hybrid 3_{10}-α and mixed 3_{10}-α helices.[15d,47] The 3_{10}-helix which is often found at the carboxyl termini of α-helices in proteins is characterized by sequential ($i + 3 \rightarrow i$) hydrogen bonds and is more tightly wound.[15d,47a] One consequence of this is that side chains are projected in different orientations.[15d,47a] Although the α-helix is generally more stable,[47a] interconversion between the α-helix and 3_{10}-helix is readily achieved.[20c,47a] Since NucSites and templates present only three hydrogen bond accepting groups, a fourth hydrogen bond must be accommodated to nucleate an α-helical turn. This can be achieved in different ways.

Kemp and co-workers have reported extensive studies on the templated peptide, Ac-Hel_1-Ala_6-OH, in a variety of solvents.[20c,e] Ac-Hel_1 is characterized by three amide carbonyl oxygens. Back-bonding from the first three amino acids appended to the template could lead to a 3_{10}- or 3_{10}-α-helical turn.[20c] To propagate an α-helical turn, the amide linking the template to peptide must forego hydrogen bonding to the template in order for the next three amide groups to hydrogen bond to the three accepting carbonyl oxygen atoms on the template.[20c] NOE data for Ac-Hel_1 in water indicated helix stabilization.[20e] In 20 mol % trifluoroethanol-water, NOEs that were characteristic of both a 3_{10} and α-helix were detected,[20e] suggesting that solvent can participate in α-helix propagation.

(46) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526−535.
(47) (a) Toniolo, C.; Benedetti, E. *Trends Biochem. Sci.* **1991**, *16*, 350−353. (b) Toniolo, C.; Benedetti, E. *Macromolecules* **1991**, *24*, 4004−4009. (c) Karle, I. L.; Flippen-Anderson, J. L.; Gurunath, R.; Balaram, P. *Protein Sci.* **1994**, *3*, 1547−1555. Karle, I. L.; Gurunath, R.; Prasad, S.; Rao, R. B.; Balaram, P. *Int. J. Pept. Protein Res.* **1996**, *47*, 376−382 (d) Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J. *J. Am. Chem. Soc.* **1996**, *118*, 2744−2745

Bartlett and coworkers[21] devised another solution. They observed little helix propagation by N-linked templates, **9-N**. However, when the amide group that linked the peptide, -AAEALAKA-NH2, to the template was replaced with an ester, the resulting **S,S-9-O** templated peptide displayed an α-helix-like CD spectrum in water. They attributed the lack of helix propagation by the **9-N** template to an initial nonproductive 3_{10}-like hydrogen bond. By replacing the linking amide with an ester, the initial hydrogen bond was eliminated, making way presumably for the following amino acids to hydrogen bond to carbonyl oxygens on the template—a carboxylate anion, carbonyl oxygen and ester carbonyl oxygen were available for the propagation of an α-helix.

A different approach to template design was taken by Müller et al.[2] They described two templates that incorporate the initial amide link to the appended peptide into the template, either as a secondary lactam or imide which replaced the initial amide hydrogen. These templates provide, in addition, either two carbonyl oxygens and a "relay" lactam carbonyl oxygen or two carboxylate oxygens and a "relay" imide carbonyl oxygen which are available for α-helix propagation.

NucSites establish an α-helix by replacing the initial N-terminal (i, $i + 4$) connection with an acylhydrazone link. The two amide carbonyl oxygens and a third "relay" carbonyl oxygen atom from the acylhydrazone (Figure 1) provide hydrogen bonding sites for the following three appended amino acids, -AEA. Thus, each of the four amide protons in the first propagated turn of an α-helix are accommodated.

α-Helix Stability. The average size of an α-helix in a globular protein is about 11 residues[48a] with 88% of helices from a representative sample of proteins between 4 and 15 amino acids in length.[25a] Although a few native peptides in this size class form detectable levels of α-helices in water at low temperatures (3−7 °C),[22h,49] many, including longer peptides (>15 residues) from the helical regions of proteins do not.[50] NucSites, templates,[2,20,21] and other methods[22,23] might therefore find use in stabilizing native peptides as α-helices. However, the amino acid sequences of protein helices are diverse and include helix-destabilizing as well as helix-stabilizing amino acids. The helix propensities of amino acids have been reviewed.[48b,51] Model studies with alanine-rich peptides suggest that very little helix formation is expected for peptides not rich in alanine.[48b] On the other hand, sequence-dependent effects have been observed to play important roles in helix stability, suggesting a complicated reality.[40,52] To sort out these effects, it may be important to stabilize native peptides as helices in water. Consequently, it is of interest that a parallel study indicates that NucSites stabilize a native peptide sequence as an α-helix.[53] Stabilizing native peptides as α-helices could aid studies on protein folding by identifying determinants of helix stability[49,54] and enhance the activities of peptides for specific reactions[8d,22a,55] or for general peptide screening.[4,5b,c]

