Design, Synthesis, and Characterization of 4-Ester CI2, a Model for Backbone Hydrogen Bonding in Protein α-Helices

Gangamani S. Beligere and Philip E. Dawson*

Contribution from the Departments of Cell Biology and Chemistry, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037

Received May 15, 2000

Abstract: The total synthesis of proteins enables unnatural groups to be incorporated into proteins to understand the molecular basis of protein stability and function. Chymotrypsin inhibitor 2 (CI2), is a small 64 residue protein consisting of an α-helix sandwiched by four β-strands. To directly evaluate the role of backbone hydrogen bonding in α-helices, an array of four amide bonds that span the length of the α-helix have been replaced with ester bonds. Residues 13, 16, 19, and 22 have been substituted with α-hydroxy acids using solid-phase synthesis, and the peptides were assembled by conformationally assisted ligation. The resulting 4-ester CI2 is a functional protease inhibitor that is destabilized by 2.93 kcal/mol compared to the all-amide protein. This study demonstrates that the formation of multiple consecutive hydrogen bonds is not required for the folding of protein α-helices.

Introduction

Structurally, polypeptides are often classified into two components, amino acid side chains and the polypeptide backbone. It is generally accepted that side chain packing and desolvation are critical elements in the folding and stability of α-helices in proteins,[1] but the role of the backbone remains more controversial.[2,3] In terms of stability, backbone hydrogen bonding has been described as being both stabilizing[4] and net destabilizing.[4b,c] In addition, while some protein folding models rely exclusively on side chain packing,[5] others describe hydrogen bonding as essential for the resolution of the Leventhal paradox[6] through a nucleation/propagation mechanism.[4c,7]

Chymotrypsin inhibitor 2 (CI2) is a well-characterized model system for studying the molecular basis of protein folding and stability.[8] This 64-residue protein conforms to a two-state folding model that has been analyzed by both equilibrium and kinetic methods.[8] Extensive studies by Fersht and co-workers on over 100 side chain mutants have described, for the folding of this protein, a nucleation site that includes a set of hydrophobic residues including Ala 16, Leu 49, and Ile 57 (Figure 1) and the N-terminal residues of the single α-helix (residues 12−24). However, due to the limitations of bacterial expression, these studies have used amino acid side chain substitutions to analyze the α-helix in the folding process.

The chemical synthesis of proteins enables a variety of noncoded elements of structure to be incorporated into proteins.[9] As a result, the polypeptide backbone can be modified to directly probe its effect on protein stability. Modification of the polypeptide backbone by ester substitution (depsipeptides) has been utilized in structure−function studies in small peptides[10] and more recently proteins.[11−13] Backbone ester bonds have many structural similarities to amide bonds; they strongly favor a *trans* conformation,[14] are planar, and have similar bond angles and lengths.[15,16] In addition, the accessible ϕ,ψ space has been shown to be similar for Ala-Lac and Ala-Ala.[14] The primary distinction between amide and ester bonds in the context of the polypeptide chain is that the ester carbonyl is a poor hydrogen bond acceptor[17,18] and the hydrogen bond donating NH is replaced with the electronegative −O− atom of the ester.[11,14,18]

Here a novel approach to directly test backbone interactions in the context of a protein is demonstrated with the design and synthesis of an analogue of CI2 in which three contiguous hydrogen bonds in an α-helix have been deleted. The 4-ester CI2 protein provides a model for studying the necessity of backbone hydrogen bonding in the folding of biopolymers.