Wider Role. In principle, a hydrogen bond mimic could be used to stabilize peptides in different conformations defined by position and sequence (Table 1). The solid-phase synthesis for insertion of the hydrazone link into a peptide provides a facile method for exploring this possibility. Medium-sized loops have been prepared by extending **Z** with increasing numbers and varied sequences of amino acids before capping with **J** as is being reported separately.[9,17]

Conclusion

The hydrazone link was substituted for an ($i + 4 \rightarrow i$) hydrogen bond at the N terminus of a short peptide stabilizing it as a full-length α-helix in water at ambient temperature as indicated by NMR spectroscopy. The link is stable for long periods under physiological conditions of pH and temperature. The advantages of the current method are that (1) it provides the first solid-phase synthesis of an α-helix nucleation site, (2) the synthesis is rapid, convenient, and capable of high yields, (3) the nucleation site can accommodate different amino acids, and (4) the same synthetic protocol can be combined with multiple peptide synthesis procedures to synthesize different structures simultaneously.

This work establishes the hydrazone link as a hydrogen bond mimic, that a hydrogen bond mimic is sufficient for stabilizing a biologically relevant conformation in water, and that it is flexible enough to allow different amino acid sequences to exert their natural conformational preferences.

Experimental Section

General. Acetone, acetonitrile, dichloromethane, *N*,*N*-dimethylformamide, *N*-methylpyrrolidone, methanol, and 2-propanol from Burdick and Jackson were of high quality and were used without further treatment. Ethyl hydrazinoacetate hydrochloride, 1-hydroxybenzotriazole hydrate (HOBt), *N*,*N*-diisopropylethylamine, 2,2,2-trifluoroethanol, deuterium oxide, chloroform-d_1, and ninhydrin were from Aldrich. *N*,*N*-diisopropylethylamine was purified by distillation and stored over 4-Å sieves. Trifluoroacetic acid (TFA) was from Pierce Chemical Co. Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexaflurophosphate (PyBOP) and 9-fluorenylmethyl chloroformate (Fmoc-Cl) were from NovaBiochem. *N*-α-Fmoc-L-amino acids, *N*,*N*-dicyclohexylcarbodiimide (DIC), and 4-(2′,4′-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy resin (Rink resin) were from Advanced Chem Tech. Fmoc corresponds to the 9-fluorenylmethyloxycarbonyl group. L-Alanine-2-d_1, l-alanine-3,3,3-d_3 and L-alanine-d_4 were from MSD Isotopes. The pH test strips 0−14.0 (P4536) were from Sigma. All amino acids are L-amino acids. Proton chemical shifts are relative to tetramethylsilane except were noted.

Resin packets were made from style 67221030 polypropylene fabric from Synthetic Industries, Gainesville, GA, by sealing with a TISH-300 Impulse sealer from San Diego Bag and Supply. Shaking was carried out on an Eberbach 6010 two-speed reciprocating platform shaker.

5,5-Dimethoxy-1-oxopentanoic Acid Methyl Ester. The compound was synthesized according to Stevens and Lee.[27]

5,5-Dimethoxy-1-oxopentanoic Acid (J). The compound was prepared by adding 5,5-dimethoxy-1-oxopentanoic acid methyl ester (10 g, 56.8 mmol) to a solution of sodium hydroxide (3.028 g, 75.7

(48) (a) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomed. Struct.* **1992**, *21*, 95−118. (b) Chakrabartty, A.; Baldwin R. L. *Adv. Protein Chem.* **1995**, *46*, 141−176.

(49) (a) Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470−2474.

(50) Reymond, M. T.; Merutka, G.; Dyson, H. J.; Wright, P. E. *Protein Sci.* **1997**, *6*, 706−716.

(51) Woolfson, D. N.; Alber, T. *Protein Eng.* **1995**, *4*, 1596−1607. Schneider, J. P.; Lombardi, A.; DeGrado, W. F. *Folding Des.* **1998**, *3*, R29−R40.

(52) Lesk, A. M. *Nature* **1991**, *352*, 379−380; Serrano, J.; Sancho, J.; Fersht, A. R. *Nature* **1992**, *356*, 453−455. Zerkowski, J. A.; J. A.; Powers, E. T.; Kemp, D. S. *J. Am. Chem. Soc.* **1997**, *119*, 1153−1154. Williams, L.; Kather, K.; Kemp, D. S. *J. Am. Chem. Soc.* **1998**, *120*, 11033−11043.

(53) A peptide (underlined) from MSP-1, a *Plasmodium falciparum* merozoite protein, was inserted into acetyl-GLAGALFQKEKMAKA-NH_2 and [**J**L(A/P)**Z**]ALFQKEKMAKA-NH_2. 1D and 2D NMR experiments, similar to those reported for **1**−**3**, show that the constrained peptide is stabilized as an α-helix in 10% D_2O/H_2O at pH 5.0 at ambient temperature: Cabezas, E. and Satterthwait, A. C., unpublished data.

(54) Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−4255. Renold, P.; Kwok, Y. T.; Shimizu, L. S.; Kemp, D. S. *J. Am. Chem. Soc.* **1996**, *118*, 12234−12235.

(55) Chorev, M.; Roubini, E.; McKee, R. L.; Gibbons, S. W.; Goldman, M. E.; Caulfield, M. P.; Rosenblatt, M. *Biochemistry* **1991**, *30*, 5968−5974.

mmol) in 100 mL of an equivolume mixture of methanol in water and stirred for 2 h at room temperature. The reaction mixture was rotoevaporated, and the residue was dissolved in a mixture of 50 mL of water, added to 100 mL of ethyl acetate, and cooled to ice temperature. The mixture was vigorously stirred and acidified to pH 2−3 with 1 N hydrochloric acid. The pH was monitored by spotting pH 0−14 test strips. The organic and aqueous layers were immediately separated, and the aqueous layer was extracted with 100 mL of ice-cold ethyl acetate. The combined organic layers were washed with 100 mL of ice-cold brine, dried over magnesium sulfate, filtered, and rotoevaporated to yield 8.2 g (88%) of a colorless oil. ^{1}H NMR ($CDCl_3$) δ 1.55−1.8 (m, 4), 2.2−2.6 (m, 2), 3.33 (s, 6), 4.39 (t, 1). Anal. Calcd for $C_7H_{14}O_4$ C, 51.8; H, 8.7. Found C, 50.58, 50.65; H 8.88, 8.75. FAB MS Calcd (MNa^+) 185.0790; found 185.0793. **J** is stable for >1 year at − 20 °C.

(1-Methylethylidene-2-Fmoc)hydrazinoacetic Acid (Fmoc-Z(Act). The compound was synthesized by dissolving ethyl hydrazinoacetate hydrochloride (15.4 g, 100 mmol) in 300 mL of an equivolume mixture of water and acetone, refluxing for 10 min, and cooling to room temperature. Sodium hydroxide (8.4 g, 210 mmol) was added to the reaction mixture and stirred for 30 min at room temperature to hydrolyze the ester. Sodium carbonate (10.6 g, 100 mmol) was dissolved in 50 mL of water and added to the reaction mixture, and the mixture was cooled to ice temperature. A solution of 9-fluorenylmethyl chloroformate (25.9 g, 100 mmol) in 50 mL of dioxane was added dropwise to the reaction mixture during 1 h with stirring. Stirring was continued for 15 h while the reaction mixture came to room temperature. The reaction mixture was rotoevaporated and the residue suspended in 200 mL of ice-cold ethyl acetate. This mixture was then vigorously stirred and acidified to pH 3 by adding 1 N HCl. The two phases were separated, and the aqueous phase was extracted with 200 mL of ethyl acetate. The ethyl acetate extracts were combined, extracted with brine, and dried with magnesium sulfate. After filtration and rotoevaporation to dryness, the product was dissolved in 100 mL of acetone, refluxed for 10 min, and again rotoevaporated to dryness. Crystallization from ethyl acetate with hexane yielded 31.4 g (80.7%) of white crystals: mp 146−148 °C; ^{1}H NMR (acetone-d_6) δ 1.5−1.75 (br s, 2), 1.7−2.0 (br s, 4), 4.19 (m, 2), 4.27 (m, 1), 7.33 (dd, 2, J = 7.4, 7.4), 7.42 (dd, 2, J = 7.4, 7.5), 7.68 (d, 2, J = 7.4), 7.87 (d, 2, J = 7.6), 11.06 (s, 1). Anal. Calcd for $C_{20}H_{20}N_2O_4$ C, 68.17; H, 5.72; N, 7.95. Found C, 67.97; H, 5.71; N, 7.85. FAB MS Calcd (MH^+) 353.1422; found 353.1511. Fmoc-**Z**(Act) is stable when stored dry at 4 °C.