(1) Dobson, C. M.; Sali, A.; Karplus, M. *Angew. Chem., Int. Ed.* **1998**, *37*, 868−893.
(2) (a) Honig, B.; Cohen, F. E. *Folding Des.* **1996**, *1*, R17−R19. (b) Srinivasan, R.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A* **1999**, *96*, 14258−63.
(3) (a) Zitzewitz, J. A.; Ibarra-Molero, B.; Fishel, D. R.; Terry, K. L.; Matthews, C. R. *J. Mol. Biol.* **2000**, *296*, 1105−1116. (b) Moran, L. B.; Schnneider, J. P.; Kentisis, A.; Reddy, G. A.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10699−10704.
(4) (a) For example: Murphy, K. P.; Gill, S. J. *J. Mol. Biol.* **1990**, *222*, 699−709. (b) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 351−365. (c) Honig, B. *J. Mol. Biol.* **1999**, *293*, 283−293.
(5) Dill, K. A. *Biochemistry* **1990**, *29*, 7133.
(6) Leventhal, C. *J. Chim. Phys.* **1968**, *65*, 44−45.
(7) (a) Zim, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526−535. (b) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963−1974.
(8) (a) Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254*, 260−288. (b) Neira, J. L.; Davis, B.; Ladurner, A. G.; Buckle, A. M.; Gay, G. Fersht, A. R. *Folding Des.* **1996**, *1*, 189−208.
(9) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.
(10) (a) Bramson, H. N.; Thomas, N. E.; Kaiser, E. T. *J. Biol. Chem.* **1985**, *260*, 15452−15457. (b) Arad, O.; Goodman, M. *Biopolymers* **1990**, *29*, 1633−49. (c) Coombs, G. S.; Rao, M. S.; Olson, A. J.; Dawson, P. E.; Madison, E. L. *J. Biol. Chem.* **1999**, *274*, 24074−24079.
(11) (a) Lu, W. Y.; Qasim, M. A.; Laskowski, M.; Kent, S. B. H. *Biochemistry* **1997**, *36*, 673−679. (b) Lu, W.; Randal, M.; Kossiakoff, A.; Kent, S. B. H. *Chem. Biol.* **1999**, *6*, 419−427.
(12) (a) Chapman, E.; Thorson, J. S.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 7151−7152. (b) Shin, I.; Ting, A. Y.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 12667−12668.
(13) England, P. M., Zhang, Y., Dougherty, D. A., Lester, H. A. *Cell* **1999**, 89−98.
(14) Wiberg, K. B.; E, L. K. *J. Am. Chem. Soc.* **1987**, *109*, 5935−5943.
(15) Ingwall, R. T.; Goodman, M. *Macromolecules* **1974**, *7*, 598−605.
(16) Ramakrishnan, C.; Mitra, J. *Proc. Indian Acad. Sci., Sect. A* **1978**, *87*, 13−21.
(17) Arnett, E. M.; Mitchell, E. J.; Murty, T. S. S. R. *J. Am. Chem. Soc.* **1973**, *96*, 3875−3891.
(18) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314−11322.

10.1021/ja001648e CCC: $19.00
Published on Web 11/28/2000

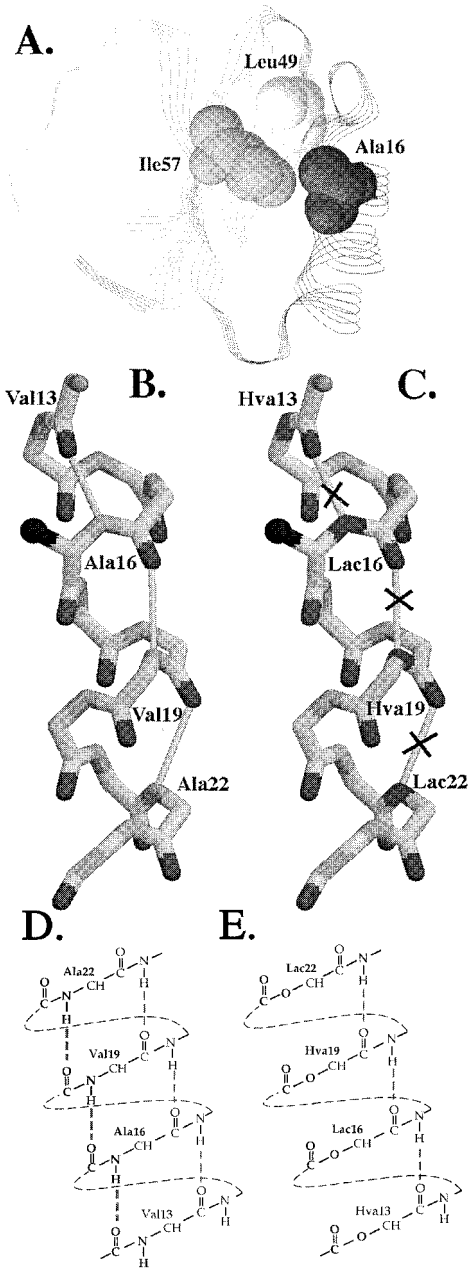

Figure 1. Illustration of CI2 conformation. (A) Space-filling representation emphasizing the hydrophobic core residues are shown, the diagram is produced using rasmol 2.6 by R Sayle(1CI2.pdb). (B and D) The CI2 α-helix(12−24) backbone showing H-bonds of interest. (C and E) The α-helix with ester incorporated in position 13, 16, 19, and 22, the H-bond being deleted are marked X.