***N*-α-Fmoc-L-alanine-2-d_1, *N*-α-Fmoc-L-alanine-3,3,3-d_3 and *N*-α-Fmoc-L-alanine-d_4.** The compounds were prepared according to Lapatsanis et al.;[56] mp 142−143 °C [*N*-α-Fmoc-L-alanine, lit. = 142−143 °C].

***N*-α-Fmoc-L-alanine chloride and *N*-α-Fmoc-L-proline chloride.** The compounds were prepared by the procedure of Carpino et al.[28] The crystalline products were dried under vacuum to remove trace amounts of hydrochloric acid. Both compounds were >97% pure by the methanol test.[28]

Fmoc-Z(Act)-Ala-Glu(*t*-But)-Ala-Ala-Lys(*t*-Boc)-Ala-Rink Resin. The compound was assembled on 4-(2′,4′-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy resin (Rink resin) (100 mg, 0.046 mmol) sealed in a 2 cm^2 polypropylene mesh packet.[57] The packet was included among several packets for simultaneous multiple peptide synthesis.[57] The resin packets were placed in capped 30-mL or 60-mL wide-mouth polyethylene bottles for common wash and deprotection steps. Wash solutions were added to the packets in the bottle, shaken for 1 min on a platform shaker, and poured off. Single packets were placed in capped 2-mL micro tubes for coupling reactions. Otherwise, when several packets required the same amino acid, they were combined and reaction solutions scaled up for coupling in a capped bottle. The following protocol reports the reagents required for a synthesis with one packet.

A standard coupling cycle involved the following steps: (1) The resin was swelled in its packet by shaking with 5 mL of dichloromethane and then washed with 5 mL of *N*,*N*-dimethylformamide. (2) The Fmoc-protecting group was cleaved from the protected peptide−resin with 5 mL of 20% piperidine in *N*,*N*-dimethylformamide for 10 min, and the resin was washed four times with 5 mL of *N*,*N*-dimethylformamide, once with 5 mL of 2-propanol, once with 5 mL of dichloromethane, and once with 5 mL of *N*-methylpyrrolidone. (3) Coupling was carried out with *N*-α-Fmoc-L-amino acid (0.092 mmol) or Fmoc-**Z**(Act) (32.5 mg, 0.092 mmol), 1-hydroxybenzotriazole (12.4 mg, 0.092 mmol), benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexaflurophosphate (47.9 mg, 0.092 mmol), and *N*,*N*-diisopropylethylamine (0.032 mL, 0.184 mmol) in 1 mL of *N*-methylpyrrolidone by shaking for 30 min. After the coupling reaction, the resin was washed twice with methylene chloride and air-dried. A small amount of resin was removed from the resin packet and tested for primary amine by the Kaiser ninhydrin test.[58] The coupling step was repeated if necessary.

J-Leu-Ala-Z(Act)-Ala-Glu(*t*-But)-Ala-Ala-Lys(*t*-Boc)-Ala-Rink Resin. Peptide assembly was continued from above in exactly the same manner except for the coupling (step 3) of Fmoc-L-alanine chloride to **Z**(Act). Since the acetone-protecting group on **Z**(Act) is labile to strong acids, it is critical that Fmoc-L-alanine chloride be freed of adventitious hydrochloric acid by briefly drying it under vacuum. *N*-α-Fmoc-L-alanine chloride (30.3 mg, 0.092 mmol) was dissolved in 1 mL of *N*-methylpyrrolidinone, added to the resin packet, and briefly shaken before adding 0.016 mL of *N*,*N*-diisopropylethylamine (0.092 mmol). Shaking was continued for 15 min. The Kaiser ninhydrin test is not applicable for this reaction. Fmoc-L-leucine (32.5 mg, 0.092 mmol) and then **J** (17 mg, 0.092 mmol) were coupled to the peptide resin using the coupling protocol for Fmoc-amino acids. Following the addition of **J**, the resin packet was washed four times with 5 mL of methylene chloride and air-dried.