Results and Discussion

Design. Incorporation of single ester bonds into proteins has been utilized to analyze the role of hydrogen bonding in protein−protein interactions,[11] protein secondary structure,[12,18] and transmembrane domains.[13] The single-ester substitution in an interior amide position of an α-helix or *β*-sheet results in a protein with an unpaired backbone amide group that is destabilizing to the protein.[11−13,18,19] In particular, the amide carbonyl is proximal to an electronegative ester -O- and the amide carbonyl is proximal to the ester carbonyl, a poor hydrogen bond acceptor. The 4-ester CI2 has been designed to avoid this limitation of single amide-to-ester substitutions. The α-helix of the CI2 has been selected to show that a contiguous array of hydrogen-bonded amides can be replaced with esters. In the context of an α-helix, this can be achieved through substitution of ester bonds in consecutive *i* to *i*+3 positions.[20] In this arrangement all ester carbonyl groups are paired with ester -O- atoms,[21] and no amide groups are left unpaired. Using these design principles, a series of helical residues, Val 13, Ala 16, Val 19, and Ala 22, have been replaced with their corresponding α-hydroxy acids to generate a 4-ester CI2 analogue.[22] This construct eliminates three contiguous *i* to *i*+4 hydrogen bonds, Ser12(CO)Ala16(NH), Glu15(CO)Val19(NH), and Lys18(CO)Ala22(NH).

Synthesis. Introducing α-hydroxy acids into peptides by solid-phase peptide synthesis (SPPS) to form depsipeptides has been described in several systems.[10,11] This approach has an advantage over in vitro translation methods[12,13,23] since multiple unnatural groups can be incorporated into a single polypeptide chain in a straightforward manner. To synthesize the 4-ester protein analogue, a thioester peptide corresponding to CI2(1−39)-COSR was assembled by SPPS. During the chain assembly, four α-hydroxy acids were incorporated: (*S*)-lactic acid (residues 16 and 22) and (*S*)-2-hydroxyisovaleric acid (residues 13 and 19). The amide couplings were carried out using standard in situ neutralization cycles for Boc SPPS using HBTU/DIEA.[24] The hydroxy acids were coupled using DIC/HOBt/NME without hydroxyl protection, and ester bonds were formed using DIC/DMAP in CH_2Cl_2.[10a] The resulting 4-ester CI2(1−39)-COSR peptide was purified by HPLC in good recovered yield (25%).

The full-length CI2 polypeptide was assembled by conformationally assisted ligation[25] of the 4-ester CI2(1−39)-COSR peptide with a peptide corresponding to CI2(40−64). This strategy takes advantage of the observation that CI2 fragments self-associate to form a nativelike tertiary structure. We have previously shown that combining the all-amide CI2 peptides 1−39-COSR and 40−62 (N-terminal Met) in a nondenaturing aqueous buffer (0.1 M sodium phosphate, pH 6.3) results in efficient ligation without the need for the N-terminal cysteine residue used in native chemical ligation approaches. As shown in Figure 2, the conformationally assisted ligation of the 4-ester CI2 proceeded to completion in less than 3 h. The efficient ligation of the 4-ester peptide fragments provides strong evidence that the pairing of the four ester bonds is tolerated in the protein fold since the fragments must fold in order to react.

Characterization. The 4-ester CI2 shows a mass of 7248 Da, 4 Da higher than the all-amide protein, indicating successful incorporation of the four α-hydroxy acid residues (Figure 2). To further characterize the folded structure, the 4-ester CI2 protein analogue was tested for functional activity as a tight-binding inhibitor of subtilisin.[26] The 4-ester CI2 inhibited subtilisin with a K_i of 2.1 nM compared to 0.25 nM for the all-amide CI2 under identical conditions. Importantly, the 4-ester

(19) Honig, B.; Yang, A. S. *Adv. Protein Chem.* **1995**, *46*, 27−58.

(20) An early example of this approach was demonstrated with polymers of $(Ala\text{-}Ala\text{-}Lac)_n$ that formed α-helical structures in aprotic solvents as characterized by CD. Wouters, G.; Katakai, R.; Becktel, W. J.; Goodman, M. *Macromolecules* **1982**, *15*, 31−35.

(21) This interaction should be somewhat destabilized from repulsion of the electronegative C═**O** and -**O**- groups.

(22) The CI2 sequence in this study differs slightly from that used by Fersht[8] (M1L, Q22A, D39T). These modifications have been shown to have little effect on the thermodynamic and kinetic properties of the protein.[8]

(23) Medel, D.; Cornish, V. W.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 435−462.

(24) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.

(25) Beligere, G. S.; Dawson, P. E. *J. Am. Chem. Soc.* **1999**, *121*, 6332−6333.

(26) Longstaff C.; Campbell A. F.; Fersht A. R. *Biochemistry*. **1990**, *29*, 7339−7347.

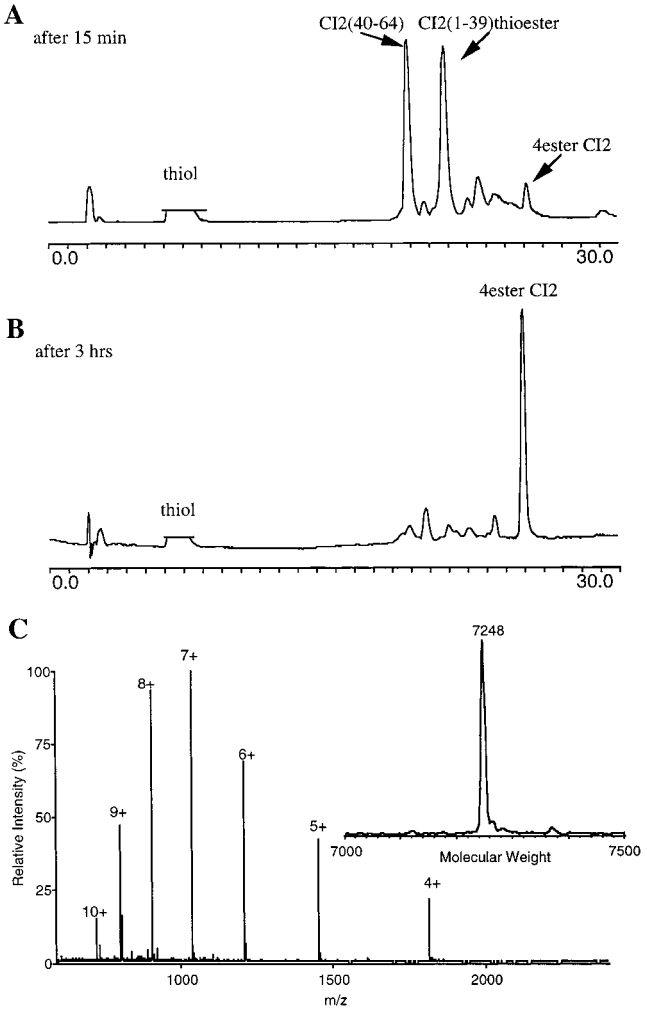

Figure 2. (A) HPLC showing the ligation after 15 min. (B) HPLC after 3 h showing completion of ligation. (C) Mass spectrum of purified 4-ester CI2 and reconstruct.

analogue was fully resistant to proteolysis over 24 h. The stable inhibition of subtilisin by CI2 is highly dependent on the folded structure since the binding loop is a good substrate when removed from the context of the folded protein.[27] These observations are consistent with a nativelike tertiary structure for 4-ester CI2.