[J-Leu-Ala-Z]-Ala-Glu(*t*-But)-Ala-Ala-Lys(*t*-Boc)-Ala-Rink Resin. The dry resin packet from the previous step was shaken for 15 min in a 15 mL solution of 20% 2,2,2-trifluoroethanol in methylene chloride that had been premixed with 0.02 mL 4 N hydrochloric acid in dioxane (5.3 mM final). The resin packet was washed four times with 5 mL of dichloromethane. A slight pink coloring of the resin was used as an indicator that sufficient acid had been added for cyclization.

[J-Leu-Ala-Z]-Ala-Glu-Ala-Ala-Lys-Ala-NH$_2$ (2). The cyclic peptide-resin from the previous step was added to 15 mL of 10% trifluoroacetic acid in dichloromethane in a capped polypropylene bottle and shaken for 15 min. The solution containing the cleaved protected peptide was removed and the packet treated again in the same manner. The acidic solutions were combined and rotoevaporated to leave a residue that was treated with 2 mL of trifluoroacetic acid for 15 min. Peptide was precipitated from the acid solution by adding 30 mL of cold diethyl ether. The precipitate was centrifuged, washed with cold diethyl ether and dissolved in 1 mL 10% acetonitrile in water.

The crude product was initially analyzed by high-pressure liquid chromatography (HPLC) on a Merck Lichrosorb analytical C-18 column (4.6 × 25 mm, 10 micron). Products were eluted with a 100% water−100% acetonitrile gradient in 0.1% TFA over 15 min at 2 mL/min. The chromatogram (Figure 2) showed one major product, **2**, and two minor products, **A** and **B**. These products were separated by HPLC on a Vydac preparative C-18 column (201TP1022, 2.2 × 25 cm, 10 μm) by eluting with a 100% water−60% acetonitrile gradient in 0.1% trifluoroacetic acid over 40 min at 5 mL/min. Fractions were lyophilized to yield 19.2 mg of **2** (47% yield, based on initial resin loading). **2** was identified by NMR spectroscopy and FAB-MS: Calcd (MH^+) 896; Obsd 896. HR FAB-MS: Calcd (M^+Cs^+) 1027.3977; Obsd 1027.4019

The **A** and **B** fractions from several HPLC runs were combined and purified further for analysis. **A** was identified as a dimer of **2** by low resolution (LR) FAB MS: Calcd (MH^+) 1791; Found 1791. **B** was identified as an isomer of **2** by FAB MS: 895.5001, Found 895.5109. ^{1}H NMR (10% D_2O/90% H_2O) of **A** but not **B** showed a signal at 7.21 ppm (dd, 1) that is characteristic of a hydrazone N═CH proton.

[J-Leu-Pro-Z]-Ala-Glu-Ala-Ala-Lys-Ala-NH$_2$ (3). The compound was synthesized by the same procedure as **2** but with *N*-α-Fmoc-L-proline chloride used in place of *N*-α-Fmoc-L-alanine chloride. Rink's resin (1 gm, 0.42 mmol) was sealed in six resin packets for this

(56) Lapatsanis, L.; Milias, G.; Froussios, K.; Kolovos, M. *Synthesis* **1983**, 671−673.

(57) Houghten, R. A.; De Graw, S. T.; Bray, M. K.; Hoffman, S. R.; Frizzell, N. D. *Biotechniques* **1986**, *4*, 522−528.

(58) As described in Stewart, J. M.; Young, J. D. *Solid-Phase Peptide Synthesis*, 2nd ed.; Pierce Chemical Co.: Rockford, IL, 1984; p 105.

synthesis. The coupling reactions were scaled up proportionately. The wash and reaction steps were carried out in 60-mL Nalgene wide-mouth polyethylene bottles. Cyclization was carried out by shaking the six resin packets for 15 min in 40 mL of 20% 2,2,2-trifluoroethanol in dichloromethane that had been premixed with 0.105 mL 4 N hydrochloric acid in dioxane. The product was cleaved from the resin by scaling up the procedure used for **2** by 10-fold.

Two major products were separated by HPLC with multiple runs on a Merck Lichrosorb semipreparative C18 column (10 × 250 mm, 7 μm) with a 100% water−100% acetonitrile gradient in 0.1% trifluoroacetic acid over 30 min at 3 mL/min. A dimer of **3** eluted at 19 min, while **3** eluted at 19.7 min. The fractions of **3** were combined, lyophilized, and purified by repeating the HPLC runs to yield 16.5 mg (4.3%); HR FAB-MS: Calcd (MH^+) 921.5158; Obsd 921.5172. The dimer of **3** was obtained in 23% yield from the first purification; LR FAB MS: Calcd (MH^+) 1842, Obsd 1842.