The stability of the CI2 analogues was determined by monitoring the intrinsic fluorescence of the single Trp residue at position 5 as a function of Gdn·HCl concentration. Similar to that of the all-amide protein, fluorescence (E_x 280 nm; E_m 356 nm) of the 4-ester CI2 is quenched by 6-fold upon folding. As shown in Figure 3, the 4-ester CI2-folds cooperatively and has a denaturation profile similar to that of the all-amide CI2. The 4-ester CI2 analogue was destabilized by 2.93 kcal/mol, corresponding to ~1 kcal/mol for each hydrogen bond substitution. The sensitivity of a protein to denaturant value, ΔG/[Gdn·HCl] is believed to be proportional to the amount of hydrophobic surface area exposed upon denaturation. The *m*-value of 4-ester CI2 increased from 1.83 ± 0.02 to 2.00 ± 0.03, consistent with the burial of a relatively nonpolar ester backbone instead of the polar amide backbone.[28] (Table 1)

The ability of the 4-ester CI2 to fold correctly is interesting in regards to current models of protein folding. 4-ester CI2

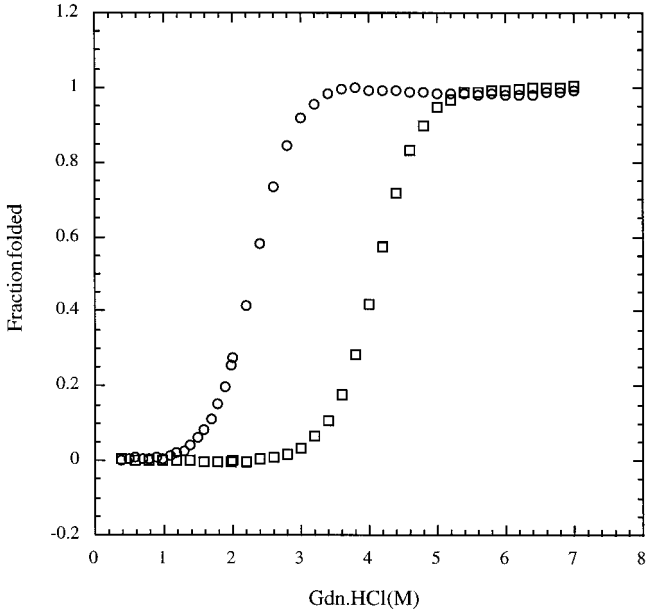

Figure 3. Guanidine denaturation of all-amide (□) and 4-ester (○) CI2, in 0.1 M phosphate pH 6.3 (intrinsic fluorescence was monitored at 356 nm with excitation wavelength of 280 nm).

Table 1: Equilibrium Guanidine Denaturation

mutant	m	$D_{50\%}$	$\Delta G°$	$\Delta\Delta G°$
all amide	1.83 ± 0.02	4.09 ± 0.004	7.51 ± 0.08	--
4-ester	2.00 ± 0.03	2.25 ± 0.008	4.58 ± 0.09	2.93

cannot form more than two sequentially consecutive hydrogen bonds. Consequently, a nucleation structure with multiple consecutive hydrogen bonds is not required for the folding of this CI2 analogue. Formation of the α-helix in CI2 is thought to occur during the rate-determining step of folding. Analysis of 4-ester CI2 using time-resolved methods will give insight into the kinetics of forming a hydrogen-bonding deficient α-helix and whether this analogue folds by the same mechanism as CI2.

Conclusions

We have demonstrated that multiple ester bonds can be incorporated into a polypeptide by total chemical synthesis to produce a folded, functional protein. These results suggest that backbone hydrogen bonding is not a required element of α-helical structures in proteins. Recent studies of foldamers[29] consisting of non-α-amino acids polymers have demonstrated that these molecules can form helical structures either with or without hydrogen bonds. This raises the intriguing possibility that an entire α-helix of a protein could be replaced with α-hydroxy acids or that a polyester sequence could be designed that would fold into a defined three-dimensional structure without hydrogen bonds.