Acetyl-GLAGAEAAKA-amide (1). The compound was synthesized in a resin packet by using the coupling cycle and cleavage procedures described above. The peptide was acetylated on the resin by first removing the final Fmoc-protecting group with 20% piperidine in *N*,*N*-dimethylformamide and then washing and shaking the peptide-resin with 15% acetic anhydride in *N*,*N*-dimethylformamide for 30 min. The peptide was purified by HPLC as a single peak and identified by NMR spectroscopy and LR FAB MS: Calcd (MH^+) 899; Obsd 899.

Acetyl-AEAAKA-amide. The compound was synthesized and purified in the same manner as **1**.

Deuterated analogues, 1a−d and 2a−c. The compounds were synthesized in resin packets as described above. Resin packets were combined for common steps and separated for individual coupling reactions[57] with *N*-α-Fmoc-L-alanine-2-d_1, *N*-α-Fmoc-L-alanine-3,3,3-d_3, and *N*-α-Fmoc-L-alanine-d_4. Products were purified by HPLC and their compositions confirmed by comparing their 1D NMR spectra with the 1D NMR spectra for the fully characterized, undeuterated compounds.

Spectroscopy. Purified peptide samples were lyophilized at least twice to remove trace amounts of TFA, weighed, and dissolved in doubly distilled water or 90% H_2O/10% D_2O. The aqueous samples were adjusted to pH = 2.9 by adding small quantities of HCl and NaOH. UV spectra were recorded in a 1-cm quartz cell (1-cm path length) with a Varian Cary 3Bio UV/Vis spectrometer. CD spectra were recorded on samples in a 0.2-cm quartz cell (0.2-cm path length) with an Aviv model 62ADS circular dichroism spectrometer. The observed spectra were corrected by subtracting a spectrum for water. Molar ellipticities for **1**−**3** are calculated for 10 residues.

Most NMR spectra of peptides were acquired with a Bruker AMX-500 spectrometer. Except where noted, all spectra were recorded on approximately 20 mM peptide, pH 2.9 at a probe temperature of 22 °C. Chemical shifts are relative to internal dioxane at δ = 3.75 ppm.

1D NMR spectra in 90% H_2O/10% D_2O were run with presaturation of water. $J_{\alpha N}$ coupling constants for amide protons were determined at 0.34 Hz/data point after zero filling. 2D ROESY experiments were run in the phase-sensitive mode using TPPI for quadrature detection in f_1. In a typical experiment, 48 transients of 1024 data points were acquired in the f_2 dimension at a spectral width of 5555 Hz and 512 t_1 increments which were zero filled to 1024 data points.

2D ROESY spectra were acquired for **2** with mixing times of 100, 200, and 400 ms and compared with a 2D NOESY spectrum on the same sample. NOEs were far more intense in the ROESY spectra than in the NOESY spectrum, indicating that the product of the Larmor precession frequency and the correlation time ($\omega\tau_c$) is close to unity. Strong NOEs with high signal/noise ratios were proportional to the mixing time. The NOE data that is reported is from 2D ROESY experiments carried out with a mixing time of 400 ms.

Temperature coefficients were determined from 1D and 2D TOCSY experiments run on a Bruker AM-300 spectrometer. Spectra were acquired from samples in 90% H_2O/10% D_2O which was suppressed using a jump-and-return selective excitation sequence. Probe temperatures were determined by using a methanol standard.

1D and 2D data was processed and analyzed on a Silicon Graphics Iris Indigo XS24 computer using Molecular Simulation, Inc. (MSI) Felix 95 software.

Acknowledgment. We acknowledge Dr. Lin-chang Chiang for his initial synthesis of Fmoc-**Z**(Act). We thank Drs. Thomas Arrhenius and James Tam for helpful conversations. Dr. John Waltho aided us in obtaining NMR measurements. The manuscript benefited from reviews by Drs. Elena Dovalsantos and Philip Dawson and from thoughtful suggestions by the final reviewers. This work was supported by grants from the United States Agency for International Development (DPE-5979-A-00-1035-00) and the National Institutes of Health (AI19499). This is publication 11533-MB.

Supporting Information Available: Tables of ^{1}NMR chemical shift data for **1**−**3**, the amide proton region from 1D NMR spectra for **1a**−**d**, fingerprint region from a 2D TOCSY spectrum of **1**, 2D TOCSY spectrum of **2**, stack plots and a 1D slice from a 2D ROESY spectrum of **3** and plotted data for the determination of temperature coefficients for **1**−**3** (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA983212T