Experimental Methods

Boc-amino acids were obtained from Midwest Biotech (Fishers, IN), α-hydroxy isovaleric acid, was obtained from Bachem Bioscience (King of Prussia, PA). L-(+)-Lactic acid was obtained from Fluka Chemie. *S*-Trityl-mercaptopropionic acid was obtained from Peptides International (Louisville, KY). 2-(1*H*-Benzotriazol-1-yl)-1,1,3,3-tetrameth-

(27) Lu, W. Y.; Starovasnik, M. A.; Dwyer, J. J.; Kossiakoff, A. A.; Kent, S. B.; Lu, W. *Biochemistry* **2000**, *39*, 3575−3584

(28) Shortle, D. *FASEB J.* **1996**, *10*, 27−34.

(29) (a) Gellman, S. H. *Acc. Chem. Res.* **1998**, *31*, 173−180. (b) Kirshenbaum, K.; Zuckermann, R. N.; Dill, K A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 530−535. (c) Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G. *Science* **1997**, *277*, 1793−1796.

yluronium hexafluorophosphate (HBTU), Boc-Gly-OCH_2Pam-resin, and *N*,*N*,-diisopropylethylamine (DIEA) were obtained from Applied Biosystems (Foster City, CA). MBHA resin was obtained from Peninsula Laboratories (Belmont, CA). All solvents of high purity were purchased from Fisher. Trifluoroacetic acid (TFA) was obtained from Halocarbon (River Edge, NJ). HF was purchased from Matheson Gas (Cucamonga, CA). Succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide and subtilisin were obtained from Sigma-Aldrich, Inc.

Peptide Synthesis. Peptides were prepared by manual solid-phase peptide synthesis (SPPS) typically on a 0.4 mmol scale using the in situ neutralization/HBTU activation procedure for

Boc chemistry as previously described. The peptide coupling was carried out with 5-fold excess (2.2 mmol) of activated amino acid for a minimum of 15 min. The N-terminal thioester peptide was synthesized on a TAMPAL resin as described.[30]

Ester Coupling. The coupling of the (*S*)-lactic acid and (*S*)-2-hydroxyisovaleric acid residues was carried out using a DIC/HOBt activation method.[10a] Lactic acid (2.2 mmol) in 4 mL of 50% DCM/DMF was activated with DIC (2.0 mmol) in the presence of HOBT-(2.4 mmol) at 0 °C for 15 min. The mixture was added to the resin, along with NEM (0.8 mmol), and coupled for 10 min, room temperature. A second coupling was carried out only if the coupling was <98%. The ester bond was formed, using DIC/DMAP activation.[10a] The next amino acid (2.2 mmol) was taken in 4 mL of 50% DCM/DMF and activated with DIC (2.0 mmol) for 15 min at 0 °C. The mixture was added to resin followed by NEM (0.8 mmol) and <10 mg DMAP then allowed to couple for 1 h, room temperature. After the chain assembly was completed, the peptides were deprotected and cleaved from the resin by treatment with anhydrous HF for 1 h at 0 °C with 4% *p*-cresol as a scavenger. The resultant peptides were precipitated with ice-cold diethyl ether, dissolved in aqueous acetonitrile, and lyophilized.

High Performance Liquid Chromatography (HPLC). Analytical reversed-phase HPLC was performed on a Hewlett-Packard HPLC 1050 system using Vydac C-18 columns (5 μm, 0.46 × 15 cm). Semi-preparative reversed-phase HPLC was performed on a Rainin HPLC system using a Vydac C-18 column (10 μm, 1.0 × 25 cm). Linear gradients of A: H_2O (0.1% TFA) and B: acetonitrile:H_2O:TFA 900: 99:1 were used to elute bound peptides.

Mass Spectrometry. Electrospray ionization mass spectrometry (ESI-MS) was performed on an API-III triple quadrupole mass spectrometer (PE-Sciex). Peptide masses were calculated from the experimental mass-to-charge (*m*/*z*) ratios from all of the observed protonation states of a peptide using MacSpec software (Sciex). Theoretical masses of peptides and proteins were calculated using MacProMass software (Beckman Research Institute, Duarte, CA).

Conformationally Assisted Ligation. The self-associating peptides of CI2, CI2(1−39) [C-terminal thioester, 4.5 mg, 0.1 mmol] and CI2-(40−64) (3.0 mg, 0.1 mmol) were ligated under folding conditions using 0.1 M phosphate buffer pH 6.3, and 2% (v/v) thiophenol to obtain the full length ligated peptide in less than 3 h (3.5 mg, 58% yield). The ligation was performed at room temperature with periodic vortexing. The final pH of the reaction was ∼5 (the pH of the solution was measured with pH indicator strips from colorpHast pH 4−7). The progress and completion of the ligation reaction was monitored by HPLC and ESI-MS.

Guanidine Denaturation. Initially, chemical denaturation experiments of CI2 and 4-ester (1−64) (0.5 g) was dissolved in 1 mL of phosphate buffer to obtain the folded protein (50 μM calculated using absorbance at 282 nm, $\epsilon = 6950$). For each data point 50 μL of protein solution was diluted into 450 μL of appropriate guanidine solutions to obtain 5 μM final protein concentration. The samples were incubated for at least 4 h, and the fluorescence emission was measured at 356 nm, upon excitation at 280 nm.

Subsequently, the denaturation experiments were repeated on an AVIV ATF 105 automated titrating spectrofluorometer using an equilibrium time of 5 min for every addition, using samples of 0 M guanidine and 8 M guanidine in 0.1 M phosphate pH 6.3.

The free energy of protein folding was calculated as described[31] using eq 1

$$\Delta G = RT \ln(f_D/f_N) \quad (1)$$

where R is the rate constant, T is the temperature, f_D is fraction unfolded, and f_N is fraction folded.

The data were analyzed by fitting to the eq 2 using nonlinear regression to obtain ΔG_{U-F}^{H2O} and eq 3 to obtain $D_{50\%}$ values.

$$F = [(\alpha_F + \beta_F[D]) + ((\alpha_U + \beta_U[D]) \exp(m[D] - \Delta G_{U-F}^{H2O}))/RT]/\{1 + \exp(m[D] - \Delta G_{U-F}^{H2O})/RT\} \quad (2)$$

$$F = [(\alpha_F + \beta_F[D]) + ((\alpha_U + \beta_U[D]) \exp(m([D] - [D]_{50\%})))/RT]/\{1 + \exp(m[D] - [D]_{50\%}))/RT\} \quad (3)$$

Data from both denaturation procedures agreed within experimental error, verifying that the 5 min incubation time used in the titration experiments was sufficient to reach equilibrium. Table 1 reflects the data derived from the titration experiments.

Chymotrypsin Inhibition Activity Assay. The assay was performed using subtilisin Carlsberg enzyme using OPTI*max* microtiterplate reader from Molecular Devices. The substrate used was succinyl-Ala-Ala-Pro-Phe-*p*-NA, and the enzyme concentration was calculated from initial rates of substrate hydrolysis. The reactions were carried out in 0.1 M Tris-HCl buffer, pH 8.4, with an enzyme concentration of 0.25 nM and substrate concentration of 1 mM in a volume of 0.1 mL with varying inhibitor concentrations. The reaction was followed at 412 nm, monitoring the substrate hydrolysis. K_i was determined as described by Fersht et al.[26] using eq 4. There was no reversal of inhibition for of the mutants over 24 h.

$$(v_o - v_s)/v_s = [I]/[K_i(1 + [S]/K_m)] \quad (4)$$

The ability of CI2 and 4-ester CI2 to inhibit subtilisin as a tight binding inhibitor was assayed by preparing an enzyme−inhibitor complex (1:2 10^{-7}M) in Tris buffer pH 8.5. At different time intervals (1, 2, 4, 8, and 24 h) the substrate (100 μL, 25 mg/mL) Suc-Ala-Ala-Pro-Phe-pNA, was added to 2 mL of complex and checked for hydrolysis of the substrate. There was no observable substrate hydrolysis even after 24 h.

Acknowledgment. This work was supported by NIH GM 59380 and The Skaggs Institute for Chemical Biology. We thank Dr. Murray Goodman for helpful discussions and Xin Jiang and Dr. Jeff Kelly for use of the titration spectrofluorometer.

Supporting Information Available: Details for TMPAL resin, ESI-MS data, and fluorescence spectra (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA001648E

(30) Hackeng, T. M.; Griffin, J.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068−10073.

(31) Jackson, S.E.; Moracci, M.; Elmasry, N.; Johnson, C.M.; Fersht, A.R. *Biochemistry* **1993**, *32*, 11259−11269.